ANNALS

OF THE

UNIVERSITY OF BUCHAREST

Philosophy Series

2 0 2 2

Number I

SUMAR • SOMMAIRE • CONTENTS

TRANSHUMANIST FUTURE TRENDS AND ETHICAL CHALLENEGES: A WORLD DEPICTED BY CYBORGS, BIO-ENHANCEMENT AND GENE TECHNOLOGIES

PHILIP LARREY, DANIELE PULEIO, Transhumanism in a Cosmic Perspective ... 3
JASON M. PITTMAN, I Think, Therefore We Are ... 29
ŽARKO PAIĆ, *Homo Kybernetes*. From Onto-Genesis to the Posthuman Condition ... 45
FRANC MALI, Human Genome Editing Technologies in the Center of Current Transhumanistic Discourse on Human Enhancement ... 69
KEVIN RÄNDI, Cyborg Networks: Freedom Through Data with Open Source and Implantable Devices ... 87
MARIUS MARKUCKAS, Sex and Transhumanism ... 111
ANDREI NUȚAȘ, The Borg – Glimpsing into a Nietzschean Transhumanist Utopia ... 129
NICOLÁS ANTONIO ROJAS CORTÉS, Twisting Transhumanism: some Scope and Limits of Sorgner's Philosophy ... 141
RADU USZKAI, EMANUEL SOCACIU, MIHAIL-VALENTIN CERNEA, Paladins with Muscle Cramps: Immersion, MMORPG's, and Nozick's Experience Machine ... 159
LORIN GHIMAN, What Society for Our Posthuman Future? A Task for the Serious Transhumanist Thinker ... 185

REVIEW

NATASHA BERANEK, Transhumanism Gets an Upgrade 199

Notes for Contributors 207

TRANSHUMANISM IN A COSMIC PERSPECTIVE

PHILIP LARREY[1], DANIELE PULEIO[2]

Abstract

This paper presents transhumanism as a post-biological philosophy that can provide a theoretical framework for scientific programmes oriented towards the search for intelligent life in the universe. The long-term thinking inherent in transhumanism synchronizes evolution and technology, heralding the genesis of post-biological entities. The first part of the paper assesses the applicability of this scheme to possible alien civilizations older than our own. The second part considers the most intriguing fictional explorations of mind uploading as a possible resolution of the Fermi's Paradox.

Keywords: Transhumanism, SETI, post-biological evolution, technological civilization, mind uploading, simulation hypothesis, Fermi's Paradox, virtual escapism

1.On the Post-Biological Fate of The Universe: From Techno-Sapiens to Techno-Alien

In the short story *The Sentinel* (1948) Arthur C. Clarke (1917-2008) describes human exploration of the lunar soil. During one of the missions to *Mare Crisium* the expedition's selenologist comes across a mysterious mechanical sparkle in the distance. Tracing the source, he discovers a small crystal artefact with a pyramid structure, protected by an energy membrane. The operating mechanism demands an intelligence beyond the human technological horizon: the pyramid turns out to be a beacon,

[1] Full Professor of Logic and Epistemology, Pontifical Lateran University.
Email: <larrey@pul.va>.
[2] Faculty of Philosophy, Pontifical Lateran University.
Email: <daniele.puleio@gmail.com>.

a sentinel – one of many scattered throughout the universe – built by a very ancient alien civilization to watch over all those worlds that, *in nuce*, promised the appearance of intelligent life, including the local portion of the Milky Way. How, however, to justify the sentinel's presence on the Moon and not directly on Earth?

> "They would be interested in our civilization only if we proved our fitness to survive – by crossing space and so escaping from the Earth, our cradle. That is the challenge that all intelligent races must meet, sooner or later. It is a double challenge, for it depends in turn upon the conquest of atomic energy and the last choice between life and death. Once we had passed that crisis, it was only a matter of time before we found the pyramid and forced it open. Now its signals have ceased, and those whose duty it is will be turning their minds upon Earth. Perhaps they wish to help our infant civilization." (Clarke 2001, 308)[3]

The protagonist's final consideration brings to light the nature of the cosmic census operated by the alien civilization, far from being a mere catalogue of the *scala naturae*. A biogenesis, the primitive presence of life (*e.g.*, microbial life), and not even *simpliciter* intelligence – rudimentary and pre-technological – both preordained mostly for mere self-preservation, was not enough: what was needed was a noogenesis[4], a technologically mature intelligence, capable of undertaking the challenges of space travel, manipulation of atomic energy, and interstellar communication. The narrative, in this way, concerns us closely: it emphasizes the mutual connection between extraterrestrial intelligence (ETI) research and the origin and future of human intelligence[5]. On closer inspection, in fact, the premise of the SETI (*Search for Extra-Terrestrial Intelligence*) research program is the same as that described in the story (but reversed in perspective): the recognition of an intelligence related to high technology,

---

[3] The story forms the basis from which Arthur C. Clarke will derive the novel *2001: A Space Odyssey*. The theme of humans in transition is present in another famous work by Clarke, whose title – *Childhood's End* (1953) – refers to a question of an eschatological nature: is humanity at the twilight of its 'infancy' or is it already the pinnacle of evolution?

[4] This term is understood as the origin of intelligence and technological civilization. Ćirković 2009, 3.

[5] The prodromes of this link can be glimpsed as early as 1971, the year of the Soviet-American scientific conference held near Buryakan. See Sagan 1973, 3; Sagan 1975, 196 ff.

capable of building radio telescopes or other astro-engineering artefacts[6] (revealing at least a level of science equal to ours), active in communication over long distances (radio transmission and reception of signals). In the words of Paul Davies, "we will recognize them by their instruments"[7].

The first SETI project can be traced back to the early days of radio astronomy, in 1959, the year the article appeared in the journal *Nature* written by physicists Giuseppe Cocconi and Philip Morrison, with the title *Searching for Interstellar Communication*[8]. The article argued for the possibility of accomplishing this kind of communication through radio waves and suggested searching at a wavelength of 21 cm., that is, tuning the receiver of radio telescopes to frequencies adjacent to the neutral hydrogen emission line (1420 MHz)[9]. The American astronomer Frank Drake (1930-2022) – who recently passed away – put the idea into practice by establishing the *Ozma project* in 1960 and, a year later, during the meeting of experts that took place at the observatory where he worked, in Green Bank, West Virginia, he proposed the eponymous equation

---

[6] ETI can be picked up by radio signals, but it is also possible to infer it through the discovery of anomalies around a planet, for instance indicative of the activity of a Kardashev Type II civilization. This is what M. Ćirković calls the *Dysonian approach to SETI* – a clear reference to the "Dyson spheres" – which emphasizes the search for technological manifestations. See Kardashev 1964, 217–20; Sagan 1975, 229 ff.; Ćirković 2009, 16).

[7] "[…] the principle of 'by their instruments ye shall know them'" (Davies 2010, 72).

[8] Cocconi & Morrison 1959, 844-846.

[9] "Mankind's first serious attempt to communicate with extraterrestrial civilizations occurred on March 3, 1972, with the launching of the *Pioneer 10* spacecraft from Cape Kennedy. […] The message is etched on a 6-inch by 9-inch gold-anodized aluminum plate, attached to the antenna support struts of *Pioneer 10*. […] The message itself intends to communicate the locale, epoch, and something of the nature of the builders of the spacecraft. It is written in the only language we share with the recipients: Science. At top left is a schematic representation of the hyperfine transition between parallel and antiparallel proton and electron spins of the neutral hydrogen atom. Beneath this representation is the binary number 1. […] Since hydrogen is the most abundant atom in the Galaxy, and physics is the same throughout the Galaxy, we think there will be no difficulty for an advanced civilization to understand this part of the message." (Sagan 1975, 17-19). As is well known, the golden plaque also contains the mapping of pulsar stars (*pulsating radio sources*), which, with their regular period radio emissions, act as cosmic clocks; and the stylized depiction of man and woman.

containing the probabilistic estimate of the number of extraterrestrial civilizations capable of establishing radio communications that we would expect in the Milky Way. From that meeting the SETI research program received the status of a scientific discipline[10]. The famous formula combines a long series of specific factors (believed to play a role in the development of such civilizations), which, multiplied together would produce *N*, the number of technological civilizations potentially in communication in our galaxy. In this sense, more than an equation, it is an organized quantification of our ignorance in the search for extraterrestrial intelligence[11]. Drake's approach is usually written in the following terms[12]:

$$N = R^* \bullet f_p \bullet n_e \bullet f_l \bullet f_i \bullet f_c \bullet L$$

Where R^* denotes the rate – the rapidity (rate) – of formation of Sun-like stars suitable for the development of intelligent life; f_p denotes the fraction of those stars endowed with planets (so-called exoplanets); n_e is the number of habitable planets in a solar system that has planets, those capable of harboring life forms; f_l refers to the fraction of biocompatible planets on which life actually appears; f_i expresses the fraction of life-hosting planets in which intelligent life emerges; f_c corresponds to the fraction of planets in which intelligent life develops a technological civilization and, consequently, the development of technology suitable for interstellar communications. This is the factor that mainly interested the alien guardians in Clarke's story. Finally,

---

[10] Since then, many projects have been carried out to intensify the research program. From the Arecibo radio telescope, decommissioned in 2016; to the "shared computing" project between private individuals around the world SETI@home; to contemporary projects in collaboration with other states: ATA (*Allen Telescope Array*), the Chinese single-parabolic FAST (*Five Hundred Metre Aperture Spherical Telescope*) radio-telescope and the international SKA (*Square Kilometer Array*) observatory. For a historical survey of the project see <https://www.seti.org/>; see Sagan & Drake 1975, 80–89.

[11] Rather than an estimate, one should speak of conjecture, since the only available example of a biocompatible planet in which life has emerged is Earth. A single sample containing only one element makes any scientific account difficult.

[12] See Drake 2013, 173-176; Dick 2013, 31-32.

L should answer the question "how long do technological civilizations last?" and indicate the average life span of a civilization capable of communication. This last factor – aside from embodying the cultural element in the formula[13] – is the most controversial and dominates over all others: the greater the assumed average life span of a technological civilization, the more civilizations will populate the galaxy and, consequently, the higher the probability of intercepting its signals[14].

Carl Sagan (1934-1996) – a pioneer in astrobiology – popularized the SETI project – including through the novel *Contact* (1985), whose protagonist, Ellie Arroway, is loosely inspired by the astronomer and former SETI director Jill Cornell Tarter. The scientific enterprise of SETI and astrobiology find their justificatory rationale in the Copernican cosmological assumption – repeatedly invoked by Sagan himself, so much so that he called it the "principle of mediocrity" (Sagan 1994, 372) – which argues in favor of the non-specialty of human existence[15]. Our place in the universe is not privileged. This would suggest, first, that the state of evolution on Earth is framed within a galactic average[16], in a context

---

[13] According to Steven J. Dick, the Drake equation incorporates astronomical (the first two factors); biological (third and fourth factors); and social-cultural (the last two) elements. See Dick 2003, 67.

[14] It should be noted that the ability to transmit signals and the possibility of self-destruction through nuclear weapons appeared in human history at the same time. Newman and Sagan suggest a social control of technologies to avoid nuclear self-destruction. During the Cold War, with the constant danger of a nuclear conflict, Sagan decided to estimate the L-factor with a low value. See Sagan & Newman 1983, 113-121; Shostak 2009, 399-414.

[15] "[…] We humans have emerged from a common evolutionary process with all the other plants and animals on Earth. We do not possess any uniquely valid locale, epoch, velocity, acceleration, or means of measuring space and time" (Sagan & Newman 1983, 113). According to the cosmological principle, derived from the Copernican principle, the universe is homogeneous (the general properties of the universe appear the same for any observer at any cosmic epoch) and isotropic (equal in all direction) on a large scale. On the epistemological value of the principle see Fano & Macchia 2020, 309-324.

[16] "Since our evolution to technological ability occurred within five billion years after the formation of our planet, it follows from this principle that other technological species would typically evolve in a similar time period" (Barrow & Tipler 1986, 586-587).

where life is a common phenomenon in the universe[17]; secondly, under this assumption, we should expect that the radio-transmitting civilizations sought by SETI are, for the most part, extremely older than our own[18].

Astronomer and philosopher M. Ćirković argues that astrobiology and ETI research are of paramount importance to transhumanist thought, as they form its scientific basis (Ćirković 2003, 2). The opposite is equally true: transhumanism can be considered preparatory to any ETI research, providing a philosophical background in which to place it. The heuristic value that the hermeneutic framework (Larrey & Puleio 2021, 14) of transhumanism suggests to the SETI research program should not be underestimated, since it embodies a long-term, large-scale thinking methodology that educates us to cultivate a cosmic horizon for the future evolution of intelligence in the universe. Transhumanism has as its perspective, cultural evolution and its technological aptitude on a cosmic scale, a thought neglected by astrobiology and SETI, intent on focusing more on the biological element. In the absence of a theory of cultural evolution[19], astronomer and historian of science Steven J. Dick proposes to acquire, in line with the theoretical framework of transhumanism, a "Stapeldonian" attitude, referring to the sci-fi epics narrated by Olaf Stapledon (1886-1950), the first, along with H. W. Wells, to preconceive a cosmic perspective for biological and cultural evolution[20]. Novels such as *Last and First Men* (1930) and *Star Maker* (1937) anticipate the possible nature of a post-biological universe, in which cultural

---

[17] If the Copernican principle invites one to consider all natural phenomena in their context of cosmic vulgarity; the alleged uniqueness of the human phenomenon, on the contrary, would reinforce an anthropic view of the universe.

[18] "Thus, all lines of evidence converge on the conclusion that the maximum age of extraterrestrial intelligence would be billions of years, specifically, A ranges from 1.7 billion to 8 billion years. Even uncertainties of a billion years would not affect the argument for taking seriously cultural evolution" (Dick 2009, 468.) This is not the same as claiming that all life evolves to intelligent and technological stages. Whether intelligent life, given an extremely long time, inevitably encounters science is debated.

[19] "[…] the problem of cultural evolution among extraterrestrials (and partially explaining why it has not been taken seriously before) is our lack of knowledge about the mechanism and direction of cultural evolution on Earth" (Dick 2009, 579).

[20] Today we can find this perspective in the novels of Greg Egan (*Diaspora*, 1997) and Charles Stross (*Accelerando*, 2005; *Singularity Sky*, 2003), for example.

evolution plays a key role and which "is one in which the majority of intelligent life has evolved beyond flesh and blood intelligence, in proportion to its longevity, *L*"[21].

Put in these terms, transhumanism is not only a philosophy of self-directed evolution; it is also a philosophy of the post-biological. If, in fact, we adhere to the first formulation proposed by Julian Huxley (1887-1975), we learn that the emblematic figure of the movement is not the unidirectional quest for immortality – which, moreover, is to be understood only in a metaphorical sense, given the underlying naturalist ontology shared by much of the movement (Sorgner 2022, 28) – but rather the attempt to transcend one's biological roots, the organic limits of the body, a burden to be technologically amended (regenerative medicine, "cyborgization", mind uploading)[22] because it is subject to entropic decay, disease and, ultimately, death.

> "Perhaps the term transhumanism will fit: man remaining human, but transcending himself, realizing the new potentials of his human nature, for his human nature" (Huxley 1957, 17).

Emphasis is placed on the verticality of the bios, on its "upward thrust"[23]. For this reason, the terminological meaning of "*trans-organic*" would be more appropriate, insofar as it accommodates a very broad definition of bios, detached from the biological substratum.

Therefore, can the reflections on the future of the human species inaugurated by transhumanism be extended to possible technologically advanced alien civilizations? If the human being is not the apex of the evolutionary scale, but only a transitional stage destined to be passed over until the "singularity" is reached, what about the possible existence of civilizations billions of years older than ours? Is it plausible that they are already in a post-Singularite and, therefore, post-biological condition?

---

[21] Dick outlines three scientific premises for a post-biological universe: "(1) the maximum age of ETI is several billion years; (2) the lifetime (L) of a technological civilization is >100 years and probably much larger; and (3) in the long term cultural evolution supersedes biological evolution, and would have produced something far beyond biological intelligence" (Dick 2009, 66).

[22] For this tripartition see Larrey & Puleio 2021, 17-20; Waters 2011, 166 ff.

[23] "[...] Interpreting evolution as an upward spiral, as a movement that is both forward and upward" (Campa 2019, 20).

Speculative fiction has accustomed us to thinking of the inhabitants of other worlds as a biological puzzle: "little green men" with thinned limbs and a highly developed brain. This is a biological version of the alien, whose technological element, while highly refined, remained foreign to exobiology and was confined to engineering instrumentation, such as spaceships. It is reasonable to expect, once the mystery of the origin of life is unraveled, that most advanced alien civilizations would be in a form of artificial superintelligence far removed from their original biological version[24]. Should there be contact, it would be with their emissaries and artificial vanguards. On this subject the literature is beginning to become important. For example, Hans Moravec – a forerunner in the field of robotics – in the prologue of *Mind Children* (1988) speaks of "children of the mind", as those who will have transcended the limiting biology of their "parents" and inherit the Earth, in the sense in which Marvin Minsky understood it in the essay *Will Robots Inherit the Earth?* (1994):

> "What awaits is not oblivion but rather a future which, from our present vantage point, is best described by the words 'post-biological' or even 'supernatural'. It is a world in which the human race has been swept away by the tide of cultural change, usurped by its own artificial progeny. […] But within the next century they will mature into entities as complex as ourselves, and eventually into something transcending everything we know—in whom we can take pride when they refer to themselves as our descendants. Unleashed from the plodding pace of biological evolution, the children of our minds will be free to grow to confront immense and fundamental challenges in the larger universe" (Moravec, 1988,1).

Then again, the fifth-dimensional entities that build wormholes and manipulate gravity in *Interstellar* (Christopher Nolan 2014) are referred to as "Them" throughout the film, except to discover at the end that they were the humanity of the future evolved beyond four dimensions. Our evolutionary descendants, in short. The implications for SETI research would be disruptive, starting with the nature of what we are looking for. There is no doubt that a key role will be played by the *GNR* (Genetics, Nanotechnology, Robotic) *revolution*[25].

---

[24] Davies calls such post-biological entities ATS (*Auto-Teleological Super-Systems*). (Davies 2010, 161)

[25] On the potential of gene or cyborgs technologies as a realistic and more mature option for the future well-being of humanity than mind uploading see Sorgner 2002, 22-29. It is often forgotten that one of the matrices of transhumanism – which, like all movements,

Philosopher of mind Susan Schneider also comes to the same conclusion:

> "[...] the technological developments we are witnessing today on Earth may have all happened before, elsewhere in the universe. That is, the universe's greatest intelligences may be synthetic, having grown out of civilizations that were once biological. The transition from biological to synthetic intelligence may be a general pattern, instantiated over and over, throughout the cosmos. If a civilization develops the requisite AI technology, and cultural conditions are favorable, the transition from biological to post-biological may take only a few hundred years." (Schneider 2019, 98)

Such a pattern could be a universal attractor – an ultimate destiny of development – to which all civilizations with techno-scientific knowledge go. Is it possible that life, although germinating in different places and contexts, arrives at the same outcome, tracing the phenomenon of evolutionary convergence[26] in nature?

Let us try to identify an outline of the development of these possible civilizations through three taxonomies.

2.Eras, Lifetimes, Civilizations

We have seen that the hypothesis underlying SETI estimates that billions of civilizations exist in our light cone of universe, including thousands or millions in our galaxy alone. As Kurzweil notes, "[t]hese civilizations must be at different stages over billions of years of development. Some would be behind us, and some would be ahead" (Kurzweil 2005, 348).

---

is not unitary – both in the context of space exploration. It is when faced with the possibility of probing places unfriendly (such as deep space) to human metabolism that reflection on *cyborgs* is born, man's ultimate symbiosis with A.I., which opens up the possibility of "deliberately incorporating exogenous components that extend the organism's function of control and self-regulation in order to adapt to new environments" (Clynes & Kline 1960, p. 27). The cyborg – short for "cybernetic organism" – is the physical (and symbolic) representation of man's taking leave of his organic structure as the exclusive support for his survival: it is a man who has undertaken a metamorphosis into a machine perspective, carried out through bionics, robotics and nanotechnology.

[26] *e.g.* the wings of birds, bats and insects perform the same function, even though they belong to distant phylogenetic lines.

Drawing on the nineteenth-century social positivism of Nicolas de Condorcet, Henri de Saint-Simon and August Comte, he interprets history as a progressive succession of paradigms, an evolutionary pattern – "Evolution is a process of creating patterns of increasing order" (Kurzweil 2016, 153) – that is divided diachronically into "Six Eras", where each uses the information-processing methods of the phase that preceded it to structure the next era.

- the first Era (*Physics and Chemistry*) is that which represents the information in the fundamental structures (of the universe) of matter and energy, from which atomic structures and, later, molecular aggregates were formed;
- the second Era (*Biology and DNA*) is marked by the presence of organic life and biological systems that have developed over time an information-preserving mechanism: DNA;
- the third Era (*Brains*), under evolution driven by DNA, sees the formation of complex structures capable of processing information and creating abstract models of the world;
- the fourth Era (*Technology*) is characterized by the advent of technology and its catalytic function with regard to biological evolution;
- the fifth Era will see the fusion of human intelligence and technology achieved;
- the sixth and final Era will be that of the awakening of the universe: "In the aftermath of the Singularity, intelligence, derived from its biological origins in human brains and its technological origins in human ingenuity, will begin to saturate the matter and energy in its midst" (Kurzweil 2016, 159).

It is in the penultimate epoch that a new evolutionary syntax emerges. Transhumanist literature is wont to point to the apogee of this process in a narratively evocative event – known as the "Technological Singularity"[27] – which, according to Kurzweil himself will occur in 2045 and will constitute:

[27] The term is first used by Vernor Vinge – a sci-fi writer and mathematics professor – in the novel *Marooned in Realtime* (1986). However, the "official" technical meaning will come a few years later (1993), at a symposium organized by Nasa. See Vinge 1993, 11-21. "[...] Vinge thinks that a singularity is likely as a consequence of advances in artificial

> "[…] a future period during which the pace of technological change will be so rapid, its impact so deep, that human life will be irreversibly transformed. Although neither utopian nor dystopian, this epoch will transform the concepts that we rely on to give meaning to our lives, from our business models to the cycle of human life, including death itself. Understanding the Singularity will alter our perspective on the significance of our past and the ramifications for our future. […] It will represent the culmination of the merger of our biological thinking and existence with our technology, resulting in a world that is still human but that transcends our biological roots. There will be no distinction, post-Singularity, between human and machine or between physical and virtual reality" (Kurzweil 2005, 7-9)[28].

According to Kurzweil's model, therefore, once a species reaches the fourth era, it would only be a matter of a few centuries before it reaches the great intelligence explosion of the Singularity. Billions of civilizations, according to SETI estimates, should lie beyond the fourth and fifth eras.

Another interesting scheme – dependent on the hardware/software dialectic dear to computational functionalism – is that of Max Tegmark, professor at the Massachusetts Institute of Technology and co-founder of the *Future of Life Institute*. In an attempt to lucidly scan the stages of this progression, he outlines a taxonomy of life – defined as "a process that can preserve its complexity and replicate itself" (Tegmark 2017, 25) – that unfolds in three evolutionary stages: in the biological stage, life's hardware and software are determined by its DNA, and change only through evolution over many generations (Life 1.0); in the cultural

---

intelligence, large networked computer systems, computer-human integration, or some other form of intelligence amplification" (Bostrom 2003, 19) (Available at <https://www.nick bostrom.com/ views/transhumanist.pdf> last time accessed in 4 December 2022).

[28] The singularity event is presented by resorting to the cosmological metaphor of the black hole, whose gravitational attraction is so intense that it alters all matter/energy structures that accelerate, with no possibility of escape, toward the event horizon. Moreover, it has been keenly observed how the matrix of the current scientist wing (Kurzweil, More, Bostrom *et al*.) of the movement is purely religious-eschatological in nature, where the concept of technological singularity itself would be nothing more than the secularization of Jesuit Teilhard de Chardin's *Omega Point*, the finalistic convergence of the Noosphere (sphere of human thought implemented by technology that permeates, enveloping it like a unitary membrane, the geosphere) with the historical-evolutionary apogee of the universe. In Teilhard there is thus a *Parousia* on a cosmological scale. See Paura 2019, 348 ff.

stage, software is subject only to human design, *i.e.*, culture: humans can redesign their software through learning new complex skills, such as learning a language (Life 2.0); finally, in the technological stage, life becomes able to design its own hardware and software, without having to wait for it to gradually evolve over generations (Life 3.0) (Tegmark 2017, 27-29). Thus, a more technologically mature civilization should definitely stick to versions after Life 2.0 (2.1,2.2, etc.) to approach Life 3.0.

The last taxonomy is the one most concerning a technological civilization. Soviet astronomer Nikolai Kardashev in 1964 published an article entitled *Transmission of information by extraterrestrial civilizations* (Kardashev 1964, 217-19) in which he proposed a criterion for classifying technological civilizations according to energy consumption. Specifically: a Type I civilization is a planetary civilization, capable of extrapolating all the energy resources of its home planet, about $4 * 10^{19}$ erg/sec. The technological level is close to that currently achieved on Earth; a Type II civilization is able to fully harness all the energy radiated by its star (about $4 * 10^{33}$ erg/sec). How? For example, through monumental technical designs, such as Dyson spheres or Penrose mechanisms for extracting energy from a star or a rotating black hole (Kerr metric) respectively. Finally, the Type III civilization would dominate the energy of its own galaxy (about $4* 10^{44}$ erg/sec).

We can conclude that a post-biological civilization lives in a Kurzwelian Era that straddles the fourth and fifth; faces a Life 3.0; is definitely a Type II civilization not far from becoming Type III.

3. Can the Simulation Hypothesis Solve the Fermi Paradox?

3.1 Mind Uploading and Simulation Argument

The choice to place the topic of simulation in the section on mind uploading is due to the fact that the latter constitutes a presupposition, very often implicit, of the former (Sorgner 2022, 25-27).

The purpose of the more techno-euphoric (or techno-gnostic) wing of transhumanism is to detach from "primitive" organic life, to arrive at "substratum-independent minds" (Koene 2013, 146-156), and to instantiate

oneself, through the technique of mind uploading, into extra-biological structures – thus long-lived and/or otherwise replaceable (*e.g.*, silicon) – or into purely digital forms within a Gibsonian cyberspace. On the potential of consciousness data transfer, Kurzweil states the following:

> "Uploading a human brain means scanning all its salient details and then re-instantiating these details in a suitably powerful computational substrate. This process would capture a person's entire personality, memory, abilities and history. If we are truly capturing the mental processes of a particular person, then the re-instantiated mind will need a body, since much of our thinking is directed at physical needs and desires. [...] by the time we have the tools to capture and re-create a human brain with all its subtleties, we will have plenty of options for 21st century bodies for both non-biological humans and biological humans using extensions of our intelligence. The human body version 2.0 will include virtual bodies in fully realistic virtual environments, physical bodies based on nanotechnology, and more" (Kurzweil 2005, 198-199).

Bostrom, who seems more aware than Kurzweil about the inherent limitations in uploading, nevertheless lists its possible benefits for humanity in the future (Bostrom 2003, 18-19):

- we will no longer be subject to biological senescence;
- back-up copies of uploads will allow us to have a life expectancy potentially as long as that of the universe;
- economically it will be more advantageous than maintaining an organic ballast, such as the body, to feed, regenerate and transport;
- we will be able to travel at the speed of light as a pattern of information, which will be convenient for future large-scale space colonization;
- there would be a radical expansion of cognitive faculties, facilitated by the continuous implementations of uploads, which are more effective than the enhancements of a simple organic brain;
- it could serve patients undergoing cryopreservation.

Mind uploading in virtual environments[29], such as simulations, does not currently appear to be a realistic option on par with other

[29] Here we will focus exclusively on simulation as the specific *locus* of the broader concept of virtual reality of cyberspace (*i.e.*, cybernetic space), for which we refer to Silvano Tagliagambe's effective, if dated, definition: "a new universe, an artificial, 'virtual,' multidimensional reality, generated, fed and made accessible by computers

hypotheses (Sorgner 2022, 25); however, it has become the subject of intriguing fictional extrapolations. Under such hypotheses, simulations would take the form of a virtual Eden (there is also talk of "paradise engineering"[30]) capable of guaranteeing a digital eternity to the contents of consciousness poured into it: "this, then, is the physical mechanism of individual resurrection: we will be emulated in the computers of the distant future" (Tipler 1995, 211). Thus, one of the "new" planes of existence considered by transhumanists is the virtual – insofar as it simulates reality or even unleashes phenomenally impractical possibilities[31] – this is because the physiological functioning of the senses is replaced by cybernetic mechanisms that generate/stimulate computerized simulations of perceptions, producing refined multisensory experiences. It is precisely in the different way of understanding the relationship between perception and experience that resides one of the most prominent philosophical spin-offs in the field of virtual reality[32].

Mind uploading is a prodigal topic from speculative science fiction, cinematic and otherwise. More or less recent productions on the subject, such as *The Lawnmower Man* (Brett Leonard, 1992), *Virtuality* (Brett Leonard, 1995), *Strange Days* (Kathryn Bigelow, 1995), *Nirvana* (Gabriele Salvatores, 1997), *Existenz* (David Cronenberg, 1998), *The Matrix* (Wachowski Brothers, 1999), *The Thirteenth Floor* (Josef Rusnak, 1999), *S1mone* (Andrew Niccol, 2002), *Serenity* (Steven Knight, 2019), *Devs*[33]

---

through global communication networks" (Tagliagambe 1997, 39). Simulation, on the other hand, as a computer-mediated epistemological tool added to theory and experiment, is a kind of *tertium organon*, a worthy successor to the Aristotelian *organon* of deductive logic and the Baconian *Novum Organon* of induction. Cf. Longo & Vaccaro 2013, 29; Deutsch 1997, 90-111.

[30] See Pearce *The Hedonistic Imperative*, available at <https://www.hedweb.com/hedethic/tabconhi.htm.> last time accessed in 2 December 2022.

[31] Consider "Embodied Simulation", a model used within neuroscience.

[32] The oxymoron "virtual reality" emerges in full force should one analyze the outcomes of research on the *Second Life* game system, in which artificial avatars "live" a virtual existence parallel to that of their real physical counterparts see Gottschalk 2010, 501-525.

[33] Project *Devs* (to be read *Deus*) embodies the "quantum Laplace demon." The series explores – implicitly following the line of research traced by physicist David Deutsch, a staunch defender of Everett III's many-worlds interpretation of quantum mechanics – the potential of quantum computers (based on qubits) to retrospectively and predictively simulate the whole of reality, including realistic quantum systems, with considerable

(Alex Garland, 2020), *Bliss* (Mike Cahill, 2021), all move from the inescapable difficulty of the protagonists to discern reality from the sophisticated simulation in which they are immersed. The idea that we are *simulacra* resulting from a reality generated by a computer simulation is today one of the most suggestive intents to explain our universe from a scientific point of view, as well as a hypothesis harbinger of consequences as metaphysically bold as they are fascinating: it may not mean the absence of an extra-simulated reality; but only that the nature of the real – unlike what we used to think – is digital, rather than physical, in deference to the Wheelerian principle of the *it from bit*[34].

The film, *The Thirteenth Floor* – a transposition of the well-known novel *Simulacron-3* (D. F. Galouye, 1964) – is a *mise en abyme* of simulated planes, where the protagonists, creators of a perfect virtual replica of 1937 Los Angeles, discover that they themselves live in a simulation designed by a technologically advanced civilization of the future. The story's twist coincides with the conclusion that Bostrom's simulation argument reaches:

> "A technologically mature 'post-human' civilization would have enormous computing power. Based on this empirical fact, the simulation argument shows that at least one of the following propositions is true: (1) The fraction of human-level civilizations that reach a post-human stage is very close to zero; (2) The fraction of post-human civilizations that are interested in running simulations of ancestors is very close to zero; (3) The fraction of all people with our kind of experiences living in a simulation is very close to one" (Bostrom 2003, 255).

---

implications for determinism and free will (See Feynman 1982, 467-488; Deutsch 1985, 97-117).

[34] "Simulation is a true creation of alternative worlds, governed by formal rules (algorithms), the phenomenal richness of which is limited, in principle, only by the nature and potential complexity of these algorithms and the amount and precision of the data on which they operate. Many questions arise in this regard: what is the relationship between a phenomenon and its computer simulation? And, more generally, what is the relationship between our world and the world inside the computer? Can it be said that our world is itself a giant-sized computer? The last question suggests a metaphysical vision in which the first principle of reality is not the *archè* of the pre-Socratics, but information, or rather algorithmic computation [...]" (Longo & Vaccaro 2013, 159-160). Concerning the informational nature of reality and/or the possibility that the universe is a quantum-mechanical computer that computes itself see Steinhart 1998, 117-125; Wheeler 1989, 354-368; Davies 2010, 65-91; Lloyd 2006; Deutsch 2004, 90-102).

Through statistical inference, Bostrom concludes that it is highly likely that we are within a simulation. For: either the human species will never reach a technological maturity to perform simulations of past eras (and, even if it does, it will still prove recalcitrant or disinterested in performing them); or we almost certainly live in a simulation. In this third hypothesis, should a posthuman civilization decide to simulate past epochs, the "original" physical universe would host such a large number of simulations that it would end up, for example, with many more simulated people from the 21st century than there are in the actual 21st century. Such a hypothesis, among others, would justify the "fine tuning" of the physical constants of the universe (*e.g.*, the fine structure constant, the intensity of the gravitational force, the electron charge, etc.), which would be precisely "calibrated" by a demiurgic superintelligence in order to generate a warp of the cosmos suitable for our existence to emerge.

> "A mature superintelligence could create virtual worlds that appear to its inhabitants much the same as our world appears to us." (Bostrom 2003, 316)

The alternative to the proposed scenario, Bostrom argues, is to suppose that a "great filter" operates on the development of a technological civilization: once a certain threshold is reached, expansion would stop, with the consequence that Fermi's paradox would be explained by the fact that no intelligent species is able to go beyond a certain rate of technological expansion (Bostrom 2008, 72-77).

Bostrom's fear leads us to the *vexata quaestio* of ETI research, exemplified by the famous, seemingly innocuous question posed by the Italian physicist Enrico Fermi in the summer of 1950, during a lunch at the *Los Alamos National Laboratory*: *Where are they?*

3.2 Silentium Universi

If the probabilistic hydra expressed by Drake's formula, optimistically interpreted, conjectures the existence of billions of alien civilizations present in the cosmos (millions in the Milky Way), many of them with a more sophisticated level of technoscience than ours, how come they have not colonized or explored the galaxy (*e.g.*, through self-replicating von Neumann probes) or we have not yet come across

any trace of them? The Fermi's paradox is otherwise known as the Great Silence[35] (or *Silentium Universi*): our galaxy is supposed to be a cacophonous cluster of signals from the most disparate places in the cosmos (but still within our cone of light), yet it is surprisingly "silent."

The SETI and astrobiology research programs – which have the Copernican principle as a presupposition[36] – seem to suffer from anthropocentric biases, chief among them that intelligent life should emerge wherever earth-like physical environmental conditions are given, expressed by the conditional probability assembled by Drake and encoded by the hierarchical sequence:

star ⟶ biocompatible planet ⟶ life ⟶ intelligence ⟶ science.

We have no reason to expect that where life emerges, after a given period of time, science is inevitable. It is objected that making a prognostic judgement about other species from the model of our own – the only species we know of so far – is an unjustified generalization; an implicit assumption of anthropocentrism (Davies 2010, 31-32). Even for the theoretical framework postulated by transhumanism – at least as it is understood here – that is, the future genesis of post-biological entities. Yet, within reflections that unfold on such a time-dilated scale (millions and billions of years), human categories themselves may lose actual consistency. Conjectures and deductions, whether rigorous or not, about the evolution of other species, at the moment, can only be made from the only data available that is truly certain, namely, that there is a technological civilization in the universe: ours.

The problem involves the *an sit* of life in the context of cosmic evolution: is it a *unicum* (Monod), a highly probable event (Sagan) or even a cosmic imperative (De Duve)? (see Monod 1996; Sagan 1995, 1-4; De Duve 1995). This leads us to one of the most curious approaches to Fermi's Paradox, reformulated by economist Robin Hanson under the

---

[35] Like the title of a wonderful Ted Chiang story on the subject or Stanislaw Lem's classic *Il pianeta del silenzio* (1986). Variation on the theme of the Fermi's paradox is the recent film *Ad Astra* (James Gray, 2019).

[36] "The Copernican reasoning suggests that we should expect evolution to occur faster at some places than on Earth (and, of course, slower at other sites as well) – where are our interstellar siblings, then?" (Ćirković 2009, 9).

hypothesis known as the "Great Filter"[37], a term that designates a "statistical probability barrier" (Bostrom 2008, 72), a wedge of fundamental evolutionary transitions that must be traversed by a life form in order for the path from lifeless matter to lead to organisms capable of mastering the technology for the exploration and mass colonization of space. Even life on Earth has had to overcome multiple obstacles in its nearly four billion years history. This evolutionary/technological check-point imposes itself as a very harsh selective criterion, suggesting that biocompatible planets are in fact filtered out before they become a niche of intelligence.

A crucial question arises: at what point in our evolutionary journey is the Great Filter? Behind or in front of us? The reason why Bostrom hopes that no (micro)organisms will ever be discovered on Mars (or anywhere else in the Solar System) is that this would imply that life is not so rare in the cosmos (Bostrom 2008, 72). In this case the filter would be in front of us, and what would be unlikely would not be the appearance of life in the universe, but the fact that no intelligent species would be able to go beyond a certain stage of technological development, for instance because it self-destructs.

As we have already seen, this is the alternative considered by Bostrom to justify the absence of past simulations by technologically mature civilizations. The *L*-factor of the Drake equation would be extremely low: technological civilizations would not only be rare but also short-lived. For us, the Great Silence would be the prelude to imminent extinction. Conversely, should the filter be located behind us, it would mean the improbability of life being a widespread phenomenon in the galaxy, or even in the entire universe. Our existence as an intelligent species would be accidental, an exceptional combination of unlikely events bordering on the unbelievable. It is the "Rare Earth" hypothesis.[38]

While it is true that "absence of evidence is not evidence of absence" (Sagan 1997, 213), the inconsistency between estimates and

---

[37] See Hanson 1998, available at <http://hanson.gmu.edu/greatfilter.html>, last time accessed in 3 December 2022; Hanson's is a refinement of an argument originally proposed by Brandon Carter in the context of the anthropic principle. See Carter 1983, 347.

[38] See Ward & Brownlee 2000; Ćirković 2018, 147.

factual reality, between the abundance of Earth-like planets in the Milky Way and the absence of ET visitors apparently persists.

Stephen Webb, in an essay with the iconic title *If the Universe Is Teeming with Aliens ... Where Is Everybody?* lays out a series of possible solutions to Fermi's conundrum, some of which Ćirković also contemplates, grouping them into three categories: 1) ETIs do not exist; 2) They exist and are already here; 3) They exist, but we have not yet discovered them (Webb 2015)[39].

The hypothesis that they simply do not exist (Fermi, Tipler, etc.) or that they are beyond our observable universe[40] (not causally connected with us) is accompanied by the range of catastrophic scenarios, from self-destruction for nuclear reasons or due to potentially risky technologies beyond our control (bioterrorism, birth of super-intelligence with purposes not aligned with ours), to ecological holocausts, and natural hazards (super-volcanoes, asteroid impacts). Among the most curious scenarios are:

- *The zoo hypothesis*. They observe us but do not contact us, perhaps because of an ethical diktat, so as not to influence our evolution or because they are disinterested in us. Much of the speculation about UFOs – criticized by Sagan – is based on this scenario;
- An *intentional panspermia*, as the film *Prometheus* (Ridley Scott, 2012) teaches us from the opening scene: our life is the grafting of genetic material from an alien race. *Ergo*, we are of alien origin!
- We are inside a simulation generated by an alien super-intelligence, *Matrix*-style, with all the skepticism around the nature of reality that goes with it.
- There is incommensurability in communication (we cannot pick up their signals and/or vice-versa), in theoretical assumptions about the universe, in their use of different physics or mathematics unknown to us, or in their better understanding of how the cosmos works.

---

[39] See also Ćirković 2009, 7-13.

[40] As early as the Renaissance, Nicolas Cusanus (1401-1464) admitted the possibility of other worlds – which are themselves inhabited – but not accessible: "[...] since all that region of the stars is unknown to us, its inhabitants also remain completely unknown to us" (Cusano 2018, *De docta ignorantia II 12* § *171*, 201).

- We are within a time window in which the astro-biological clock of many civilizations is undergoing a major reset. We therefore live in the time horizon of their phase transition.

3.3 Virtual Escapism

Here we would like to briefly propose the thesis of *Virtual Escapism* (or Digital Diaspora)[41], a variant of the *Transcension hypothesis* originally formulated by John Smart. Both tread a ground that is in danger of collapsing into science fiction, but it is nevertheless worth exploring its fictional implications. The hypothesis of virtual escapism would have the merit of reconciling Fermi's paradox, the argument of simulation (of which mind uploading is a prerequisite) and the circumstance that the great filter is behind us (short duration of technological civilizations). Let us start in order. In brief, Smart's transcension[42] scenario involves

> "[…] that a universal process of evolutionary development guides all sufficiently advanced civilizations into what may be called "inner space", a computationally optimal domain of increasingly dense, productive, miniaturized, and efficient scales of space, time, energy, and matter, and eventually, to a black-hole-like destination." (Smart 2012, 55)

In other words, following a common trajectory of development (so-called *systemic teleology*), civilizations, once they have reached a critical threshold of technological expansion – presumably after the singularity –, mature the realization that transcending elsewhere – into the domain of inner Planck scale space, through the manipulation of Einstein-Rosen bridges and black holes – is a better choice than expanding. Indeed, Smart argues, the expansion hypothesis, as the standard expectation of SETI and Active SETI (or METI, *Messaging to Extra-Terrestrial Intelligence*) is impractical due to the physical constraint of the finite size of the speed of light. As a counterbalance to the Kardashev scale adopted by expansion theorists, Smart adopts the "Barrow-scale" where the reference metric is

---

[41] The reference is to Greg Egan's 1997 speculative novel (*Diaspora*), one of the first to apply mind uploading to the entire human species. See Greg Egan 1997.

[42] "Transcension" seems to be the fusion of the terms "transcendence" and "ascension".

> "[…] the miniaturization (spatial localization) of a civilization's engineering, perhaps terminating at the Planck scale. Due to STEM (space, time, energy, matter) compression, intelligent civilizations can presumably continue to develop exponentially more localized, miniaturized, dense, efficient and complex structures and energy flows to generate greater computational and adaptive capacity, right up to the black hole limit and presumably even beyond, as black hole event horizons in stellar mass and supermassive black holes are still well above the apparent Planck-scale limits of universal structure. Thus, if the hypothesis is correct, the Barrow scale, and more generally, STEM efficiency, STEM density, and computational growth scales would be much more appropriate measures of civilization development." (Smart 2012, 59)

To saturate the local region of the universe, in its articulations of space, time, energy and matter, to create the densest and most efficient computational substrate possible and to exploit the (hypothetical) properties of black holes to shape the topology of surrounding space, to travel in time, through the mechanisms of wormholes[43]. As can be seen, the transcension hypothesis embraces numerous assumptions, including a civilization's mastery of the subatomic scale, its computational aspects (and limits), and those properties of the space-time fabric still being explored in cosmology: black holes[44]. Transcension remains a science-fictionally fascinating hypothesis.

If in Smart's escapism the diaspora occurs in Kardashev's counter-scale, in virtual escapism it occurs in the virtual realm of simulations. The idea is that having reached the stage of the singularity, a phenomenon of evolutionary convergence passes through the various civilizations at that level, leading them to the choice of transcending technologically into virtual platforms. Disconnected from their physical form, they would definitively embrace the consequences of mind

---

[43] See Morris *et al.* 1998, 1446–49; Morris & Throne 1988, 395–412.

[44] One of these limits is related to the Bekenstein-Hawking entropy in the context of black hole thermodynamics: there is a maximum amount of information that can be stored in a region of space that depends not on the volume of the region but on the area of its surface (holographic or Bekenstein bound). In the case of the black hole, the entropy is proportional to the area of its event horizon, the surface within which even light cannot escape the gravity: the swallowed information would therefore remain confined to the surface (the area of the event horizon) of the black hole and not in its volume. See Bekenstein 2003, 58–65.

uploading (and its premise, computational functionalism[45]), becoming virtual post-biological entities or digitized civilizations.

There may be a narrow window of time (a few centuries) between the moment when a race reaches interstellar flight and the moment when it ceases to have corporeal existence due to the onset of the Singularitian stage. Compared to the Great Filter, then, each intelligent species either self-destructs shortly after the discovery of nuclear weapons or survives perhaps a couple of centuries and then decides to transcend. This explains why the search for intelligent beings using radio-telescopes is failing. The Great Silence would be the imperceptible hum of the computers in which they are running.

Ultimately, the possibility of a transformation of civilizations through mind uploading (and its underlying assumption, functionalism) and the simulation argument was examined, providing, among other things, a possible explanation for the Great Silence: it is not so much a question of whether or not we are inside a complex simulated reality (Bostrom's hypothesis), but rather of examining the possibility that having reached a certain degree of development, civilizations deliberately decide, for reasons of convenience or survival, to pour themselves into a virtual realm. It is "They" who are inside a simulation, not us.

Acknowledgments:
We would like to thank Flavia Marcacci for helpful comments and critical remarks.

REFERENCES

Barrow, J.D., Tipler, F.J. (1986). *The Anthropic Cosmological Principle*. Oxford: Oxford University Press.

Bekenstein, J.D. (2003). "Information in the Holographic Universe: Theoretical Results about Black Holes Suggest That the Universe Could Be Like a Gigantic Hologram." In *Scientific American* 289(2): 58–65.

[45] On functionalism as an anthropological model of transhumanism see Larrey & Puleio 2021, 24-27.

Bostrom, N. (2003). "*Transhumanist FAQ: A General Introduction*." Version 2.1, World Transhumanist Association: 4-56. Available at <https://www.nickbostrom.com/views/transhumanist.pdf>, last time accessed in 4 December 2022.

Bostrom, N. (2008). "Where Are They? Why I hope the search for extraterrestrial life finds nothing." In *MIT Technology Review* May/June: 72–77. Available at <https://www.nickbostrom.com/extraterrestrial.pdf>, last time accessed in 4 December 2022.

Bostrom, N. (2016[2]). "The Control Problem. Excerpts from Superintelligence: Paths, Dangers, Strategies." In Schneider, S. (eds.), *Science Fiction and Philosophy*, Chichester: Wiley-Blackwell, 308-330.

Bostrom, N. (2003). "Are We Living in a Computer Simulation?" In *The Philosophical Quarterly*, Vol. 53: 243-255.

Campa, R. (2019). *Credere nel futuro. Il lato mistico del transumanesimo*. Krakow: Orbis Idearum Press.

Carter, B., McCrea, H. W. (1983). "The Anthropic Principle and Its Implications for Biological Evolution." In *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, Vol. A 310: 347-363.

Ćirković, M. (2003). "On the Importance of SETI for Transhumanism." In *Journal of Evolution and Technology* 13(2): 1-14.

Ćirković, M. (2018). *The Great Silence. The Science and Philosophy of Fermi's Paradox*. Oxford: Oxford University Press.

Ćirković, M. (2009). "Fermi's Paradox: The Last Challenge for Copernicanism?" In *Serb. Astron*. J., Vol. 178: 1-20.

Clarke, A.C. (2001). "The Sentinel." In *The Collected Stories of Arthur C. Clarke* (ed. A.C. Clarke). New York: Tor Books, New York.

Clynes, M., Kline, N. (1960). "Cyborgs and Space." In *Astronautics*, September: 26-27 and 75-76.

Cocconi, G., Morrison, P. (1959). *Searching for Interstellar Communications*. In *Nature*, Vol. 184(4690): 844–846.

Cusano, N. (2018). *De docta ignorantia*. In Cusano, N., *Opere filosofiche, teologiche e matematiche*, translated by Enrico Peroli (eds.), Milan: Bompiani.

Davies, P. (2010). "The Eerie Silence." *Phys. World* 23 (3):31–32.

Davies, P. (2010). *The Eerie Silence. Renewing Our Search for Alien Intelligence*. Boston: Houghton Mifflin Harcourt.

Davies. P. (2010). "Universe from bits." In Davies, P., Gregersen N. (eds.), *Information and the Nature of Reality: From Physics to Metaphysics.* Cambridge: Cambridge University Press, 65–91.

De Duve, C. (1995). *Vital Dust: Life as Cosmic Imperative*. New York: Basic Books.

Deutsch, D. (1985). "Quantum theory, the Church-Turing principle and the universal quantum computer." In *Proceedings of the Royal Society*, 400(1818): 97–117.

Deutsch, D. (1997). *La trama della realtà*. Torino: Einaudi.

Deutsch, D. (2004). "It from Qubit." In Barrow J., Davies P.C.W., Harper C.L. (eds.), *Science and Ultimate Reality. Quantum Theory, Cosmology, and Complexity*. Cambridge: Cambridge University Press, 90–102.

Dick, S.J. (2003). "Cultural evolution, the post-biological universe and SETI." In *International Journal of Astrobiology* (2)1: 65-74.

Dick, S.J. (2009). "The postbiological universe and our future in space." *Futures* 41(8): 578-580.

Dick, S.J. (2009). "Cosmic Evolution. History, Culture, and Human Destiny." In Dick, S.J., Lupisella M. (eds.), *Cosmos and Culture: Cultural Evolution in a Cosmic Context*. Washington D.C.: NASA History Series, 25-62.

Dick, S.J. (2009). "Bringing Culture to Cosmos. The Postbiological Universe." In Dick, S.J., Lupisella M. (eds.), *Cosmos and Culture: Cultural Evolution in a Cosmic Context.* Washington D.C.: NASA History Series, 463-487.

Drake, F. (2013). "Reflections on the Equation." In *International Journal of Astrobiology,* 12(3):173–176.

Egan, G. (1997). *Diaspora*. London: Gollancz.

Fano, V., Macchia, G. (2020). "Historical-epistemological considerations on the cosmological principle." In *Aquinas*, LXIII, vol.1-2. Città del Vaticano: Lateran University Press, 309–324.

Feynman, R.P. (1982). "Simulating physics with computers." In *International Journal of Theoretical Physics*, 21(Nos. 6/7): 467–488.

Gottschalk, S. (2010). "The Presentation of Avatars in Second Life: Self and Interaction in Social Virtual Spaces." In *Symbolic Interaction*, 33(4): 501–525.

Hanson, R. (1998). *The Great Filter: Are We Almost Past It?,* available at <http://hanson.gmu.edu/greatfilter.html>, last time accessed in 3 December 2022.

Huxley, J. (1957). *New Bottles for New Wine*. London: Chatto & Windus.

Kardashev, N.S. (1964). "Transmission of Information by Extraterrestrial Civilizations." *Sov. Astron*, 8(2): 217–20.

Koene, R. A. (2013), "Uploading to Substrate-Independent Minds". In More, M., Vita-More N., (eds.), *The Transhumanist Reader: Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*, Chichester: Wiley-Blackwell,
146-156.

Kurzweil, R. (2005). *The Singularity Is Near: When Humans Transcend Biology*. London: Penguin Books.

Kurzweil, R. (2016[2]). "Superintelligence and Singularity." In Schneider S. (eds.), *Science Fiction and Philosophy. From time travel to superintelligence*. Chichester: Wiley-Blackwell, 146-170.

Larrey, P., Puleio, D. (2021). "Una scansione del transumanesimo: dalla science-fiction alla filosofia." In A. M. Pezzella, A. Calcagno (eds.), *Challenges of the new frontiers of the human: trans- and post-humanism?*, Per la filosofia. Filosofia e insegnamento, XXXVIII, no. 113, 12-28.

Lloyd, S. (2006). *Programming the Universe: A Quantum Computer Scientist Takes on the Cosmos*. New York: Alfred A. Knopf.

Longo, G.O., Vaccaro, A. (2013). *Bit Bang. La nascita della filosofia digitale*. Milano: Apogeo.

Monod, J., (1996). *Il caso e la necessità*. Milano: Mondadori.

Moravec, H. (1988). *Mind Children. The Future of Robot and Human Intelligence*. Cambridge: Harvard University Press.

Morris, M.S., Thorne, K.S. (1988). "Wormholes in Spacetime and Their Use for Interstellar Travel: A Tool for Teaching General Relativity." In *American Journal of Physics* 56(5): 395–412.

Morris, M.S., Thorne, K.S., Yurtsever, U. (1988). "Wormholes, Time Machines, and the Weak Energy Condition." In *Physical Review Letters* 61(13): 1446–49.

Paura, R. (2019). "A Rapture of the Nerds? A Comparison between Transhumanist Eschatology and Christian Parousia." In *Forum Philosophicum* 24(2): 343-367.

Pearce, D. (1995). *The Hedonistic Imperative*, available online at <https://www.hedweb.com/hedethic/tabconhi.html>, last time accessed in 2 December 2022.

Sagan, C. (1973). *Communication with Extraterrestrial Intelligence*. Cambridge: MIT Press.

Sagan, C. (1975). *The Cosmic Connection, an Extraterrestrial Perspective.* New York: Dell Publishing Co., Inc.

Sagan, C. (1994). *Pale Blue Dot: A Vision of the Human Future in Space*. New York: Random House.

Sagan, C. (1995). "The abundance of life-bearing planets." In *Bioastronomy News* 7(4):1–4.

Sagan, C. (1997). *The Demon-Haunted World: Science as a Candle in the Dark*. New York: Ballantine.

Sagan, C., Drake, F. (1975). "The Search for Extraterrestrial Intelligence." In *Scientific American*, 232(5): 80–89.

Sagan, C., Newman, W.I. (1983). "The Solipsist Approach to Extraterrestrial Intelligence." In *Quarterly Journal of the Royal Astronomical Society*, 24: 113–121.

Schneider, S. (2019). *Artificial You. AI and The Future of Your Mind*. Princeton: Princeton University Press.

Shostak, S. (2009). "The Value of L and the Cosmic Bottleneck." In S.J. Dick, M. Lupisella (eds.), *Cosmos and Culture: Cultural Evolution in a Cosmic Context*. Washington D.C.: NASA History Series, 399–414.

Smart, J.M. (2012). "The Transcension Hypothesis: Sufficiently Advanced Civilizations Invariably Leave Our Universe." *Implications for METI and SETI, Acta Astronaut*, 78: 55-68.

Sorgner, S.L. (2022). *We Have Always Been Cyborgs. Digital Data, Gene Technologies, and an Ethics of Transhumanism*. Bristol: Bristol University Press.

Steinhart, E. (1998). "Digital Metaphysics." In T.W. Bynum, J.H. Moor (eds.), *The Digital Phoenix: How Computers Are Changing Philosophy*. Oxford, New York: Blackwell, 117–125.

Tagliagambe, S. (1997). *Epistemologia del Cyberspazio*. Cagliari: Demos.

Tanzella-Nitti, G. (2002). "Vita Extraterrestre." In G. Tanzella-Nitti, A. Strumia (eds.), *Dizionario Interdisciplinare di Scienza e Fede*. Roma: Città Nuova - Urbaniana University Press, 591-605.

Tegmark, M. (2017). *Life 3.0. Being Human in the Age of Artificial Intelligence*. New York: Vintage Books.

Tipler, F. (1995). *La fisica dell'immortalità. Dio, la cosmologia e la risurrezione dei morti*. Milano: Mondadori.

Vinge, V. (1993). "The Coming Technological Singularity: How to Survive in the Post-human Era." In *NASA Technical Reports, Vision 21: Interdisciplinary Science and Engineering in the Era of Cyberspace*. Lewis Research Center, 11–21.

Ward, P.D., Brownlee, D. (2000). *Rare Earth: Why Complex Life Is Uncommon in the Universe*. New York: Springer.

Waters, B. (2011). "Whose Salvation? Which Eschatology? Transhumanism and Christianity as Contending Salvific Religions." In Cole-Turner, R. (eds.), *Transhumanism and Transcendence: Christian Hope in an Age of Technological Enhancement*. Washington D.C.: Georgetown University Press, 163–175.

Webb, S. (2015). *If the Universe is Teeming with Aliens...Where is Everybody? Seventy-Five Solutions to the Fermi Paradox and the Problem of Extraterrestrial Life*. New York: Springer.

Wheeler, J.A. (1989). "Information, physics, quantum: the search for links." In *Proceedings of the III International Symposium on Foundations of Quantum Mechanics*, Tokyo, 354–368.

I THINK, THEREFORE WE ARE

JASON M. PITTMAN[1]

Abstract

Humanity is transforming from a pure biological organism to a biodigitally converged being. The transformation has been hastened by rapid increases in power and commercialization of advanced technologies such as brain-computer interfaces and artificial intelligence. All indicators point towards a trend of continual acceleration away from baseline biology as the human becomes the source of computing rather than simply a user. While mainstream transhumanists laud such development as exemplative of progress towards a grand vision of man and machine becoming one, a variety of concerns, both philosophical and technological, are readily observable. This paper focuses on one such concern associated with the intersection of brain-computer interfaces and embedded artificial intelligence on those devices. Within this intersection, the biodigital convergence of a brain-computer interface is known to produce issues related to user agency. However, what is not known is how brain-computer interfaces with embedded AI may dampen or amplify such issues and what potential technological solutions may exist to ameliorate the agency problems. Furthermore, there is a lack of an applied epistemology to guide any such understanding. To that end, this paper describes a propositional metadata schema which may serve as a brain-computer interface action labeling mechanism for the information output by the biodigital system. The theoretical and practical significance of the schema are explained and ideas for future work are provided.

Keywords: transhumanism, cyborg, biodigital convergence, brain-computer interface, artificial intelligence

[1] University of Maryland, Global Campus Adelphi, Maryland USA.
Email: <jason.pittman@umgc.edu>.

Introduction

We find transhumanism where biology and digital technology meet. The goal of this meeting, even if by happenstance, is to maximize the well-being and longevity of the biological. Such a meeting, or convergence, is enabled by modern society's treatment of biology as information (Reader 2022, 23). The line of reasoning then is if biology is information, and if information is the result of computation[2], then there must be a biological function capable of being the source of such computation. In the views of some transhumanists, this meeting is the essence of what it is to be both a transhuman and a cyborg. Such transhumanists perceive humanity as inevitably transforming from a purely biological create to a converged organism, part biology and part digital technology.

The position is not without merit. Today's digital technologies are accelerating away from what existed even months ago. Biology, whether integrated with digital components or modified through non-digital technology, has become intertwined with digital technology. Thus, the technologist transhumanist sees humanity as *becoming* a cyborg.

Other transhumanists, principally Sorgner (2022), emphasize carbon-based innovations such as gene enhancement over silicon-based technology. In actuality, this view creates a smoother glide path for carbon-based innovation going forward because our species has exercised it in one form or another through natural means. To this end, Sorgner points to language education and sexual selection as transhumanist, and thus cyborg, forces. To Sorgner's point, we have always been cyborgs and simply are becoming more so as we refine the tools, techniques, and technologies.

The emphasis on carbon-based transhumanism is understandable given the far future-leaning focus of mainstream transhumanists (Kurzweil 2005; Aithal and Aithal 2018) on extreme technologies such as consciousness transfer, artificial general intelligence, and technological singularities. Thus, we should not view transhumanists such as Sorgner as anti-technology. Rather, the point acknowledges a potential white space between the different views.

---

[2] Perhaps more appropriately we can use the phrase *information processor*.

A recent review (Pittman 2022, 8-9) of *We have always been cyborgs* (Sorgner 2002) mentioned the augmentative role of narrow artificial intelligence as a silicon-based transhuman technology. Narrow AI encompasses the types of AI we have today and is line with Sorgner's assertion that silicon-based technologies represent a, "pragmatic necessity" (2022, 22) for increasing lifespan and general good. By extension, adjacent digital technologies which take advantage of narrow AI perhaps reveal clues about what is possible in the transhumanist white space. This is the thread I wish to pull at in this work.

An example of one such adjacent digital technology are brain-machine interfaces (BCI). Because the BCI establishes a direct link between a brain and an external computing system, I consider BCI to be a strictly biodigital converged technology. Moreover, because BCI extends baseline human biology in a "self-overcoming" (Sorgner 2022, 1) manner, the technology is transhumanistic. Ross (2020) articulated a similar philosophy with the notion, "[e]nhancement is a key idea within transhumanism since it involves the elevation of human capacities beyond a baseline of normal human limitations" (25).

There can be little doubt the BCI existing today are powerful human-computer augmentations. What is more, the BCI of tomorrow, integrated with AI (Portillo-Lara *et al*. 2021), will be demonstrably cyborg. Still, there are unique challenges in the biodigital convergence of BCI and AI, however (Zhang *et al*. 2020). Foremost of these challenges is BCI user agency. More distinctively, existing BCI are known to elicit a sense of "disembodied agency" (Steinart 2019, 459). When BCI technologically evolve to include embedded AI for delegated decision-making and autonomous *intentional* action, we must anticipate the intensification of agency issues.

To that end, this work describes a *thought metadata schema* based on emergent biodigital convergence themes. The goal is to extend the silicon-based transhumanism philosophy presented by Sorgner (2022). I attempt to provide a bridge between Sorgner's work and near-future brain-computer interface technology as a definitive cyborg augmentation. I then employ a propositional framework to describe the schema with goal of establishing a foundation for future work.

1. Related Work

Biodigital convergence is inherently a multi-disciplinary construct. The construct is sufficiently developed to warrant discussion. Additionally, we must account for the specific technology involved in the convergence. Specifically, the research domains most related to this work are artificial intelligence, brain-machine interfaces, and agency. I highlight prominent ideas from these fields as well as open problems and challenges.

Biodigital convergence

Technology, being the scientific application of a knowledge domain to address a practical challenge, marries well with an evolutionary driven field like biology. The Nature drives the latter, the former selected through market pressures and paradigmatic innovation. The similarity in forces allows *bios* and *techne* to converge into the transhuman, the cyborg.

The convergence of biology and digital technology has engendered a variety of innovations. These innovations are externally invisible in the case of gene modification by CRISPR-CAS9 and externally observable in instances such as biohacking to implant digital devices. One unifying technology layer across the majority innovations is the use of AI, machine learning specifically, to enhance the fidelity of the biodigital information.

Furthermore, as with Heidegger's views, the technology segment of the biodigital convergence shapes the biology. Such shaping, even when viewed within Sorgner's concept of positive "self-overcoming" (2022, 1), comes with trade-offs. For instance, O'Riordan (309) suggests biodigital convergence subjects the biological to the norms of the digital. Others, such as Jandrić and Knox (2021) rightfully question the assumptions inherent in assuming biodigital convergence is determinist or instrumentalist.

Interestingly, the innovations coming out of biodigital convergence, while rooted in developments from the 1980s and 1990s, have all appeared with the dawning of the 21st century. What we know think of as biology started with the human genome project (Reader 2022, 24) and evolved into advanced genetic manipulation techniques. Moreover,

what we think of digital technology has transitioned from large beige boxes on our desktop to system-on-a-chip computers embeddable in the human body (Belwafi, Gannouni, and Aboalsamh 2021). Overall, the biodigital domain is in its infancy but has demonstrated an aggressive growth curve.

There can be little doubt biodigital convergence in any form will impact the individual, culture, society, and the species. To what extent and in what manner are questions open? The field is not yet close to being mapped out. Thus, answers to such inquiries necessitate being more specific with regard to the digital technologies involved in the convergence under examination. Within the bounds of this work, AI and BCI are those technologies.

Artificial Intelligence

The general notion of AI is to imitate human intelligence (Carter 2007). Thus, AI is a broad discipline and can be an ambiguous term. In the context of this work, AI can be taken as synonymous with the specific sub-discipline, machine learning (ML). I choose to narrowly scope AI in this way for three reasons.

First, other sub-disciplines such as natural language processing and computer vision are AI fields but have little or nothing to contribute to the purpose of this work. BCI benefit from learning algorithms that reliably generalize from broad inputs to specific outputs. Such algorithms are necessary to consolidate a continuous brain activity signal into a discrete compute input signal. At the same time, there is no technological impetus behind extending BCI to include other forms of AI.

Second, ML algorithms implement a specific inductive logic to achieve a limited set of outcomes. In broad terms, ML either classifies inputs into trained categories or uses statistical regression to assert predictions. Again, other sub-fields within AI do not perform function similarly. Interestingly, the algorithmic induction implemented through ML is tantamount to experimental philosophy of science (Bensusan 2000).

Lastly, stemming from the prior two reasons, the associated philosophy becomes clearer. Starting with Hume, we can trace a clear line to modern thinking as represented by Strawson, Korb, Searle, Dennett,

Chalmers, among many others. To be clear, the specific line connects the Humean root of inductive reasoning to the inability for a single, universal ML capable of solving every type of learning problem.

This last point also allows us to distinguish ML from AI to avoid larger, systemic challenges in the broader AI context. For instance, a goal of transhumanists is artificial general intelligence (Kurzweil 2005). The ethics of pursuing the goal are controversial and the likelihood is hotly debated. ML avoids the controversy and debate by virtue of being a narrow AI construct. Furthermore, AI has epistemological issues related to embodiment (Luvaanjalba and Su 2022) as well as two-minds and brain-in-a-vat challenges. Again, ML avoids such issues because the algorithms do not attempt to *know*, only update previously trained values. This makes ML a powerful augmentative digital technology for biodigitally converged systems.

Brain-Computer Interfaces

Brain-computer interfaces (BCI) are an augmentative biodigitally converged technology with applicability to a variety of applied domains (Ramadan & Vasilakos 2017). Put simply, a user's brain activity is converted to computing inputs through invasive (*i.e.*, surgically implanted in the brain) or non-invasive (*i.e.*, scalp surface electrodes) devices. Applications of digital technology range from consumer toys, military-based human-machine teaming, to medical augmentation to offset various disabilities. On one hand, BCI are a form of human-computer interaction not substantively different from a keyboard and mouse. On the other hand, the material difference in BCI enables behaviors Buller (2021) described as *phenomenologically similar* to ordinary behavior. Thus, while the BCI gives us the ability to interact with a computer as if we're using a keyboard and mouse, the experience is as if we are using the keyboard and mouse, not using the BCI.

A majority of BCI research within the philosophy domain is concerned with the ethics of the technology (Burwell, Sample, *and Racine* 2017; Orito *et al*. 2020; van Velthoven 2022). Outside of ethical inquiries, work by Tamburrini (2014) explored the epistemological facets of BCI, concluding there is a, "need for effective and responsible communication

strategy" (160) in relation to the human user, the BCI, and the computing device being controlled. This claim seems to stem from what Heersmink (2011) articulated as a, "strong symbiosis and reciprocity" (2) between BCI and human. In other words, without tight coupling between all three components in the BCI system, and appropriate communications therein, there may be epistemological deficiencies.

Within the technology domain, research is focused on developing the BCI hardware and increasing BCI communications fidelity. The integration of AI and BCI is solved, at least in general. Specifically, AI is used to extrapolate a user's EEG patterns into discrete computer inputs (Aggarwal, and Chugh 2022). Notably, the AI is not autonomous in these integrations. The system cannot behave independent of a user's active engagement despite what the user may believe the BCI can do.

The cutting-edge of BCI technology is represented by Neuralink (Fiani 2021). One novel aspect of the solution is the ability to carry information to the user's brain. While future-leaning and speculative, I suggest it is not far-fetched to imagine an AI embedded in the BCI. There is a limited literature supporting this suggestion (Zhang *et al*. 2020). The narrow AI would be useful in learning user patterns of computer interaction. Such an AI then might augment the interaction by providing additional signaling and the AI may act autonomously to produce interactions we offload by delegation. Such an AI will be situated to filter or otherwise mediate input to the brain. However, an open problem, and the fundamental issue addressed in this work, is the applied epistemology associated with transmitting and receiving phenomenological information[3] through a brain-computer interface with an embedded augmentative AI mediating our knowledge in a simulation. Consequently, I feel some foundational understanding of agency in this context is required.

Agency

Not only humans affect technology while using it, but also such interactions ultimately affect humans by consequences of such use. In

[3] I purposefully avoid using the term *qualia* here because the focus is on the information, not the experience of the information.

fact, technology mediated experience and impact to one's sense of agency is an area of active research. Indeed, there have been a number of studies investigating agency in the context of BCI. Haselager (2013) provides a robust summary.

For this work, I suggest agency be conceptualized as a metadata label, applicable to the phenomenological artifact arising when one intends to act and observes the act embodied in the world. That is, we don't normally think of agency prior to acting. Rather, we ascribe agency to behaviors post hoc. The foundation for the type of agency related to the present work is Libet's (1985; 1999) experimentation. Philosophy's reaction to Libet has been to separate agency into two forms.

The differentiation is based on (a) basic versus nonbasic behavior and (b) mental versus bodily behavior. The difference between mental and bodily is defined by Mele (1997) as occurring either entirely within the mind, without involvement of the physical body or with the physical body, respectively. Both involve a sense of agency albeit with differing characteristics. For example, thinking about using a calculator to sum two integers and imagining the output gives one the sense of having used a calculator to some degree. While similar, there is a material difference in the sense of agency associated with physically using the calculator.

Importantly, Metzinger (2013, 2) further refined mental agency to subordinate *attentional* and *cognitive* types. Taking the mental-cognitive route, the agency focus shifts to planning and decision-making. Both planning and decision-making fall within the sphere of narrow AI expertise. Consequently, there may be an opportunity to attach an artifact to outbound BCI communications originating from the embedded AI which provides redress to the agency problem.

Concurrently, we must be cognizant of the in-kind difference between actions happening to us and actions we cause. This *philosophy of action theory* is relevant to the unique mediation of action BCI provides. As Steinert *et al*. (2017) explained, "[a] motionless body, the epitome of inaction, might be acting" (457). Therefore, according to Lynn *et al*. (2010, 5) it is possible to craft an illusory sense of agency in BCI users. This effect has been replicated and expanded upon (Lopez-Sola, Moreno-Bote, and Arsiwalla 2021).

Haselager (2013, 415) suggests a root cause of negative BCI agency phenomena is a lack of reliability in users' determination of causation. Some work (Vlek *et al*. 2014; Haselager, Meacacci, and Wolkenstein 2021) has explored options to extend BCI to include some form of self-recognition for the user. Put differently, work is underway to develop a means for a BCI user to recognize the result of an action as the effect of his or her intent to action in the same way. This is meaningful work, but it has considerations for us to explore. To be precise, the cutting-edge of BCI agency solutions do not account for the presence of an augmentative AI with semi-autonomy (Haselager 2013, 416).

2. Thought Metadata Schema

If we assume a near-future possibility of brain-computer interface with an embedded AI, we need to begin developing a metadata labeling schema. The initial goal of the schema is to provide a metadata level differentiation between:

(1) an action stemming from BCI user intent;
(2) an action stemming from BCI user intent and filtered by the embedded AI and
(3) an action stemming from the BCI AI absent BCI user intent.

It follows the schema requires three functions, mapped one-to-one with the above goals. One function is bound to actions originating from user brain activity wherein the BCI embedded AI sends the communications without modification. A second function is bound to those actions which originate from the BCI user but are *filtered* by the embedded AI. Here, we take filtering to mean *classification* in the ML context for example. Lastly, a function is required for communications the AI sends autonomously. In all three cases, I assume the receiving computing system is capable of receiving the communication and reacting according to the schema.

Overall, I describe *actions* using the set $\mathbf{A}$. This is the set containing both user intended actions as A_u as well as embedded AI intended actions as A_i. Additionally, the embedded AI possesses a set of filters as $\mathbf{F}$. Meanwhile, the receiving computing systems have a companion set $\mathbf{R}$ containing $\mathbf{A_u}, \mathbf{A_f}, \mathbf{A_i}$ that have a relation to the stated goals. These will be used by the receiving system to signal source of action to the BCI user.

Then, more specifically, when the BCI communication passes through the embedded AI without modification, we can apply the following propositional description.

$$(1)\ \forall a \in A_u \wedge \neg A_i : \mathrm{a_u} \rightarrow \mathrm{a}$$

The (1) function encapsulates the need for a BCI user to act without the BCI with embedded AI augmentation and while acknowledging the mediation of the biodigital convergence. Principally, the function communicates all actions in the set of user intended actions – and not embedded AI intended actions – as *actions* to the receiving system. This is necessary because in advanced applications the BCI cannot be shut down without severing all communications. However, the embedded AI can forward actions without modification. I anticipate the receiving computer system will display a notification indicative of $\mathbf{A_u}$ being the cause of the action.

In the event a user intended action needs to be mediated by the embedded AI, the BCI can leverage the following propositional description:

$$(2)\ \forall a_u \in A_u \wedge \exists f \in \mathbf{F} : a_u \cup f \rightarrow \mathrm{a_f}$$

This (2) function is operationally similar to the prior. However, the addition of at least one *filter* is meaningful because the output $\mathrm{a_f}$ is not the original intent of the user. Rather, $\mathrm{a_f}$ is a synthesis of user intended action. This is necessarily so because the embedded AI abstracts fuzzy intent into discrete computer input. For example, the continuous EEG values for moving a mouse cursor up on a screen must be synthesized

into the discrete *up* input code. Indeed, such synthesis is how existing BCI operate. Here, the function is (a) executed within the BCI and (b) attaches the a_f marker to the output. Reciprocally, the receiving computer system can display a A_f notification.

Finally, we have a function (3) for when the BCI embedded AI acts autonomously. Importantly, this function describes the means for the biodigital convergence to integrate with an external computing system without user intent or intervention. Importantly, the function blocks any potential inclusion of user intended action in the same operation as autonomous AI action by negating the A_u set. Further, I want to point out prior training is a strong constraint on the function. That is, the set A_i contains only those actions the AI is trained to perform. Thus, I suggest the receiving system notification of A_i is sufficient to avoid a truly disembodied sense of agency for the user.

$$(3)\ \forall a_i \in A_i \wedge \neg A_u : a_i \rightarrow a$$

I anticipate the reasonable question at this point, *how does the embedded AI understand which function to apply*? The simple answer is the AI does not understand. In fact, understanding is not required. Instead, the AI classifies input according to the categories established through previous training. The basis of the classification is nothing more than a computational instantiation of the mapping between *relata*. More completely, the information processing (*i.e.*, classification) lacks a distinct epistemology on its own. This is expected because the AI augments the human user, it does not replace him or her.

Therefore, based on the goals and schema, I suggest we can harness the embedded AI to produce a recognizable artifact in the communication streams between human and computer. I would point out again the value of using a narrow AI (*i.e.*, ML) algorithm is we can sidestep the host of (legitimate) philosophical issues related to artificial general intelligence. Thus, there is little danger in adding our metadata schema functions to the embedded AI[4].

[4] I recognize machine learning, despite being a mature technology, has challenges such as bias, adversarial attacks, and so forth. However, there is an active research literature

In a similar manner, I suggest the outcome of this metadata schema is a narrow, but sound applied epistemology. The applied epistemology is narrow insofar as the metadata schema leads to true beliefs about the user's actions and the embedded AI's actions *only* in the set of BCI use cases. Further, the soundness of the metadata schema approach is substantiated in associated literature. For instance, the metadata schema encodes a *cybernetic reflexive* relation between human, BCI, and actions consistent with Glanville (1975). As well, the metadata schema facilitates *self-organisation* (Espejo 2015) through *self-observation* (Krippendorf 1984) and a *self-steering* positivism (Sorgner 2022).

3. Conclusion

Humanity is transforming into a biodigitally converged organism. To mainstream transhumanists, this development points towards progress on the way to silicon-based cyborg goals such as mind transfer and digital immortality. Like Sorgner, I view mind uploading or consciousness transfer as unnecessary as a defining feature of transhumanism. There are a variety of non-extreme biodigital technologies contributing to our continued evolution as cyborgs. More importantly, whether the ultimate mainstream transhumanism goals are obtainable or not does prevent other innovations from being realized.

For instance, BCI are available today. These devices use brain activity to control computing systems. While the noninvasive type may not seem like a transhumanist silicon-based technology, the invasive type certainly should since it is implanted into the user's brain. The future-leaning innovation in this space is represented by Neuralink which is notably planned to allow bidirectional- brain to system and system to brain- communication.

BCI are not without challenges, however. Users can develop an illusory sense of agency. While not conclusive, the literature suggests problems related to agency stem from how the BCI mediates the users'

---

developing countermeasures to these issues. My intention here is to delineate between technological issues and philosophical issues.

experience of their own behavior. More specifically, users can become decoupled from the experience of causing a change in the external environment because it is not perceptively clear whether intent to act resulted in the act.

Concurrent to BCI, AI has emerged from the scholarly realm and become a mainstream talking point. Further, the cutting-edge of BCI technology already incorporates AI to limited degrees. In particular, ML has demonstrated a capability to augment decision-making and planning mechanisms because of the constrained classification and regression algorithms present.

This paper explored the intersection of BCI with an embedded AI. I described a propositional metadata schema which may serve as a decision labeling mechanism for the information output by the biodigitally converged system. My goal was to provide a potential mechanism to bridge the gap between the underlying philosophy of transhumanism, the problem of user agency, and advancing biodigital convergence.

With this in mind, the metadata schema has theoretical significance as it represents a foundational knowledge architecture for addressing agency issues with BCI as well as BCI with embedded AI. The architecture can be used by philosophers and technologists alike to establish more extensive logical frameworks. Moreover, as a propositional artifact, the schema has practical significance as it can be readily translated into a programmed algorithm. Experimentation building from the baseline provided in the literature for BCI agency may be of interest to computer scientists and technologists.

Additionally, future work is necessary in this area. BCI and related biodigitally converged transhumanist technology will not cease to advance in sophistication. On one hand, a systematic review of open challenges and issues will benefit tomorrow's research by providing a stable foundation. As well, because biodigital convergence and BCI are multi-disciplinary endeavors, all relevant knowledge domains should come together to develop a general taxonomic road map for critical technologies, terminology, and transhuman objectives. Such necessarily includes computer science, information technology, philosophy of science, philosophy of technology, and social epistemology and perhaps should also adopt futurists, ethicists, and biologists as well.

REREFENCES

Aggarwal, Swati, and Nupur Chugh (2022). "Review of machine learning techniques for EEG based brain computer interface." *Archives of Computational Methods in Engineering*, 29(2022): 3001–3020.

Aithal, P. S., and Shubhrajyotsna Aithal (2018). "Nanotechnology based Innovations and Human Life Comfortability–Are we Marching towards Immortality?" *International Journal of Applied Engineering and Management Letters* (IJAEML),2(2): 71-86.

Belwafi, Kais, Sofien Gannouni, and Hatim Aboalsamh (2021). "Embedded brain computer interface: state-of-the-art in research." *Sensors* 21(13): 1-20.

Bensusan, Hilan (2000). "Is machine learning experimental philosophy of science." In *ECAI2000 Workshop notes on scientific Reasoning in Artificial Intelligence and the Philosophy of Science*, 9-14

Buller, Tom (2021). "Actions, Agents, and Interfaces." In *Clinical Neurotechnology meets Artificial Intelligence*, 11-23. Springer, Cham.

Burwell, Sasha, Matthew Sample, and Eric Racine (2017). "Ethical aspects of brain computer interfaces: a scoping review." *BMC medical ethics*, 18(1): 1-11.

Carter, Matt (2007). *Minds and computers: an introduction to the philosophy of artificial intelligence*. Edinburgh: Edinburgh University Press.

Espejo, Raul (2015). "The cybernetics of self-organisation." *Kybernetes*. Online source, available at: http://wosc.co/wp-content/uploads/2015/12/Editorial-for-CYBERNETICS-OF-SELF-ORGANISATION-Kybernetes-Vol-42-Nos6-7.pdf, accessed December 16, 2022.

Fiani, Brian, Taylor Reardon, Benjamin Ayres, David Cline, and Sarah R. Sitto (2021). "An Examination of Prospective Uses and Future Directions of Neuralink: The Brain-Machine Interface." *Cureus* 13(3): 1-4.

Glanville, Ranulph (1975). "A cybernetic development of epistemology and observation applied to objects in space and time (as seen in architecture)." PhD diss., Brunel University School of Engineering and Design. Online source, available at: https://bura.brunel.ac.uk/bitstream/2438/4286/1/FulltextThesis.pdf, accessed December 16, 2022.

Haenlein, Michael, and Andreas Kaplan (2019). "A brief history of artificial intelligence: On the past, present, and future of artificial intelligence." *California Management Review* 61(4): 5-14.

Haselager, Pim (2013). "Did I do that? Brain–computer interfacing and the sense of agency." *Minds and Machines* 23(3): 405-418.

Jandrić, Petar, and Jeremy Knox (2021). "The postdigital turn: Philosophy, education, research." *Policy Futures in Education* 20(7): 780-795.

Krippendorff, Klaus (1984). "An epistemological foundation for communication." *Journal of Communication* 34(3): 21-36.

Kurzweil, Ray (2005). *The singularity is near: When humans transcend biology*. London: Penguin.

Negroponte, N. (1998, January 12). *Beyond digital. Wired.* Online source, available at: https://web.media.mit.edu/~nicholas/Wired/WIRED6-12.html, accessed December 16, 2022.

Orito, Yohko, Tomonori Yamamoto, Hidenobu Sai, Kiyoshi Murata, Yasunori Fukuta, Taichi Isobe, and Masashi Hori (2020). "The ethical aspects of a 'psychokinesis machine': An experimental survey on the use of a brain-machine interface." In *Societal Challenges in the Smart Society*, pp. 81-92. Universidad de La Rioja.

Libet, B. (1999). "Do we have free will?" *Journal of Consciousness Studies*, 6(8–9), 47–57.

Libet, B. (1985). "Unconscious cerebral initiative and the role of conscious will in voluntary action." *Behavioral and Brain Sciences*, 8(4): 529–539.

Lopez-Sola, E., Moreno-Bote, R., & Arsiwalla, X. D. (2021). "Sense of agency for mental actions: Insights from a belief-based action-effect paradigm." *Consciousness and Cognition*, 96(2021): 103-225.

Luvaanjalba, Batnasan, and Bo-chiuan Su (2022). "An Epistemological Analysis of the 'Brain in a Vat' Approach for the Philosophy of Artificial Intelligence." In *International Conference on Human-Computer Interaction*, 97-111. Springer, Cham.

Lynn, Margaret T., Christopher C. Berger, Travis A. Riddle, and Ezequiel Morsella (2010). "Mind control? Creating illusory intentions through a phony brain–computer interface." *Consciousness and Cognition* 19(4): 1007-1012.

O'Riordan, Kate (2011). "Revisiting digital technologies: Envisioning biodigital bodies." *Communications* 36 (2011): 291-312.

Pittman, Jason M. (2022). "Review of Stefan Lorenz Sorgner's We Have Always Been Cyborgs." *Social Epistemology Review and Reply Collective* 11(9): 38-45. Portillo-Lara, Roberto, Bogachan Tahirbegi, Christopher AR Chapman, Josef A. Goding, and Rylie A. Green (2021). "Mind the gap: State-of-the-art technologies and applications for EEG-based brain–computer interfaces." *APL Bioengineering* 5(3): 1-16.

Ramadan, Rabie A., and Athanasios V. Vasilakos (2017). "Brain computer interface: control signals review." *Neurocomputing* 223(2017): 26-44.

Reader, John (2022). "Biodigital Becoming." In *Bioinformational Philosophy and Postdigital Knowledge Ecologies*, 23-37. Springer, Cham.

Ross, Benjamin (2000). *The philosophy of transhumanism: A critical analysis*. Bradford: Emerald Group Publishing.

Sorgner, Stefan Lorenz (2021). *We have always been cyborgs: Digital Data, Gene Technologies, and an Ethics of Transhumanism*. Bristol: Policy Press.

Wood, C. C. (2019). "The computational stance in biology." *Philosophical Transactions of the Royal Society* 374(1774): 1-5.

Steinert, Steffen, Christoph Bublitz, Ralf Jox, and Orsolya Friedrich (2019). "Doing things with thoughts: Brain-computer interfaces and disembodied agency." *Philosophy & Technology* 32(3): 457-482.

Tamburrini, Guglielmo (2014). "Philosophical reflections on brain–computer interfaces." In *Brain-Computer-Interfaces in their ethical, social and cultural contexts*, 147-162. Springer, Dordrecht.

van Velthoven, E. A. M., O. C. van Stuijvenberg, D. R. E. Haselager, M. Broekman, X. Chen, P. Roelfsema, A. L. Bredenoord, and K. R. Jongsma (2022). "Ethical implications of visual neuroprostheses—a systematic review." *Journal of Neural Engineering* 19(2): 1-19.

Vlek, Rutger, Jan-Philip van Acken, Evine Beursken, Linsey Roijendijk, and Pim Haselager (2014). "BCI and a User's Judgment of Agency." In *Brain-Computer-Interfaces in their ethical, social and cultural contexts*, pp. 193-202. Springer, Dordrecht.

Zhang, Xiayin, Ziyue Ma, Huaijin Zheng, Tongkeng Li, Kexin Chen, Xun Wang, Chenting Liu *et al*. (2020). "The combination of brain-computer interfaces and artificial intelligence: applications and challenges." *Annals of Translational Medicine* 8(11): 1-9.

HOMO KYBERNETES. FROM ONTO-GENESIS TO THE POSTHUMAN CONDITION

ŽARKO PAIĆ[1]

Abstract

The author deals with the consideration of what he calls the post-metaphysical complex of the technosphere, for which the ontogenetic framework of classical metaphysics from Aristotle to Hegel is no longer valid. Instead, the question of the essence of man posits his transcendence in the emergence of *Homo kybernetes*, with which he removes the posthuman condition. To explain how the permeation of human and non-human appears in all areas of contemporary life as a connection and relationship of information, feedback, control and communication, it seems necessary to open the issue of the cyborgization of life itself. At the same time, the analysis of transhumanism in the opinion of Stefan Lorenz Sorgner proves to be stimulating for understanding what the digital age of the technosphere presupposes, and that is the question of the possibilities of freedom beyond its metaphysical framework as it was considered in the speculative thinking by Schelling. Instead of freedom as absolute, the time has come to establish the possibility of thinking about freedom as an experiment of thought and action beyond human reduction and its associated anthropologies that have lost their credibility in an encounter with the complexity of *Homo kybernetes*.

Keywords: *Homo kybernetes, metaphysics, information, technosphere, onto-genesis, posthuman condition*

1. Metaphysics and Technosphere

Without thoughtful reflection on the challenge of the relationship between *technology and (post)humanity*, it seems impossible to persist in the relationship between technoscience and humanities that would not

[1] University of Zagreb, Professor of Aesthetics and Media Studies.
Email: <zarko.paic@ttf.unizg.hr>.

become non-contemporary in one way or another. Although the prefix post is in parentheses, and humanity is without it, it seems obvious that today there is still something that does not reduce human prerogatives and achievements to the logic of technological singularity. As if this play between what comes after the human and what is here and now human-too-human, to use Nietzsche's famous statement, points to the necessary difference between what constitutes the essence of posthumanism and humanity at all, which becomes the highest reach of *technosphere* rule in managing societies and culture with the help of artificial intelligence. But what if this small difference no longer exists or if it is suspended and neutralized? What, furthermore, if a man has never been able to separate from technology in historical development, *i.e.* what if *we have never been humane*, as Donna Haraway claims, or what if *we have always been cyborgs*, as Stefan Lorenz Sorgner puts it? (Haraway 1990, Sorgner 2022)

While, namely, there is a relationship between the human and the non-human, the metaphysical structure of thought is at work, even if Being is replaced with the concept of information, and beings with a multitude of pure contingencies. A relationship denotes always a possibility of a union at a distance. Synthesis, therefore, presupposes a real state of analytical separation, for example, of different substances in some hybrid assemblies such as contemporary technoscience with biogenetics, biocybernetics and biosemiotics. When we talk about the relationship between the human and the non-human, we must start from the assumption that the former is always the model for understanding the latter. Man in so-called human nature or its essence cannot be understood within metaphysics in any other way than essentialist. This means that his position within Being and his being are always given to him spiritually and biologically, *i.e.* vertically, following Aristotle's famous definition of *animal rationale.* Furthermore, the relationship between the two, starting from the human primacy, is therefore not only horizontally reduced to biologism, from which psychologism and anthropology follow, which Heidegger in *Being and Time (Sein und Zeit)* tried to eliminate within the framework of the systematic "destruction of traditional ontology". (Heidegger 2018) The reason should be seen in the fact that he considered that by doing so "man" or *being there* (*Dasein*)

completely fails in his thinking and thereby reduces to exactly what he wanted to critically overcome, namely Aristotle's definition of *animal rationale.* Biologicism denotes the reduction of man to naturalism, that is, to the so-called biological factors of the whole of his own; psychologism becomes a reduction to the so-called unconscious articulation of language, which reduces man to a pre-and-post-rational living being, while anthropology means the reduction of man to the subjectivity of the subject by which all other beings appear at a lower level of cosmic-biological evolution and thus make him the master of nature as a whole. Heidegger's critique of the reduction of the meaning of Being to the so-called life processes in their inclusive vitality affects the triad of biology-psychology-anthropology because they think a man insufficiently and inauthentically according to this conceptual-categorical "rationalization" of language. Any reduction of the human in man to the non-human necessarily fragments his integrity and irreducibility as a being among beings. However, every reduction cannot be the same as the annulment of human authentic existence in the openness of freedom. To reach the purity of the human essence, phenomenologically speaking, it is necessary to understand that such an essence comes from the dark core of the non-human, which was already known by pre-Socratic thinkers when they spoke about the relationship between demons and humanity. We should remember that Socrates was also constantly followed and urged to freely express his thoughts in public by his demon.

So, we shall see that the question of the relationship between humans and non-humans cannot be simply resolved by a vulgar distinction between the biological and the post-biological when it comes to the relationship between humans and post(humanity). It is not difficult to understand that the problem of this relationship is that both relata in the relationship are derived from the difference against the life force of what comes from the concept of nature (*physis*), which is life (*bios*). The human in the metaphysical sense is always rising above the limits of biologically understood life as mere survival at the animal level. Therefore, it seems self-evident that from this every form of animalism will be considered as a loss of the so-called human dignity. This term has its origin in connection with the ethical-aesthetic doctrines of the Renaissance notion of humanism in Marsilio Ficino and Pico della

Mirandola because dignity is associated with the high rank of a human being against the bare life of an animal and the brutality with which the meaningful life of man in the historical-epoch sense is threatened by the fall to the level of bestiality, refers to the good and the beautiful that transcends the limits of life. Therefore, "human dignity" represents an ethical-aesthetic ideal of living. This transcends the limits set by biological affects and drives. After all, in borderline situations when a person defies tyrannical forms of government with his own life when he is thrown into shackles or slavishly humiliated, the last thing that remains will be the human dignity of a person as an existential experience of indomitable freedom. It seems obvious from this that life in its complex contingency has no meaning without the originally conceived freedom. However, this does not mean that the biological organization of human existence is considered null and void for man and humanity in general. (Paić 2022b, 1-36)

The human versus the non-human denotes clearly what unites spirit and life understood in a metaphysical sense, even when it comes to the realization of metaphysics in cybernetics. Namely, the spiritual always refers to the "rational", and the vital to the "animal", if we follow the logic of Aristotle's definition of man. This problem of the relationship between the cognitive possibilities of man's existence and the biological-physical limits of his physicality should by no means be lightly declared a relic of Cartesian metaphysics. This assemblage has been established in that way as a system of signifiers with which the concept of self-consciousness presupposes the necessary predominance of *res cogitans* over *res extensa*. In contemporary neuroscience and neurophilosophy, this dualism becomes critically overcome by referring to the complex structure of interaction between thought and its natural environment, starting from the notion of a network that has neither centre nor edge. But despite this "new" anti-essentialism that should abolish the distinction between human and non-human by requiring a different notion of all post-metaphysical concepts such as network, interactor, communication and cybernetic corporeality, we still live in two worlds. These worlds touch and permeate each other. For this reason, it seems that the most important philosopher of cybernetics, Gilbert Simondon, was right when he considered that the future of the

relationship between humans and thinking machines does not mean the abolition of humanity in the dominance of superior technology, but an attempt to build coexistence as the only meaningful way of coherence of the analogue and digital world.[2]

If therefore, we try to understand the non-human by starting with the human, this is self-explanatory. After all, Heidegger in *Being and Time* tried to show that the meaning of Being in the horizon of original temporality can only be explained if the meaning of the essence of being there (*Dasein*) is first shown in its existential-ontological spheres of the project, care and openness towards the coming future. What is different, alien and foreign to the human being is not possible to think about itself, from the own logic of the thing itself, *par lui-même*, but only in a roundabout way. The non-human is therefore no longer just a pure negation of man's ability to spiritually-biologically rule his existence, system and environment of beings as such. There is something uncanny (*Unheimlich*) in that matters. Namely, if a man cannot be the result of any biological-animal intervention and reduction of his essence to the non-human in the sense of things and objects, then his ontological-temporal openness encompasses a completely different condition, both from the

---

[2] If for Simondon, the machine in the cybernetic "third order" appears precisely as a self-propelled and self-managed mechanism of purposeful action, and man has only the function of overseer and mediator of the process of the autonomous "concert of machines", it should be obvious that there can no longer be the talk of superiority-subordination in terms of the articulation of power in the physical-biological-psychosocial framework. Therefore, man and machine are no longer determined from a metaphysical point of view, but from a cybernetic set of relationships between living systems and the environment. Living systems are possible in the technical world but only when they incorporate technical objects in the same way as other living beings - plants and animals. When onto-genesis becomes techno-genesis, we enter a period of the disappearance of all previous standards of becoming a "Being", "beings" and "essence" of man. Anthropology that would adequately correspond to Simondon's utopian demand set back in the 1950s by the "philosophy of individuation" truly no longer corresponds to what this term meant for Kant, but even more so for many social scientists (sociologists and ethnologists) during the 20th century. Instead of the science of man, which in one way or another, logocentric or anti-essentialist, still considers man "the measure of all things", Simondon proposes concretization (*concretude*) or the openness of the machine itself as a condition for the possibility of a new homeostatic, metastable, informational balance. (Simondon 1994)

reduction to God in the sense of pure spirituality and incorporeality, as well as from the reduction to technology in the sense of pure applicative objectivity that arises from the subject-substance of modern sciences. God and technology are ways to metaphysically understand the notion of Being as *téchne* and *poiesis* from Plato to the end of metaphysics in cybernetics. (Severino 2016, 208-209)

When we have this in mind, we can see that the problem with the understanding of the non-human is primarily that what is in the *posthuman condition* can no longer be measured by the parameters in which man traditionally thinks and expresses himself metaphysically. What are these parameters? Undoubtedly, the autonomy of the human essence was still for modern speculative thought from Kant to Hegel a question of the sustainability of anthropology in a pragmatic way. Man as a subject of knowledge of the external world was considered as a substantial set of reflective practices. Even for Blaise Pascal, the so-called human dignity consisted only of rational thinking. In other words, rationality becomes a condition for the possibility of human existence in the technical world, which at the same time means that science and technology are the key parameters of the human possibility of progress and development in the historical movement towards reaching *object X*. The paradox is, therefore, that already with the entry into the scientific-technological age, the human, *i.e*. humane, became an artificial construction, and not a natural act of revealing the secret of Being. Therefore, the *posthuman condition* cannot be just the realization of the potential of Western metaphysics, which is completed in cybernetics. What is non-human as opposed to human is nothing non-human in the sense of animalism and the bestiality of a creature or monstrous being as the embodiment of evil in a sublime state of chaos and entropy. On the contrary, the *posthuman* denotes a becoming (*Werden, devenir*) of something that transcends any distinction between living and non-living, nature and culture, Being and beings. It is therefore not at all accidental that in posthumanism, instead of God in the metaphysical sense of the first substance understood, the creator of the world and man in it, we speak of *singularity* as the Omega-point of joining the onto-theo-cosmo-anthropological perspective in the form of an absolute *technosphere*. (Paić 2022a)

What does it mean? *Singularity* assumes that all becoming of the universe, as evidenced by astrophysics and cosmology, is a unique and irreversible event of the process of creation and destruction of matter and energy. Theories about black holes, therefore, represent a cognitive-theoretical path in solving the mystery by which supernova explosions turn into an implosion of dark matter. From a philosophical point of view, we encounter a clear movement of the very thing of thought from the speculative-reflexive analysis of the progress and development of the primordial and its form as *eidos* and *morphé* to the last station in life, regardless of whether it is about stars, plants, animals or man and his self-consciousness. The second law of thermodynamics with the concept of entropy shows us exactly what Nietzsche metaphorically named the statement that the *desert grows*. This means that the order in its stability inevitably collapses and becomes, as Simondon claims, only sustainable if it is in a state of *metastable equilibrium*. Simondon always thinks starting from (1) trans-duction (of Being); (2) trans-individuation (of beings) and (3) metastability (of information) in the transformation of the order of metaphysics. (Simondon 2005) The fundamental concept by which singularity can be measured in the cosmic-human perspective is the fundamental parameter of the *posthuman condition*. Of course, this should be the term with which cybernetics begins and ends, while at the same time the entire original structure of Western metaphysical thinking collapses with it. It is about the concept of *information* as an absolute measure of all measures in the age of the *technosphere.* Without information, there is no feedback loop in the life of homeostasis systems, and there is no possibility of external and internal control of the process of the creation of events, nor, finally, is there any possibility of interactive communication. It is therefore no coincidence that Norbert Wiener, the father-founder of cybernetics, said that the key, third concept for self-rule and self-management of systems and the environment after matter and energy becomes precisely the concept of information. Hence, informational entropy has become a categorical framework for the path to absolute singularity, in which the cosmo-biological merges with the anthropo-technological. (Rieger 2003)

If, therefore, Aristotle's definition of man as an *animal rationale* placed at the centre of the entire thinking of the West from the ancient

Greeks to the beginnings of cybernetics in the relationship between the concept of a living being or an animal that thinks with what makes such thinking operative and pragmatic, then this connection between the spiritual and the material, or biological has become in the age of artificial intelligence not only obsolete, but also fundamentally inappropriate for the explanation of the essence of that which is no longer non-human, nor can it be determined by something that only exceeds the limits of the human, as established by the concept of *transhumanism* (Sorgner 2020). Why? Precisely because a "man" can no be longer at the centre of the world, as it was in the modern era, in which he appeared as the subject of thought at the origin of all possible and real paradigms for understanding the relationship between Being, God and the world. Today's discursive congestion with the concept of the end of the *Anthropocene* represents a credible testimony to this (Stiegler 2018). The loss of the centre of man goes hand in hand with the gain of his so-called substitute being, that is, becoming something that can not be longer merely human as opposed to a non-human. When this distinction between man and machine is gradually lost, we encounter the hybrid status of all beings in the *posthuman condition* of the world's technological worldliness. The *animal rationale* can no longer be a condition for the possibility of all kinds of anthropology, from philosophical, structural, and cultural to cybernetic. In its place has come a concept that no longer elevates man to divine heights, as did Renaissance humanism, from which today's humanities originated. After the descending period of the establishment of metaphysics announced by Nietzsche with the term "superman as the meaning of the Earth" (Nietzsche 1999, 14), the dizzying realization of the project of becoming an interplanetary nomad, which instead of static inertia rests in the dynamic notion of *infinite speed,* began (Deleuze and Guattari 1991/2005, 198).

To be able to fly and transform our appearance in the physical sense, we must first of all become Other, as unconditionally different from the historical-epoch worldview. This also means that Being is replaced by the process of infinite becoming. *Instead of onto-genesis, techno-genesis is at work.* "Man" no longer "is" something in the sense of intentional consciousness, which, with Husserl, distinguishes between *noesis* and *noema*, consciousness and the object of consciousness. His

"fate" now lies in the contingency and emergence of life itself as a construction of events. If, in contrast to Simondon, who says that a robot does not exist, we can say that with the help of technological uploading, man is transformed into a producer of information and a user of an operating system in which this information circulates and becomes pragmatic knowledge only in a certain context and situation, then his "case" might be always possible to think as pure singularity and individuation. Therefore, the term *Homo kybernetes*, which corresponds to this "case" of a complete transformation of the ontological place of the human in post-history, can no longer be possible to determine from the linear development of a robot controlled remotely by artificial intelligence. On the contrary, this "man" in his cognitive self-organization and self-management becomes an autonomous synthesizing of the logic of the *technosphere.* Beyond all metaphysical definitions, its ontological fundamental feature is that it connects information and communication into a system of global-planetary code of networked life. In the tendency-latency, *Homo kybernetes,* therefore, become a visualized event of synthetic thinking as the production of virtual worlds. So, that means the final farewell to every dichotomy of human-non-human, living-artificial, natural-cultural. At the end of history, as a metaphysically determined journey from the first cause to the last purpose (*arché-eschaton*), there remains only the endless wandering of "man" who, instead of the secret of language, is mastered by the absolute visualization of life from its origin to its disappearance. The image now precedes language as the *technosphere* precedes the *biosphere* and becomes its only remaining pledge of redemption and deliverance from ruin.

2. Stefan Lorenz Sorgner on Transhumanism

The emergence of that great *Homo kybernetes* with which both posthumanism and transhumanism operate in their kinship and differences certainly cannot be understood without the following essential assumptions. I carried out my analysis of this problem explicitly in five volumes *of Technosphere*. (Paić 2018-2019) For this consideration, it will be sufficient to indicate the following interrelated main ideas. After

all, throughout the history of human thought within the circle of Western metaphysics, we encounter five fundamental, as Ernst Cassirer would say, symbolic forms in which the issue of the meaning of Being and time is crystallized. These are myth, religion, art, philosophy and science. Each of these matrixes of thinking is based on correlative concepts – myth based on primordial tales (*Sages*), a religion on faith in otherworldly gods and God, art on the representation and presentation of the world in images and the language of the senses, philosophy is based on the conceptual articulation of Being, beings and essence of man, and science on the idea of knowledge as a path towards reaching the absolute truth about *object X.* This quintet of historical-epoch thinking corresponds to the number of senses in Human. It is not at all surprising that the sense of sight will become key to what the technosphere signifies as the visualization of events in synthesizing the world as an a priori-a posteriori form of thought that includes calculation, planning and construction. But what would be now decisive for moving to the other shore beyond metaphysics is that the *technosphere* represents the possibility, reality and necessity of the rule of the cybernetic fourth of information, feedback, control and communication. This suspends and neutralizes the four causes in the traditional Aristotelian ontology, such as *causa formalis, causa materialis, causa efficiens and causa finalis*. At the same time, the *technosphere* as a "sixth sense" or a post-metaphysical symbolic form of thought transcends the entire history of metaphysics in all aspects of its appearance. Hence, it is its realization in the sense of the dialectical notion of the abolition of opposites and contradictions (*Aufhebung*) in the synthesis of the triad (*tollere, conservare, elevare*), but it is, even more, the becoming of a new process of thinking as immanence in the irreversibility of the movement of thinking itself. In this way, paradoxically, Hegel, Heidegger, Wittgenstein and Deleuze are reconciled, because what transcends any form of rootedness in space and time and thoughts intuitively starting from the rational core of artificial intelligence, can only be "founded" on *visual pragmatics knowledge*, and not in language as a telling of Being. Therefore, *Homo kybernetes* cannot be longer a "man" or a "being there" (*Dasein*), nor can his post-humanity be understood only by starting from the transhuman turn in the essence

of the cosmic-technological evolution of nature itself as the source of all kinds of cognitive-sensitive transformations.

To show how the logic of *Homo kybernetes* actions unfolds in the process of this "new" thinking that starts from the premise that we have always been cyborgs, we will briefly look at the analysis of Stefan Lorenz Sorgner, the leading philosopher of transhumanism today. As a reliable witness of this way of thinking, all the more so because he has behind him a study of Nietzsche and the concept of the superman, and because in a creatively stimulating way he combines the analytical and continental philosophy, Sorgner has in mind that the concept of humanism became not only inoperable already in the 20th century after all the interventions of Heidegger and Sloterdijk, but primarily because the emergence of the emergent digital technologies contributed to the grounding of metaphysics and the dismantling of speculative metaphysics with Hegel as the pinnacle of that direction of thought. What Sorgner calls "philosophical problems" in the discussion of transhumanism refers to the ontological, cognitive-theoretical and practical-production aspects of the change in the very essence of human substance and essence. In the book with the programmatic name *We Have Always Been Cyborgs: Digital Data, Gene Technologies, and an Ethics of Transhumanism,* we can discern how the relationship between cybernetics, biogenetics and bioethics is tried to be considered, conditionally speaking, since the categorical-conceptual order is established with the formal primacy of information. It should be clear that this is not about the Kantian way of establishing the triad of thinking, but about the necessary convergence between ontology, epistemology and ethical-political postulates. How, then, does Sorgner determine what the essence of transhumanism is concerning metaphysics in general and especially concerning different positions on the same "philosophical problem"?

Interestingly, Sorgner points out that, starting from man's desire for a good life since Aristotle's *Nicomachean Ethics*, the problem we face with transhumanism cannot be in the utopia of immortality, but in an attempt to lead such a life with improved health in the physical and cognitive sense.

> "Yet neither cryonics nor mind uploading seem to be the most promising techniques for promoting this goal. This being the case, it needs to be pointed out

> that immortality is not a realistic option. We cannot even conceptualize the notion of immortality on the basis of a naturalistic understanding of the world. Immortality implies either that humans cannot or that they must not die. Both options are absurd, if we think of the world on a naturalistic basis. In this case, the universe began with the Big Bang. It has expanded since then, and eventually, the expansion process will slow down so that the entire universe will come to a standstill. A different theory is that at a certain stage the expansion process will become revised into a contraction process that will end up in a cosmological singularity, a point of immense density. In both cases, there does not seem to be a realistic option for human beings to survive. (...) Hence, it is clear that immortality in its literal sense cannot be a realistic option for most transhumanists." (Sorgner 2022a, 6-7)

Immortality, therefore, is excluded as the realistic solution. However, it should be quite obvious that the aspiration to reach the perfect state of a good life led healthily and effectively cannot stop at the virtues of prudence (*phronesis*) and perhaps even stoic *ataraxia* in the face of the inevitable fate of being chained to nature and the biological constitution of the human species. Moreover, this version of transhumanism advocated by Sorgner tends to be both "realistic" and "anti-naturalistic", and at the same time cannot deny the fact that the concept of information in cybernetics becomes the result of digital constructivism. What is offered in the utopian horizon of posthumanism, on the other hand, for example with the postulate of singularity in the sense of the transition of the human to the *posthuman condition* by transferring the individuation of man as a person into the operative program of cyborgization and computer simulation, is not obvious to Sorgner the right measure between ontology, cognitive theory and ethical-political actions. Therefore, it is self-explanatory why the utopian dimension of transhumanism as a possibility of achieving immortality is rejected. Instead of this impossible goal, which from the point of view of the natural-cosmic evolution of the universe is even absolutely unattainable and not just absurd, Sorgner defend

> "a non-utopian version of transhumanism, as I regard utopias as extremely dangerous. Thereby, the present becomes sacrificed for a future which most probably will never be actualized." (Sorgner 2022a, 7)

If, however, we see transhumanism as a philosophical consideration of the possibility of crossing the boundaries of the human-too-human,

then it will not at all be surprising why the only two techniques in the utopian horizon of reaching post-humanity are related to the notion of immortality or the singularity of man as a species. Namely, both cryonics and uploading of the human mind are technological postulates, not actualized means of technological development at the current level of knowledge about the impossible. Moreover, these techniques seem mutually convergent because they assume the revival of a dead body and the transfer of its soul/spirit into the logic of *Homo kybernetes.* When we talk about the primordial notion of man, then we have in mind a mythical-religious image of the world in which this being is in mediation with other beings with the help of that natural-biological and technological approach to the supernatural, *i.e.* the divine, like the mythical Icarus as the prototype of human flight into space or like a centaur as a mythical creature composed of two substances. Transhumanism is to that extent the logic of the transgression of metaphysics at its end. What these opinion crosses are the boundaries of the animal-human body and the human soul-spirit in creating a hybrid complex of what is no longer human in the traditional sense of metaphysics, but which is not even a posthuman being similar to a cyborgized creature from SF-dystopia. It becomes, therefore, a transitional state, between two mutually close and distant worlds. For Sorgner, the analogue age of the *Anthropocene* and the digital age of the *technosphere* turn out to be the Carbon Age and the Silicon Age. (Sorgner 2022a, 22-109) Of course, these are metaphors that point to a distinction in the understanding of transhumanism, starting from what belongs to the technological postulate and what belongs to bioethical judgments about the limits to which the imaginable installation of implants in the body can go, as well as surgical-pharmacological interventions to improve the structure of the human body.

However, everything that has been said is still not enough for us to start with a true philosophical discussion about the essence of transhumanism in the ontological-epistemological sense. Why? Because there is no answer to the question of why we have always been cyborgs. Sorgner claims that

> "Our turning into cyborgs is a development which has taken place since we became *Homo sapiens.* Acquiring a language is the first human upgrade. (...) The

> term 'cyborg' means cybernetic organism. The word *cyber* comes from the ancient Greek kybernήτης, which means helmsman or pilot. So a cyborg is a governed, a steered organism." (Sorgner 2022a, 9)

The governed simultaneously governs by being connected in interaction with other interactors in the network of life processes. The cybernetic way of thinking exists always in dualities, and not in trinities as in the previous metaphysics. The logic of the digital world, after all, rests on the binary code 0-1. Cyborgization, therefore, means programming the world as a contingent case. It is possible to manage systems and the environment only when the condition of the possibility of cybernetic thinking is realized in language as a pragmatic form of life. Such a language becomes a visual conceptualization of the relationship between structures and functions. Since the image in the digital environment no longer shows or represents anything, the action of metaphysically established thinking in philosophy, art and sciences have been suspended. The image cannot be a function of something external, nor does it arise from an otherworldly source like God's hidden presence, as it was during the reign of the referential frame of myth and religion. Instead, the image shows and expresses what happens in the process of the emergence of life itself in the phenomenon. The connection of genes and information as the emergence of new life determines, for example, the essence of bio-art and trans-gene art. (Sorgner 2022b, Paić 2022b, 191-219) When we have all this in mind, it should be not difficult to understand that the essence of transhumanism in the philosophical sense necessarily appears within the framework of what belongs to posthumanism, which is the realization of the essence of the *technosphere in a posthuman condition*. If for Nietzsche the last man was a preparation for the arrival of superman as the meaning of the Earth, then in the case of *Homo kybernetes* we find ourselves in a position of thinking that can no longer be philosophical or artistic, nor scientific research into the first reason and the last purpose of the human-too-human in the age of the global-planetary development of the technosphere. Cyborgization denotes the last instantiation of the production of technologically perfected and improved "beings" that have become hybrid structures and applications of genes and information in everyday life.

Why are we talking about "the last instance" of this border crossing between the human and the posthuman? Because what inevitably comes can no longer be non-autonomous and non-self-regulating. What is coming as a technological singularity must necessarily be freed from any external management of the system and environment. And this means that technosphere thinking requires overcoming the limits of metaphysical thinking as a whole starting from the introduction of the concept of *autopoiesis* (Paić 2021, 431-456). Sorgner, as we can see, from the so-called realistic perspectives removes the futuristic-utopian techniques that one direction of transhumanism explicitly promotes. These are cryonics and uploading. Body and mind are no longer considered Cartesian, but anti-essentialist. If the approach to the new cybernetic physicality is derived from a different relationship to the term "enhancement", then here we must distinguish between technique as a means and emergent technology as a means-end of the entire process of creating a cyborgized *posthuman condition*. Since the invention of language at the dawn of human evolution was the beginning of what we call *animal rationale*, one should be extremely careful with the assumption about the future of language in a *posthuman condition*. The reason is that the pragmatics of *know-how* thinking only needs formalized rules of use for a thing or set of events to be realized virtually. Language thereby loses its meaning or purpose in what Heidegger attributed to the telling thought of human existence (*Dichten*). (Heidegger 2002) Instead, all humanities and what Sorgner calls *posthuman art* need a different linguistic articulation of a meaningful horizon in the interaction of system and environment. Such a changed character of the language in the logic of new media instead of syntactic-semantic clarity of speech becomes a kind of *white noise* in communication. In this way, the pictorial structure of language in the age of the technosphere, as already announced by Ludwig Wittgenstein in *Philosophical Investigations*, becomes a *language game* (*Sprachspiele*). (Wittgenstein 1998)

What seems particularly interesting in Sorgner's consideration of the philosophical problems of transhumanism is his position on the so-called nihilistic perspective of this opinion and the so-called positive pessimism. Of course, he disclaims that all directions and tendencies in this interdisciplinary movement today are "nihilistic, pessimistic-positive",

but according to the statistical analysis, they are mostly "types of naturalism, atheism or secularism" (Sorgner 2022a, 11). Schopenhauer was the only one who took pessimism seriously, even asserting that art in its deep ontological truth is imbued with pain and suffering, and its sublimity becomes the last possibility of deliverance from the abyss of life's disappointment and death's agony. But, as we know, Nietzsche assigned to art the attributes of the most magnificent possible turning point in the effort to overcome the nihilism of Western metaphysics and therefore removed any pessimism from this doctrine of self-affirmation of life. Overall, Sorgner shows how obvious paradoxes and aporias occur with transhumanism at the level of "worldview". The denial of God and complete naturalism in the development of assumptions about the possibility of man's transition beyond biological limits at the same time may lead to a precisely nihilistic attitude about the historical-epochal dimensions of this transformation of life. Namely, although pessimism is mostly the negative attitude about the future of the human species and its prospects, the positive one comes from the trust in the power of the mind as "common sense" which is crystallized in what modern technology provides. In this way, namely, it is shown that the key point of the synthesis of mind and technology is that opinion which, although Sorgner does not carry it out in his analysis and although he does not have a mostly positive attitude towards Heidegger, tries to connect the very beginning and modernity starting from the metaphysical frameworks of thought in which cybernetics was born.

Heidegger, therefore, said the following about it in the most concise way:

> "Therefore, the circular rule applies as a basic feature of the cybernetic world. The possibility of self-management, automation of the movement system rests on this. In the cybernetically represented world, the distinction between automatic machines and living beings disappears. It is neutralized in the undifferentiated event of information. (...) Humans are also drawn into this monotony of the cybernetic world. And even in a distinguished way. Because in the point of view of cybernetic representation, man has his place in the widest circle of rules. According to the modern idea of man, he is, namely, a subject who relates to the world as an environment of objects by processing them. Hence, all kinds of changes in the world come back to a man. The subject-object relationship is, cybernetically represented, a mutual relationship of information, and feedback

loops in a distinguished circle of rules, all of which can be described under the title 'man and the world'." (Heidegger in Jaeger and Lüthe 1983, 16)

From what Heidegger said about the essence of cybernetics, which becomes in the technosphere a planetary-global space of virtualization and implosion of information, can we arrive at the possibility that what transhumanism has as its postulate, regardless of its differences in the so-called directions, open as an *autopoiesis* in which the possibility of preserving both "man' and the world" remains beyond their almost inevitable reductions? This question seems decisive for the prospects of the coming philosophy. Any distinction between humans and non-humans presupposes that a third exists (*tertium datur*). This "Big Third" throughout the history of metaphysical thinking has always been thought of as *God, Nature, causa sui, superego, sublime object X.* But isn't it astonishing that even in Hegel's science of logic, the third is nothing more than a subject-substance or a thing of thought that sets itself in motion? After all, isn't freedom precisely that autopoietic event of thought that sets in motion every corporeality understood in one way or another, just as it sets in motion every soul-spirituality (*psyché-logos)* understood in one way or another? Isn't freedom a condition for the possibility of absolute in one way or another?

3. Freedom beyond Absolute

Freedom as absolute? This is exactly the essence of posthumanism in a philosophical sense. Take the so-called destiny into one's own hands, to change the necessity of historical events into the possibility of control over contingent events in the process of becoming (*Werden, devenir*) different from the Same. Almost all ardent theorists of post-and-transhumanism programmatically, manifestly and triumphantly, show that thanks to the progress and development of the mind in history, the time has come for humans to free themselves from the shackles of the historical tragedy of its Being. Mind, therefore, is identified with knowledge, which in the form of technoscientific construction of new beings fundamentally changes the found natural-biological order of relations, which was almost unchanging for metaphysics from Plato to

Hegel. That this is so might be evidenced by the attitude of most representatives of technoscience, who talk about the fact that in laboratory conditions, thanks to synthetic biology, an artificial being such as the so-called oncomouse has been generated. All this points to the real possibility of introducing artificial life into the upcoming aspects of the complexity of nature and human habitation on Earth and beyond its natural-biological boundaries. Moreover, that event represents a turning point in the understanding of the difference between the human and the posthuman, and in the above example also in the difference between pre-human, human and post-human. History in a different, post-metaphysical sense starts from the logic of artificial intelligence and irreversibly refers to the cosmic-natural evolution of the animal-man-machine.[3]

Freedom as the thought-life autonomy of human existence does not contradict this event of absolute techno-genesis. This was already clearly seen by Gilbert Simondon when he defined the essence of substance in a completely different way, which was traditionally Cartesianly explained by inertness in the expansion of matter, the *res extensa*. We have seen that the problem with the cybernetic paradigm is that it is the realization of metaphysics, on the one hand, and on the other, its transgression by introducing the principle of *tertium datur*. Matter can no longer be left to vulgar materialism in this new thinking, just as it cannot be idealized with some spiritualized model of vitalistic materialism. Matter, therefore, has not only an energy potential, without which there would be no complex network of interpenetrating bodies in motion, nor

[3] "The most important reorientations of recent science does not arise at the methodological level but at the ontological one. It concerns the 'techno-scientific turn' regarding the objects of scientific research. Many objects of scientific study are no longer part of untamed nature; they are produced artificially. The technoscientific turn at the ontological level means that most of the entities and processes addressed by scientific research are not found or uncovered, but are rather created. Laser processes, chlorofluorocarbons (CFCs), nanotubes, or layered systems of ferromagnetic materials and chromium in a magnetic field (in which giant magnetoresistance makes its appearance) would have hardly ever come to exist by nature left to its own devices. They are deliberately produced by humans but explored, like objects generated by whims of nature." (Carrier 2011, 51-52).

any freedom of movement of monads in a reestablished harmony à la Leibniz. To be able to think freely in the sense of *causa sui*, so that human thinking in the form of living-and-cybernetic physicality could reach the idea of absolute self-consciousness not only of man but also of a machine as a cyborgized "creature" with a soul-spirit, it is necessary to bring that connect the concept of information with the concept of *autopoiesis*. In this way, two necessary cybernetic reductions take place: the one that refers formally to the metaphysically understood *res cogitans* and the one that affects the *res extensa*. This necessary anti-substantialism Simondon tried to be carried out radically in his reflections on the relationship between man and machine. He claims the following:

> "Information, on the one hand, is something infinitely variable and something that, if transmitted with minimal loss, allows energy to be effectively conserved so that it is in no way reduced to the rank of possibilities... But information, on the other hand, is something that, by transmitting go beyond the level of pure chance phenomena, such as white noise and heat stroke; therefore, information is something common, which has its own position, a certain sphere and a certain stereotype that distinguishes it from pure chance... This contrast represents a technical antinomy that seems to be a problem of philosophical thinking: information is like a random event, and yet it differs from it. An absolute stereotype, which excludes any novelty, therefore excludes any information. However, the difference between information and interference is based on the reduction of uncertainty limits." (Simondon 1958/1969/1989, 234-236)

Neither the mind nor the body is separated by an abyss anymore. Therefore, information is neither a form nor a substance in the Aristotelian sense, but that which interferes, which creates new information and thus produces events as pure contingencies. When the mind in its absolute freedom thinks in the human form of self-consciousness about an *object X*, it is not mere contemplation and speculation about the Kantian thing-about-itself (*Ding-an-sich*) that affects our senses. Thinking cannot be just the indifferent observation of a being in nature in the sense of the mere objectivity of the object, on which the ideal of the Greek paradigm of knowledge about the world rested. Instead, already in the new Era, thinking becomes a pure construction of the object of cognition and hence the subject-substance way of producing new beings or objects in the world. But with the advent of the cybernetic paradigm, due to the concept of information and the possibilities of machine information

processing, we are no longer dealing with something "in" the world and we are not thinking "about" something intra-worldly or supra-worldly. *We*, as a union of natural mind and artificial intelligence, produce objects *X* that exist only when changed substances and the appearance or form of these substances are joined by techno-scientific construction. This process in which the power of the technosphere reigns auto-poetically as a new absolute is therefore not nihilism without the meaning of history and the complete loss of the human-too-human. On the contrary, this process with the potential possibilities advocated by radical transhumanism such as cryonics and uploading techniques leads to the postulate of technological singularity as absolute freedom of a new way of thinking beyond metaphysics in general. It seems as if we are encountering something at the same time monstrously fascinating and almost uncanny (*Unheimlich*). Namely, we are used to defining freedom as something that is neither a Being nor a being but as a condition for the possibility of their relationship in common existence. Thus, for example, in Schelling's definition of the difference between the essence of human freedom and human freedom itself, the question of freedom of will, Good and Evil, personality, etc. appears. However, this difference can only be thought of when it is taken into account that absolute causality in one essence allows all others only unconditional passivity.

> "Through freedom a fundamentally unlimited power is asserted next to and outside of divine power, which is unthinkable according to these concepts. As the sun in the firmament extinguishes all the lights in the sky, even more so does infinite extinguish every finite power. Absolute causality in One Being leaves only unconditional passivity to all others. This entails the dependence of all beings in the world on God, and that even their continued existence is only an ever-renewed creation in which the finite being is produced not as an undefined generality but rather as this definite, individual being with such and such thoughts, strivings, actions and no others." (Schelling 2006, 11-12)

If this position of Schelling's is understood as one of the essential insights into what makes freedom the abyss of human existence and cannot be reduced to the definition of man by Aristotle as *animal rationale*, then we are faced with something even more monstrously than what Schelling himself could have anticipated when he spoke about free will, Good and Evil, personality, etc. Why? If, namely, for Simondon, the robot does not exist, but has a different way of being, then from the point

of view of absolute metaphysics, we should see how the transmission of human freedom to the possibility of the freedom of the non-human in the *posthuman condition* arise. All beings in the world are a living unity of physical-mental-spiritual synthesis of the natural-biological and autopoietic in the sense of homeostasis. When, therefore, Simondon assumes in his defence of the cybernetically understood coexistence of animal-man-machine a kind of secularized idea of the Good beyond all pantheism and atheism, he is talking about the belief that progress and development in history are only possible from the pre-established harmony of the organic and the mechanical. But what if, for the understanding of the essence of the technosphere as absolute freedom of self-movement and self-management of life, it might be impossible to neutralize and suspend what is the negation of freedom of will, and rests on it – namely, evil in its pure contingency? What, then, if freedom in its absolute construction of events is indifferent to good and evil, and shows itself only in the pragmatic form of one's own self-consciousness that combines the living and the artificial in something that formally has the characteristics of evil and destructive tendencies against an evolutionarily lower stage of development in the form of animals and humans? The famous astrophysicist Stephen Hawking warned about this near the end of his life, expressing his fear of uncontrolled progress and the development of artificial intelligence. Now we see that the philosophical problem with post-and-transhumanism becomes a problem that was posed precisely by Schelling with the speculative thinking of German idealism. This is the problem of freedom as an absolute that starts from unconditional causality and thus subjects everything else to unconditional passivity. The technosphere as an autopoietic system of absolute freedom of information production in tendency-latency represents also a system of information control and as such has the same logic of action as the "human freedom" of mental action. However, this is not the essence of freedom, because the information can not be longer reduced to the logic of organically understood causality and accordingly to teleology, *i.e.* the ultimate purpose of all events. Freedom is neither a Being nor a being, and its "essence" is not reduced to absolute causality, but the event of an onto-cybernetic network of pure contingencies.

Evil is a first-class philosophical problem because it connects Being-God-world-and-man. However, evil is neither "in" nor "outside" of the world, but in the midst of it as a constellation of relationships that in some contexts and situation creates evil creatures and violence as a threat to the destruction of the world in general.

How then is it possible to overcome that which affects the fate of the world, which man cannot abolish and which was considered a tragic event by the Greeks since ancient times, precisely because a man had the freedom of will to autonomously choose his destiny, like, for example, King Oedipus who defiled his own mother and killed his father? It will be said, that he did it unconsciously, in ignorance of the causes and consequences that led to the tragic act and the realization of the truth. That's exactly the problem. Unconsciousness and ignorance are excluded in the logic of the operation of autopoietic thinking systems. When information has an interfering influence on the interactors in communication, then this might be the way to create what Simondon calls a *metastable equilibrium*. What is the meaning of that attitude except that, paradoxically, the unstable in the form of new chaos now enables the system to balance, and this is only possible when entropy as the second law of thermodynamics regulates processes in the cybernetically-based relationship between man and the world. The result is that we do not rule the environment in the sense of absolute control over evil as an event capable of destroying not only the biosphere but also the technosphere as such. If it seems to someone that this could be a post-metaphysical "Big narrative" about the *war of the worlds*, then they might be truly right. The philosophical problem with the emergence of the *posthuman condition* is not in any relationship between humans and non-humans, but in how freedom in its essence opens up in the realized world of the technosphere where the logic of absolute causality and teleology no longer rules. Instead, we encounter a metastable game of chance in which pure contingencies produce the coming that is never devoid of the possibility of being unpredictable and beyond the control of the human mind. *Homo kybernetes* denotes the last station on the way of human evolution and at the same time the end of metaphysics in cybernetics in general. It only means that for the first time in history, freedom has become a contingent necessity of thought,

and not always an eternal fate that the gods or God assigned to man as a reward and punishment for his dignity as an irreducible being under the stars.

REFERENCES

Carrier, Martin (2011). "Knowledge is a Power, or how to Capture the Relationship between Science and Technoscience." In *Science Transformed? Debating Claims of an epochal Break*, in A. Nordmann, H. Radder and G. Schiemann (eds.). Pittsburg: Pittsburg University Press, 43-53.

Deleuze, Gilles, & Guattari, Felix (1991/2005). *Qu'est-ce que la philosophie?* Paris: Minuit.

Haraway, Donna (1990). *Simians, Cyborgs, and Woman: The Reinvention of Nature.* London-New York: Routledge.

Heidegger, Martin (1983). "Die Herkunft der Kunst und die Bestimmung des Denkens". *Distanz und Nähe: Reflexion und Analysen zur Kunst und Gegenwart*, in P. Jaeger & R. Lüthe (eds.). Würzburg: Konigshausen + Neumann, 11-22.

Heidegger, Martin (2002). *Was heisst Denken?* Frankfurt/M.: V. Klostermann.

Heidegger, Martin (2018). *Sein und Zeit.* GA. Vol. 2. Frankfurt/M.: V. Klostermann. 2nd ed.

Nietzsche, Friedrich W. (1999). *Also spach Zarathustra. Ein Buch für Alle und Keinen*. Sämtliche Werke. KSA. W. de Gruyter: Berlin.

Paić, Žarko (2018-2019). *Technosphere*. Vol. I-V. Zagreb: Sandorf and Mizantrop.

Paić, Žarko (2021). "Thinking and creation: *Arché-Ergon-Autopoiesis.*", In *Nihilism and Contemporaneity – In Nietzsche's footsteps*. Zagreb: Litteris, 431-456.

Paić, Žarko (2022a). *Art and the Technosphere: The Platforms of Strings.* Newcastle upon Tyne: Cambridge Scholars Publishing.

Paić, Žarko (2022b). "Metaphysics and Cybernetics: About the *Technosphere* or from the thing of thought to the thing that thinks."*The Technosphere as a New Aesthetic*, Žarko Paić (ed.). Newcastle upon Tyne: Cambridge Scholars Publishing, 1-36.

Paić, Žarko (2022b). "Synaesthetic Without Sensitivity? The Body as a Technological Construction." In *The Technosphere as a New Aesthetic*, Žarko Paić (ed.). Cambridge Scholars Publishing: Newcastle upon Tyne. 2022b, 191-219.

Rieger, Stefan (2003). *Kybernetische Anthropologie: Eine Geschichte der Virtualität.* Frankfurt am Main: Suhrkamp.

Schelling, F. W. J. (2006). *Philosophical Investigations into the Essence of Human Freedom*, trans. and with an introduction by Jeff Love & Johannes Schmidt. New York: SUNY.

Severino, Emanuele (2016). *The Essence of Nihilism.* London-New York: Verso.

Simondon, Gilbert (1958/1969/1989). *Du mode existence des objets techniques*. Paris: Aubier.

Simondon, Gilbert (1994). *Une pensée de l'individuation et de la technique*. Paris: Chatelet.

Simondon, Gilbert (2005). *L'Individuation à la lumière des notions de forme et d'information*. Grenoble: Jerôme Millon.

Sorgner, Stefan Lorenz (2020). *On Transhumanism*. Penn State University Press.
Sorgner, Stefan Lorenz Sorgner (2022a). *We Have Always Been Cyborgs: Digital Data, Gene Technologies, and an Ethics of Transhumanism*. Bristol: Bristol University Press.
Sorgner, Stefan Lorenz (2002b). *Philosophy of Posthuman Art.* Basel: Schwabe Verlag.
Stiegler, Bernard (2018). *The Neganthropocene*. London: Open Humanities Press.
Wittgenstein, Ludwig (1998). *Philosophical Investigations.* Oxford: Willey-Blackwell. 4th ed.

HUMAN GENOME EDITING TECHNOLOGIES IN THE CENTER OF CURRENT TRANSHUMANISTIC DISCOURSE ON HUMAN ENHANCEMENT

FRANC MALI[1]

Abstract

The article deals with certain challenges connected with the progress of human enhancement technologies, where genome editing technologies form part of them. Today, the idea of human enhancement is the subject of considerable debate among various streams of transhumanist thinkers. The author rejects pessimistic and dystopian views on human enhancement technologies that see these technologies as one of the main risks for modern civilization, and instead defends the opposite view. It is claimed that although such technologies continue to be relatively poorly defined, they do and will belong to the 'good side' of current (future) technological progress. Namely, in the context of pragmatic and moderate transhumanism, human enhancement is primarily viewed as a way of improving the quality of human and social life. The article mainly refers to human genetic enhancement, noting that the invention of CRISPR/Cas9 technology about 10 years ago paved the way for a revolution in gene research and gene therapy. Before the emergence of CRISPR/Cas9 technology, bioethical discussions had mostly concentrated on classical bioethical dilemmas (autonomy, privacy, etc.). Yet, the situation has changed radically in the last 3 years when the first practical misuses of CRISPR-Cas 9 technology (the He Jiankui affair) in germline genome editing were reported. Although the author agrees with views calling for more scientific and broader social control of germline genome editing, strong support is given for the claims made by moderate and pragmatic transhumanists that modern democratic societies must ensure that objective and unbiased assessments of the benefits/risks of all types of human enhancement technologies are retained.

Key words: *human enhancement, genome editing technology, bioethics, moral entrepreneurship, CRISPR-Cas 9 technology, He Jiankui affair.*

---

[1] University of Ljubljana, Slovenia. Email: <franc.mali@fdv.uni-lj.si>

Introduction

Overly critical views on the transhumanist discourse can often be found today. It seems the aversion to transhumanism is firmly linked to the threats before the current and regarding the future progress of human enhancement technologies. In the thinking of various social groups, one frequently encounters deeply rooted, considerably pessimistic, even dystopian views on the whole future of technological progress, not simply the specific technology of human enhancement. These pessimistic and dystopian views sometimes even lead to technophobia, *i.e.*, a hate of technology that rises above a reasonable consideration of the possible negative effects of technological progress. They tend to present technology as one of the main risks for modern civilization. In these views, newly emerging technologies are accompanied with metaphors like "Brave new world", "Frankenstein's science", "playing God", etc. Not just ordinary people, but even scientists, experts, and the media are inclined to such a negative standpoint on these technologies.

In this contribution, the opposite view is defended. The basic assumption is that there is needed to overcome the (unproductive) disqualifications proposed in modern transhumanistic discourse. Namely, if technological progress is seen not simply as something that develops with a deterministic logic of its own, but as something that depends considerably on the overall social context, such kinds of technophobia can be shown to be unfounded. A look back at history reveals that many earlier fears about technology proved to be exaggerated. The resistance to new technologies gradually disappeared when they in the meantime became widely accepted in society. After the new technologies became established, there was no longer a need to repeatedly keep finding a new consensus on their acceptance.

Today, the controversial debates concerning transhumanism must be put on the right track. Unlike the simplified disqualifications found in transhumanist discourse, it is better to understand transhumanism as an effort for ensuring human development and social well-being on the grounds of technological progress. The philosophy of transhumanism has advanced greatly ever since Julien Huxley first uttered that word (Huxley 1968). Today, transhumanist discourse is increasingly exploring

the boundless positive possibilities of technological progress for generations of people now and in the future. The opinions of the frontrunners of transhumanism are far from an uncritical presentation of technological progress. They also stress possible misuses of the newly emerging technologies and pay substantial respect for individual safety, freedom, privacy, etc. (Newton 2019). As Nick Bostrom, a leading representative of transhumanism, wrote in his article Transhumanist Values: "Transhumanism advocates the well-being of all sentience, whether in artificial intellects, humans, and non-human animals (including extraterrestrial species, if there are any)" (Bostrom 2005, 4).

The world-renowned German philosopher Stefan Lorenz Sorgner also defends the realistic and pragmatic concept of transhumanism. Instead of far-flung future technological hype, he presents the role of technological progress primarily as a means of gradually contributing to the social welfare of society. In this sense, his transhumanism has nothing together with disreputable thinking such as the idea that human beings may become immortal, etc. For him, the idea of immortality functions in the best case as a rhetorical metaphor. In all of his most recent books, instead of the idea of the immortality of human beings he places in the forefront the idea of the happiness of human beings on the planet (Sorgner 2022; Sorgner 2021; Sorgner 2016).

The thinking of certain leading transhumanists today makes it hard to agree with the critics of transhumanism who declare that this philosophy is in the best case a pseudo-scientific ideology oriented to making baseless speculations about a distant future. The worst tendency in the discussions on transhumanism is insisting on fixed ideological statements, which have remained notwithstanding the open dialogue. This is also the case with discussions regarding the progress of human genome editing technologies. These technologies give the option of radically changing the (biological and social) conditions of human beings and their lives. Such possibilities options are not just projections relating to distant future societies, but already appear today. Therefore, while considering the issues of human genome editing technologies it is necessary to take more holistic views into account. In this contribution, some aspects of human genome editing technologies are highlighted, as

connected with broader questions concerning the role of newly emerging technologies in current human enhancement processes.

Role of Newly Converging Technologies in Enabling the Shift to Human Enhancement

One of the main ideas behind the pragmatic idea of transhumanism is that strategic interventions using newly emerging technologies into both our reproductive capacity and nature's selection processes will permit us to direct the course of human evolution. Newly emerging technologies have started to show their dominance over spontaneous biological processes and are finally overtaking biology as the science of life. These processes of the "technization of nature" are best seen in current genetic engineering research (Endy *et al*. 2014; Calvert 2010).

Genetic engineering no longer relies on ad hoc technological procedures. Instead, it is transforming into a systematically pursued undertaking where the operations of biological systems can be described in a standardized form with the use of technology; namely, a precondition for making them measurable, comparable, exchangeable, etc. Yet, human genome editing technology does function in this way. For example, the relative ease of applying the latest ground-breaking technology – CRISPR-Cas9 technology – means that it can be quickly incorporated using the equipment, skill sets, practices, and routines of a wide variety of research groups, in turn opening up fresh experimental possibilities. Finally, the biggest expectation concerning the entire field of genetic engineering ever since the 1980s has been that the rearrangement of organic life will entail routine and standardized procedures. Drew Endy, a leading synthetic biologist, described it this way: "Biotechnology remains complex because we have not so far made it simple" (Endy 2005, 451). The processes of technicization are responsible for the nostalgic image of the biological world and for nature slowly disappearing. While throughout humankind's long intellectual history the main idea about nature was that it conceals its secrets from human view, it seems that today nature no longer has any secrets. Nature has largely been reduced to matter that can and must be technically manipulated in manifold

ways. Modern biogenetics and biomedicine can claim the most merit for this new conception of nature. Contrary to the outdated belief that altering the genes of a human amounts to 'a violation of its essence' (as constantly complained by conservative bioethicists), genetic research and therapy are no longer understood as an unnatural intervention in the fixed essence of human beings. Accordingly, they are not just part of evolution, but one of the engines driving it. Stefan Sorgner claims that there is "... a close connection between transhumanism and two assumptions, *i.e.* evolutionary (naturalistic) thinking and belief in technological progress" (Sorgner 2021, 84).

Today, several potentially valuable applications for current and future society are emerging as the newly emerging technologies become unified, *i.e.*, biotechnology, nanotechnology, information technology, and cognitive technology. These applications cover practically every area of human life. We are living in a time of more intense processes aimed at unifying scientific and technological knowledge – the crucial factors in encouraging human enhancement. To describe these processes of merging newly emerging technologies, the umbrella term "converging technologies" is often used. The first and most comprehensive elaboration of the concept of converging technologies was provided about 20 years ago by William S. Bainbridge and Mihail C. Roco (Bainbridge and Roco 2002). Rogers Hollingsworth has dealt extensively with the converging technologies concept and claims that today they are in the development phase, which requires the building of a strong common scientific-technological core made up of shared theoretical frameworks plus a common stock of models and mechanisms that integrates a broad range of domains normally analyzed by different scientific disciplines and technological areas (Hollingsworth 2006).

The initial conceptualization of the converging technologies concept was followed by rapid progress and its operationalization. Today, these processes may best be described as a combination of enabling scientific discoveries (genetics, nanoscience), techniques (informatics, gene splicing), and advances in associated tools (computing power, robotics) that is strongly accelerating the basic science involved and leading to practical technological applications across a breathtaking range of subjects, from human health through to material science

(for more, see Fuller 2012). The convergence of various disciplinary fields is especially important for genomic science and technology that support several types of genomic therapy. For instance, today it is extremely important to integrate genomic data into electronic health records, *i.e.*, to combine the knowledge of all scientific fields to allow researchers to improve the clinical impact of the genomic data, demonstrate their utility, and make it accessible for clinical decision support tools (Venner *et al*. 2021). In genomic sciences dealing with extending the healthy lifespan of human beings, strong processes involving the cross-fertilization of scientific disciplines have also been noticed. Namely, in the context of transhumanism and human enhancement, prolonging the human lifespan is no longer simply a purely theoretical topic. It is becoming a practical question related to the further interconnectedness of scientific and technological areas where synthetic biology, nanotechnology, and micro-fabrication, along with advanced computational, modeling, and visualization capabilities, can ultimately be enlisted to solve the problem of degenerative ageing and how to extend the healthy lifespan of humans (Stambler 2019).

The concept of converging technologies refers to the radically new possibilities that are opening up for modern societies. While in many centuries in the past small-scale convergences were underway in scientific and technological development, the current trans-disciplinary convergences are radically altering the character of modern science and technology. These trends towards the unification of scientific and technological knowledge hold huge implications for the current and future social life of humans. They are sometimes described in hyperbolic terms. Kurt Kurzweil is a well-known transhumanist thinker who seems to advocate that such ideas of converging all spheres of life are the most provocative. He assumes that already in the near future an evolutionary process in which the union of all entities of the world (not only scientific knowledge) will occur. At this point of singularity, the knowledge and skills embedded in our brains will be combined with the vastly superior capacity, speed, and knowledge-sharing ability of our own creations. As he states: "Just as a black hole in space dramatically alters the patterns of matter and energy accelerating toward its event horizon, the impending

singularity in our future is increasingly transforming every aspect of human life" (Kurzweil 2005, 7).

Idea of Human Enhancement Supported by a Pragmatic and Moderate Transhumanist Discourse

In contrast to conservative bioethicists who seek to guarantee the permanence of human life and the fixed human essence as it is broadly recognized today, even moderate transhumanist thinkers support the proactive stance taken by innovation-oriented individuals. Their overriding maxim is that the future must be made by active and innovative human activity, which in itself assumes that, regardless of how normal and successful an activity may be, there are always better ways of achieving the same ends, if not better versions of the same ends. Steve Fuller and Veronika Lipinska invented the concept of "moral entrepreneurship" to describe this proactive stance that people adopt (Fuller and Lipinska 2014). According to Fuller and Lipinska, the proponents of moral entrepreneurship are future-oriented innovators who are leading the way into the unknown territory of technological progress. Their purpose is not to do only pure research with the intention to understand the ways in which human and other living systems will develop in the future. Their purpose is instead chiefly to produce new types of human enhancements in an evolutionary way. The crucial message of arguments claiming that we need "moral entrepreneurship" is that future-oriented innovators should be free to use their creative efforts to contribute to social welfare on the grounds of redesigning the biological conditions of humans, but without interfering with the rights possessed by someone else.

In the context of pragmatic and moderate transhumanism, human enhancement is mostly seen as a way of improving human and social life. It seems that this variant of transhumanist thinking is consistent with the fundamental assumptions made in bioethics literature more generally. For example, both mainstream bioethics and transhumanists encourage conceiving of the self within a framework of choice and agency, and explicitly or implicitly claim that more choices and more

capabilities are obviously improving the quality of human and social life (see, *e.g*. Hall 2019). Finally, even Francis Fukuyama, one of the fiercest opponents of transhumanism, acknowledges the use of cognitive enhancement, claiming that "...a society with higher average intelligence may be wealthier, insofar as productivity correlates with intelligence" (Fukuyama 2002, 97).

It is sometimes difficult to talk about technologies dedicated to human enhancement as a specific type of newly emerging technology. Potentially, any newly emerging technology that allows individuals to make themselves better regarding human beings' species-typical functioning could be labelled human enhancement technology. Advances made in the broad spectrum of newly emerging technologies are creating the possibility of improving or enhancing many human capabilities beyond what up until recently were considered 'average' or 'normal'.

Extremely promising fields of enquiry when it comes to human enhancement technologies are genetic enhancement, morphological enhancement, cyborg enhancement, and pharmacological enhancement. Namely, technological progress typically refers to one of six categories of enhancements: cognitive, affective, moral, physical, cosmetic, longevity. (Longevity enhancements are interventions that prolong a human's lifespan) (Newton 2019; Sorgner 2018; Savulescu *et al*. 2011).

In recent times, especially genetic human enhancement has led to many bioethical questions being raised. The big reason for this is that such enhancement has taken a huge step further. According to Stefan Sorgner, CRISPR/Cas9 technology could make it easier for people to develop in a carbon-based way, *i.e.*, they could become organic trans- or posthumans where the posthuman either still belongs to the human species, but has at least one characteristic which goes beyond the limits of current human capacities, or can become a representative of a new species (Sorgner 2021).

Although (bioethical) debates on the social role of human enhancement technologies are constantly underway, these technologies remain relatively poorly defined by experts. The main reason is that no strict demarcation exists between human enhancement and medical therapy that could guide research inquiries made in these fields. As concerns genetic enhancement, the line between medical treatment and

enhancement is extremely complicated. Another problem is that it is changing rapidly over time. Past development of human genome editing technology and its use in medical therapy shows that it holds the capacity to alter the scope very quickly of what is considered normal. The use of medical treatment for longevity enhancements is such a case. Since, on one hand, the processes of ageing have proven to be a "disease of the cells" (De Grey and Rae 2008, 12) and, on the other, the progress of biomedicine is acquiring the capability to manipulate what is happening inside cells, and thus tackling disease by resetting the cellular clock raises the important question of whether it makes sense to strictly distinguish therapy from enhancement. The processes of biological ageing must namely be addressed at the level of the cell, and the molecules that constitute it, in genes, in both the nucleus and the mitochondria, as well as through epigenetic modulation. For example, Alzheimer's disease research may have 'side effects' for efforts to enhance cognition.

Elizabeth Parish describes an even more trivial case to show how difficult it is to draw a strict dividing line between human enhancement and medical therapy: how to define the function of antibiotics? Does the fact they are effective just long enough to fight off a pathogen mean that they belong to medical therapy, or do they belong to enhancement if such treatment amounts to building a super body with abilities that the natural body does not possess? Elizabeth Parish concludes that an unambiguous answer here is impossible (Parrish 2019).

It is expected that the further progress of various forms of gene therapies oriented to techniques of inserting or deleting genes will make the distinction between gene therapy and gene enhancement even more difficult. This is especially true in a situation where gene therapy will be able to address a complex genetic disease, *i.e.*, diseases, caused by many genes, not just one.

The very simple cases presented above reveal the difficulties in defining a strict dividing line between medical treatment and human enhancement. These are extraordinarily complex, knowledge-intensive and controversial topics that pose a challenge for all bioethicists and transhumanist thinkers. Genome enhancement appears not only at the crossroads of different theoretical views, but also at the intersection of very different social interests. For instance, if these technologies permit

more genes that cause undesirable conditions or characteristics in humans to be identified, societal pressure could then increase to eliminate those conditions or characteristics. This approach could shrink the category of 'normal' species functioning and broaden what is considered to be a deleterious mutation. Some former normal human variants could thereby be recategorized as pathological.

This means that in the context of current debates on the further progress of human enhancement technologies it is vital to follow pragmatism and pluralism. On the one hand, we can avoid, in this manner, a total ban of many gene therapies that promise to heal various diseases and, on the other hand, we can also avoid a laissez-faire approach. We need more realism and fewer ideological statements while assessing genomic enhancement. When genomic therapy in biomedicine also opens the way to enhancement interventions, a bioethical assessment based on empirical data and not simply moral and ideological disqualifications will be needed.

The Need to Encourage an Open and Critical Intellectual Discussion on All Types of Human Genome Editing Technologies

In the last 10 years, the progress made by human genome editing technology has meant human organisms can be transformed by direct manipulation of genomes within a broad range of applications. The invention of CRISPR/Cas9 technology about 10 years ago paved the way for a revolution in gene therapy. Put simply, CRISPR-Cas9 works as a type of molecular scissors that can selectively trim away unwanted parts of the (human) genome and replace them with new stretches of deoxyribonucleic acid (DNA) (Tyagia *et al*. 2020). In recent times, this revolutionary genome editing technology has become extremely popular in biomedicine.

Before the invention of CRISPR-Cas 9 technology, biomedical gene therapy had a longer history based on genetic recombination. This type of gene therapy can correct deleterious genetic mutations that cause various diseases by inserting multiple copies of the nonmutated version of the gene either directly into the body (in vivo) or into stem cells that

are subsequently infused into the body (ex vivo). With genetic recombination, the nonmutated copies of the gene are usually carried into human cells by a 'vector' (modified viruses or plasmids as delivery vehicles) that is able to facilitate the transport of the genetic material into the patient's cells (Keeler *et al*. 2017). A big problem with this therapy is that the behaviour of added genes inserted into the human body is not always predictable. These genes cause many 'off-target' effects, what may bring several negative consequences for patient health. At some point, this growth even ended abruptly following the death of a young boy called Jesse Gelsinger, a patient who was receiving an experimental therapy in 1999 (Mukherjee 2016).

The discovery of CRISPR-Cas9 technology made it immediately apparent that this technology is far more accurate and precise than the older genetic intervention technologies, that its editing power can be assessed in ways not possible with genetic recombination, and that it can be applied at any stage of human development. CRISPR-Cas9 technology has opened up new dimensions with respect to the accessibility and the scope of therapeutic applications.

Still, the great expectations regarding the potential of human genome editing technology have been accompanied by growing awareness about its risks. The main concerns refer to human germline editing. Since such editing creates changes that can be passed on to future generations and the implications of these heritable changes have not been tested in practice, this approach is presently not safe enough to justify its use in medical therapy.

Before CRISPR/Cas9 technology emerged, bioethical discussions connected with human genome editing were mostly concentrated on ethical arguments known as "principilism" (Bauman 2016). Here, the ethical dilemmas have generally been restricted to the questions of autonomy (one cannot experiment on another individual unless that person has agreed to be experimented upon), beneficence, and non-maleficence (the possibility of harming a person must be lower than the possibility of helping the person), etc. The reason for this is that the somatic genome editing therapies developed over decades were based on longstanding experimental and clinical research practices and healthcare. The difference between somatic and germline genome

editing was first introduced in A Report on the Social and Ethical Issues of Genetic Engineering with Human Beings prepared by U.S. President's Commission for the Study of Ethical Problems in Medicine (Splicing Life, 1982). In this report, the difference was described in the following way: mutations that occur in somatic cells only affect the progeny of that mutant cell, which means that the effects of such mutations are restricted to the individual in whom they occur. In germ cells, however, mutations result in the altered DNA being transmitted to all cells—somatic and germinal—of an offspring.[2]

For example, French W. Anderson, the leading scientific advocate of gene therapy in 1980, differentiated three types of genetic engineering of future generations: therapy on somatic cells for treating genetic disease, the enhancement of simple characteristics of human beings (*e.g.*, height), and eugenics to improve complex human traits. Anderson saw somatic gene therapy as ethically acceptable whereas enhancement and eugenics were not (Anderson 1985).

The situation has changed radical in the last few years when the first practical misuses of CRISPR technology in germline editing were reported (Martin *et al*. 2020; Turkmendang and Martin 2021). Namely, the hardly controversial and speculative bioethical debates on germline human genome editing at the time of the rise of genetic recombination completely changed following the announcement of the birth of the first gene-edited babies in China. The scientific and wider social community was shocked when biophysicist He Jiankui from China announced the birth of twin babies whose genes he had modified with the gene editing technique. He had used CRISPR in an attempt to make a specific change in the DNA (CCR5 gene) of the human embryos. By so doing, He Jiankui had hoped that HIV would be unable to infect the white blood cells of the babies born from the embryos. The edited version of the CCR5 gene

[2] As stated by many authors, the distinctions made between somatic and germline editing might still be too approximate. It is worth noting that the effects of heritable changes only occur when the modified cells are used to establish a pregnancy that is carried to term. Thus, a distinction has been made between germline genome editing (GGE), which may only affect in vitro embryos in research activity, and heritable genome editing (HGE), which is used in reproductive medicine (Almeida and Ranische 2022) .

might be passed down to those babies' children (if any) and down through the generations (for more, see: He Jiankui 2018; Sand *et al*. 2019).

Although an independent investigation of the He Jiankui affair is still pending due to a lack of information (Meyer 2022; Greely 2019), it has provoked an extremely strong response from most leading genetic scientists, bioethicists, and international institutions around the world. While aspects of the opinions of the actors involved vary, they have generally joined in the call to ban clinical uses of human germline editing.

The birth of the first gene-edited babies is certainly connected with the threat of a new type of eugenics. Since germline editions can be inherited by future generations, their effect on future human life is unpredictable. Still, even though about 30 nations currently have legislation that directly or indirectly bans all clinical uses of germline editing, it is hard to expect that all of the actors involved will be able to fully agree on the bioethical positions on human germline editing.

It is interesting that some critics of human genome editing contend that the real change in the germline modification discourse already came with the new technique of mitochondrial replacement therapy that is used to avoid inherited mitochondrial disease. The UK was the first jurisdiction in the world to legally permit the technique of mitochondrial replacement therapy before the controversial "He Jiankui affair".[3]

Moreover, there are different views on how to cope with the issue of the extreme conditions of human life. People are suited for life on earth, but not for the extreme conditions in outer space. Therefore, each future imaginary provides a wide range of arguments for what and how to do with such extreme situations. For example, even serious scientists,

[3] The report of Nuffield Council on Bioethics maintains that there are no categorical reasons to prohibit some kinds of research in germline genome editing (Nuffield Councl of Bioethics 2018). By emphasizing the value of procreative freedom, the report stresses that in some cases human genome editing might be the only option for couples to conceive genetically-related, healthy offspring. In the UK, the Human Fertilisation and Embryology (Mitochondrial Donation) Regulations permit genetic modification to enable the conception of children without the mitochondrial disease. It assesses that, compared to nuclear DNA, making changes to mitochondrial DNA is relatively insignificant in the makeup of human genes.

including some working with NASA, have started to ask whether humans can be genetically enhanced for space travel or the possible colonization of other planets. They will need a powerful immune system to protect them from possible deadly new pathogens, etc. Georg Church, the world-renowned genetic scientist who cofounded Harvard Medical School's Consortium for Space Genetics, is promoting the further exploration of human health in space as well. He has identified around 40 genes that might be advantageous for long-term spaceflight (and would also benefit those who stay behind) (Pontin 2018).

It is legitimate to ask whether people should still be human beings or new creatures. Taking account of the complexity of the current situation in germline genome editing, it would be intellectually and socially unproductive to entirely halt any more of such research activities, let alone further intellectual discussions on this topic. A difference does exist between research and clinical application. Many historical cases in the field of biosciences suggest that it is pointless to impose categorical bans on any kind of newly emerging technology. In the current atmosphere of a strict ban of germline genome editing, calls for the measured analysis of all the (positive and negative) implications of all kinds of human genome editing approaches are not always heard. Finally, the next generation of personalized medicine is accompanied by many expectations that it will be based just on gene therapy, whether targeting the germline or somatic cells. The issue of whether to permit (or not) the further progress of any kind of emerging technology must be resolved following a proper assessment of the risks and benefits. Namely, when the benefits are found to outweigh the risks scientists (or physicians) must determine whether to continue, modify, or immediately stop the continuing progress of technology. All new technologies should not be analyzed simply in terms of their risks, but also in terms of the opportunities they provide. New governance experiments and settings are hence needed in order to foster a rational discourse on the further development of all kinds of human genome editing technologies. Here, it will also be necessary to re-define the function of expert knowledge and its role in risk assessment in the intermediary space between cognitive and practical values. If we wish to prevent the complex relations between values and technologies, among different values, and

the competing interpretations of values from leading to ever more conflicts about new and emerging technologies, we must find new ways to foster a rational discourse on these technologies and the relevant values.

Conclusion

The main aim of the discussion was to address the issue of human genome editing as it relates to the recent progress made by human enhancement technologies. In modern transhumanist discourse, much attention is paid to the issue of human enhancement as supported by the progress of newly emerging technologies. Today, a number of potentially valuable applications are emerging from the processes of unifying newly emerging technologies, *i.e.*, biotechnology, nanotechnology, information technology, and cognitive technology. The convergence of various disciplinary fields is especially important in human genomic technology related to genomic therapy and genetic enhancement. In the article, some cases of the converging of technology were presented. Human enhancement as supported by transhumanist thinkers was shown to still be relatively poorly defined. The prime reason for this is that no strict demarcation between human enhancement and medical therapy exists that could guide research inquiries in these fields. This issue becomes extremely complicated when it comes to genetic enhancement. The discussion included a presentation of certain bioethical implications of the current rapid progress being made by human genome editing technologies. The invention of CRISPR/Cas9 technology some 10 years ago opened the door for a revolution in gene therapy. While before the emergence of CRISPR/Cas9 technology bioethical discussions were generally concentrated on classical questions, the invention of CRISPR/Cas 9 technology and the practical (mis)use of this technology in the context of germline genome editing are creating totally new bioethical dilemmas. When taking account of the complexity of the present situation with germline genome editing, it was argued that it would be intellectually and socially unproductive to entirely halt further research into various types of human genome editing. It is vital to keep

in mind the maxim of moderate and pragmatic transhumanists who claim that in modern democratic societies, and despite the contested future of scientific-technological progress, pluralistic views and public discussions of the benefits and risks of any newly emerging technology (including germline genome editing technology) must not be stifled.

REFERENCES

Almeida, Mara and Ranische, Robert (2022). "Beyond safety: mapping the ethical debate on heritable genome editing interventions." *Humanities and Social Sciences Communications*, 9 (139): 1-14.

Anderson, French (1985). "Human Gene Therapy: Scientific and Ethical Considerations." *The Journal of Medicine and Philosophy,* 10 (3): 275–291.

Baumann, Martina (2016). "CRISPR/Cas9 genome editing – new and old ethical issues arising from a revolutionary technology." *Nanoethics,* 10(2): 139-159.

Bostrom, Nick (2005). "Transhumanist Values". *Journal of Philosophical Research* 30 (Supplement): 3-14.

Calvert, Jane (2010). "Synthetic biology: constructing nature?". In *Nature after the genome*, edit by Sarah Parry and John Dupre, 95–113. Oxford: Wiley Blackwell.

De Grey, Aubrey and Rea, Michael (2008). *Ending Aging. The Rejuvenation Breakthroughs That Could Reverse Human Aging in Our Lifetime.* New York: St's Martin Griffin.

Drew, Endy (2005). "Foundations for Engineering Biology". *Nature,* 438 (7067): 449–453.

Drew, Endy (2014). "Introduction. How would you design nature?". In *Synthetic aesthetics. Investigating synthetic biology's designs on nature, edit by* Endy D., Ginsberg A., Calvert J., Schyfter P., Elfick A., 7–22. Cambridge: The MIT Press.

Fuller, Steve (2012). *Preparing for Life in Humanity 2.0.* Hampshire and New York: Palgrave Macmillan.

Fuller, Steve and Lipinska, Veronika (2014). *The Proactionary Imperative. A Foundation for Transhumanism.* Hampshire and New York: Palgrave Macmillan.

Greely, Henry (2019). "CRISPR'd babies: human germline genome editing in the 'He Jiankui affair'." *Journal of Law and the Biosciences,* 6 (1): 111–183.

Hall, Melinda (2019). "Second Thoughts on Enhancement and Disability." In *The Oxford Handbook of Philosophy and Disability*, edit by Adam Cureton and David T. Wasserman, 633–650. Oxford: Oxford University Press.

Hollingworth, Rogers (2006). *Fostering Scientific Excellence. Organizations, Institutions, and Major Discoveries in Biomedical Science.* New York and Cambridge: Cambridge University Press.

Huxley, Julien (1968). "Transhumanism". *Journal of Humanistic Psychology,* 8 (1): 73-76.

Jiankui, He (2018). *About Lulu and Nana: Twin girls born healthy after gene surgery as single-cell embryos.* Available at <https://www.youtube.com/watch?v=th0vnOmFltc&t=81s>

Keeler, Allison and Mallah Mk and Flotte Tr (2017). "Gene Therapy: Progress and Future Directions." *Clinical and Translational Science*, 10(4): 242–248.

Kurzweil, Ray (2005). *Singularity is Near. When Humans Transcend Biology.* New York: Penguin Group.

Martin Paul, Michael Morrison, Ilke Turkmendag, Brigitte Nerlich, Aisling McMahon, Stevienna de Saille and Andrew Bartlett (2020). "Genome editing: the dynamics of continuity, convergence, and change in the engineering of life." *New Genetics and Society*, 39 (2): 219-242.

Meyer, Morgan (2022). "Taking responsibility, making irresponsibility. Controversies in human gene editing". *Social Studies of Science*, 52 (1): 127-143.

Mukherjee, Siddhartha (2016). *The Gene. The Intimate History*. New York: Scribner.

Nuffieled Council on Bioethics (2018). Annual Report 2018. Available at <https://www.nuffieldbioethics.org/assets/pdfs/NCOB-Annual-Report-2018.pdf>

Newton, Lee (Ed.) (2019). "The Brave New World of Transhumanism". In *Transhumanism Handbook*, edit by Newton Lee, 3-49. Cham: Springer Nature Switzerland AG.

Parrish, Elizabeth (2019). "Extending Healthy Human Lifespan Using Gene Therapy." In *Transhumanism Handbook*, edit by Newton Lee, 423 – 433. Cham: Springer Nature Switzerland AG.

Pontin, Jason (2018). "The Genetics (and Ethics) of Making Humans Fit for Mars". *Wired,* July 8.

Roco, Mihail and Bainbridge, William (2002). *Converging technologies for improving human performance: nanotechnology, biotechnology, information technology and cognitive science.* Available at <https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/bioecon-%28%23%20023SUPP%29%20NSF-NBIC.pdf>

Sand, Martin, Annelien, Bredenoord and Karin Jongsma (2019). "After the fact—the case of CRISPR babies". *European Journal of Human Genetics,* 27 (June):1621–1624.

Savulescu, Julian, Ruud ter Meulen, Guy Kahane (2011). *Enhancing Human Capacities*. Malden and Oxford: Wiley Blackwell.

Sorgner, Stefan L. (2022). *Philosophy of Posthuman Art*. Basel: Schwabe Verlag.

Sorgner, Stefan L. (2021). *We have always been cyborgs.* Bristol: Bristol University Press.

Sorgner, Stefan L. (2016). *Transhumanism. Die gefaerlichste Idee der Welt!?.* Freiburg in Breisgau: Verlag Herder GmbH.

Splicing Life (1982). *A Report on the Social and Ethical Issues of Genetic Engineering with Human Beings.* Washington, President's Commission for the Study of Ethical Problems in Medicine. https://repository.library.georgetown.edu/handle/10822/548744

Stambler, Ilia (2019). "What Do We Need to Know to Treat Degenerative Aging as a Medical Condition to Extend Healthy Lifespan?" In *Transhumanism Handbook*, edit by Newton Lee, 381-397. Cham: Springer Nature Switzerland AG.

Tyagia Swati. Kumarb R., Dasc A., So Youn W., and Pratyoosh S. (2020). "CRISP R-Cas9 system: A genome-editing tool with endless possibilities." *Journal of Biotechnology,* 319 (10. August): 36-53.

Turkmendag, Ilke an Martin, Paul (2021). "Thinking the unthinkable: how did human germline genome editing become ethically acceptable?" *New Genetics and Society,* 40 (4): 384–405.

Venner, Eric, Yi Victoria, David Murdock, Ge Zhang, Lan Zhang and Hana Zouk (2021). "Neptune: an environment for the delivery of genomic medicine". *Genetics in Medicine*, 23(10): 1838-1846.

CYBORG NETWORKS: FREEDOM THROUGH DATA WITH OPEN SOURCE AND IMPLANTABLE DEVICES

KEVIN RÄNDI[1]

Abstract

Implantable computing technologies are the blueprint for the ongoing process of cyborgization. Already with the present devices, design decisions and computational issues blend together with a living organism, so that our health, for example, depends on well-adjusted information security, which requires constant work and decisions by many actors of different expertise. This means that implantable technologies form important sociotechnical systems. I will explore cyborg vulnerabilities found in sociotechnical systems, but take seriously the proposed development model of open-source technologies, promising existing and future cyborgs more autonomy and freedom. However, I want to take this vision further. Given the number of technologies that matter in the context of cyborgs, it becomes necessary to look at a wider application of open source, needing more experts and quicker data exchange. While offering a technological solution to make the data exchange possible, granting the users freedoms and autonomy, I also find it important to address how the debates around open source move from software to a large-scale social consideration, should the solution be implemented.

Keywords*: cyborg, open source, sociotechnical systems, implantable technologies, data exchange*

Introduction

In the first half of 2022, following an article published in *IEEE Spectrum*, news spread about people risking losing their sight for a

[1] PhD Student in Studies of Cultures, School of Humanities, Tallinn University.
Email: <kevin.randi@tlu.ee>

second time due to a company going bankrupt (Strickland & Harris 2022). Evidently, the company behind valuable implantable retinal and brain technologies had faced financial difficulties, discontinuing any further support for the parts, and thus causing uncertainties for its users. This is only one example of what could be called a cyborg problem with implantable technologies. However, it is an important issue that should make us think about such technologies beyond the interaction between the user and the machine. The relevant questions concerning our health, well-being, and vulnerabilities are also questions about development and design processes of technologies. In such cases, I take the view that dependency on technologies is at the center of a complex interplay between several actors and forces, including clinical experts and engineers, together with those who maintain, manufacture, design, use, and live in this complexity. All these human and technological actors, but also regulations and rules form what is known as sociotechnical systems (Cooper & Foster 1971; Franssen & Kroes 2009).

A view from sociotechnical systems, where the focus is on existing implantable technologies and its stakeholders, uncovers vulnerabilities with implantable technologies in more detail, but it also discloses why we need to give more philosophical attention to the development of computing devices. Due to the surgical nature of such technologies, the wearers may wish, for example, that they had more certainty about its maintenance and access, if needed. Furthermore, as in the example above, the way technologies are manufactured, raises concerns about autonomy for the wearers. One possibility is to design technologies with the principles of open source, important in the field of software and computing. I commit to this view. However, by assuming that implanting wireless and networked computational devices with both software and hardware will continue to be relevant on the path to our cyborg future, I want to further examine the use and the ethos of open source. If we remain true to both sociotechnical systems and open source, having an open-source piece of hardware and software in our bodies might not be enough. We would require different communities of experts together with quick and reliable ways of exchanging a large amount of data. While I propose an existing technological solution, able to co-join private and public institutions as an example, I will show how some new

philosophical issues and considerations might emerge. Implantable technologies already make information security an integral part of our understanding of health. With the changed understanding of health, both novel vulnerabilities and possible mitigations need only look to the existing debate between closed and open source having an impact on social and political matters once the decision is made.

Cyborgization Now

Cyborg, a portmanteau for a cybernetic organism, refers to a deeper technological embodiment forming a hybrid that is a "self-regulating human-machine system" (Featherstone & Burrows 1995, 2). The term "cyborg" was first coined by Clynes and Kline (1960) to propose a new way for humans existing in space by incorporating "exogenous components" that extend "the self-regulatory control function of the organism in order to adapt it to new environments." This entity has become an eminent figure in today's academic literature that captures ways we should rethink our being in the existing technological world, as famously proposed by Donna Haraway (1991). For others, like Ray Kurzweil (2005), cyborg also denotes one possible posthuman outcome of the ongoing technogenesis. Reconceptualization of the human with cyborg ontology is often attributed to posthumanism, whereas the view of cyborg as the outcome is more closely aligned to transhumanist ideas. Due to different ontological and anthropological assumptions, both are seen in many respects as irreconcilable. Without seeking any unification of the two, there is, nevertheless, a question concerning technology that requires attention from both. Namely, what could be significant for the future of cyborgization once we start from the lifeworlds of those currently needing cyborg technologies?

Cyborg technologies are often discussed in the light of human enhancement, with the emphasis on RFID chips, and brain-computer interfaces that move beyond the species-typical biological functioning (See, *e.g.*, Barfield & Williams 2017). However, it cannot be assumed that current wearers of prostheses and medical implants are able to see their lives merely having the species-typical functioning restored. Much

about their lives becomes different, as do certain meaningful concepts about their bodies. As Nelly Oudshoorn has argued, instead of equating cyborg technologies with human enhancement or emergent technologies, we need to also consider the lives of the already existing "and more familiar" cyborgs for whom technology is a matter of life and death (Oudshoorn 2015). Medical implantable devices, to begin with, are the clearest instance of how the future of humanity could be dependent upon the design and how all the relevant actors come to affect us, revealing already novel forms of vulnerabilities from which to re-think the cyborg future.

Also, by mentioning vulnerabilities next to existing medical devices and "familiar" cyborgs, I do not mean that the latter need to be taken as more vulnerable because of a person's condition, or because of some specific medical function of these devices. Rather, these devices help to reveal already existing vulnerabilities that could characterize other implantable technologies of the future. The idea that vulnerabilities form our unwanted companion for the whole road to the future has been explicitly stated by Mark Coeckelbergh (2011, 2013). Even the currently imagined posthumans would be relational, embodied within something and, as such, dependent on other factors; only the form of vulnerability is apt to change (Coeckelbergh 2011).

Implantable Technologies and Sociotechnical Systems

Wireless and networked implantable medical devices (IMDs) have been used for years. Without referring to any specific device now, these devices within the overall e-healthcare systems, including also the Internet of Medical Things (IoMT), have provided medical professionals better data monitoring, and remote access to health. Moreover, some devices, like implantable cardioverter defibrillators (ICDs), save lives with an automated monitoring and regulating system to shock and pace the heart to avoid the worst. As it is with many technologies, while there are benefits, there are also trade-offs. With all the connectivity and software in a wireless IMD, biomedical considerations have also become cybersecurity concerns. Signals, databases, and the device itself can be

exposed to malicious actors (Pycroft & Aziz 2018). Because of the cybersecurity issues, people with IMDs might need the extra layer of protection from other (wearable) technologies, such as wristbands, UV-visible tattoos, centralized databases (See, Denning *et al*. 2014). Furthermore, to satisfy utility, safety, security, and privacy goals, IMDs need to meet a broad body of criteria, from software updates to data protection (Halperin *et al*. 2008). Not only is (medical) cybersecurity a relational matter of persons, devices, and professionals, but a successful mitigation of such issues requires even more actors, such as cybersecurity experts, and their constant work on innovation. Furthermore, technology itself is not the only thing that constitutes all the factors and vulnerabilities. An unexpected shock, and at the wrong time, from an ICD, as I will briefly show, sets many relevant sociotechnical practices in motion. The wearers of ICDs are the stakeholders who are impacted by the design, but also dependent on medical professionals who interact with these technologies by referring to the existing regulations. What this all means is that networked devices form a complex interaction and interdependencies between technologies and many actors, highlighting what we should mean by sociotechnical systems. To solve emerging problems of implantable technologies, or even cast light on overall emerging cyborgization, we should investigate these interactions and interdependencies as such.

I want to emphasize that in the context of existing IMDs, it is the wearer who is ultimately affected by the overall "setup". Oudshoorn (2016) has analyzed and highlighted how people experience and attempt to handle an inappropriate shock—the one not needed, and often unexpected—from their defibrillators. According to her study, both the technicians and technomedical culture determine whether an inappropriate shock could be likely. A technician, under uncertainty, might tune the machine to have more agency, and risk an inappropriate shock (noted in the USA and previously in the Netherlands), or disable a couple of functions on the device, and allow the heart to attempt to regain itself (as it is in the Netherlands now) (Oudshoorn 2016). She then further points out how frustration with the technology and technicians has led some people to use magnets to tame the device, having learned it themselves or suggested to do so by medical professionals or other

users. In case of an emergency, physicians may use round-shaped clinical magnets to bypass the programming and wait for an appropriate professional to analyze it. However, there are risks involved with using magnets (Rodriguez-Blanco *et al*. 2013), and even more so when used outside medical institutions. These technologies reveal just how interdependencies generate the need for more artifacts as well as the need to modify the existing technology because the medical device itself depends on certain actors. But it is also the opposite. The technology itself, programmed and accessed in a certain manner and by certain professionals, generates the layout for how actors are involved.

Besides devices and medical practitioners, we also must account for the software and, more importantly, the code that the device uses. Karen Sandler's story and research shows why it is such an important matter for IMDs. Sandler, formerly in the GNOME Foundation, and currently an executive director of Software Freedom Conservancy, highlights an inappropriate shock experience with her ICD. In her "cyborg-to-cyborg" interview with Marie Moe (2021), she mentions her hypertrophic cardiomyopathy being the reason for having an ICD, and tells of an instance during pregnancy. When the ICD registered the heart palpating as worthy of a shock, it made her take countermeasures to calm it down. As ICDs are computing devices, Sandler sees the fault in software, which is usually owned by manufactures, and as a result, closed for patients and medical professionals to access and test its features (Sandler *et al*. 2010). Closed source means that software makers have decided not to provide the code which can be read and modified by others (third parties). It is very often the case with the so-called propriety licensing where the authors function as the owners, setting the restrictions.

Instead, Sandler argues for open-source software that allows others to review and modify the code. It is mainly because some essential technical problems might be overlooked, and the closed source is potent to vulnerabilities as the only ones able to investigate and mitigate are the owners (Sandler *et al*. 2010). There is another benefit, according to Sandler, besides having more eyes and hands on technical problems. Open source would not ultimately mean that wearers themselves access it, but that they have the freedom to choose the professionals and

companies that have access to the device (Sandler & Moe 2021, 105). Moreover, as she mentions, because of proprietary software in the ICD, there is no way to ensure that the device can adapt to the changes in her life—this technology is meant to last for years, however, in a world with rapid technological changes (Sandler & Moe 2021, 103). What Sandler proposes here, is a distinct concept of freedom, applicable to implantable technologies. It is the freedom of relationality, important for our sociotechnical systems' context. This gives one the choice to decide the assemblages between actors. It implies that an individual should have the freedom to decide among the professionals by allowing them also to access and, if needed, modify the device. To generalize, this could become relevant for all kinds of implantable technologies in the process of cyborgization. If such freedom presupposes the development model of open source, it deserves further investigation.

From Open-Source Software to Clinical Engineering

The previous section provided a look into some problems we need to face with implantable technologies. However, it also uncovered open source as a development model that could be needed for existing and emerging cyborg technologies. Furthermore, without open source, cyborgs could face further difficulties to their quality of life if, by having an implantable technology, one is tied to a closed relation in terms of actors, technicians, or a company's decisions. Nevertheless, an account for software might not be enough to argue for the need and value of open source that needs to include many elements of cyborg technologies. Furthermore, particularly when it comes to other aspects of biological life where open-source development could be used or is needed, there are already worries and benefits pointed out. This section and the following one will examine the development model across relevant human and technology realms.

Drafted by Bruce Perens, and later modified by his colleagues, the Debian Free Software Guidelines (Debian Project 2004) became "The Open Source Definition" of the Open Source Initiative (OSI). The document states the terms according to which software counts as open

source (Open Source Initiative 2007). Anyone should be allowed to access the source code for software, and be able to study, use, modify, and distribute in under numerous licenses available. Eric S. Raymond, in his famous essay *The Cathedral and the Bazaar* (2001) examined the development model of Linux's operating system kernel that happened via the internet. According to Raymond, before and outside the development of the Linux operating system, software with open-source code reached only a small group of developers. It resembled people building cathedrals. However, the development model of Linux took advantage of the internet. The latter indicated that everyone could see and access the development early on. Thus, Linux took the open-source development to the public, resembling something of a bazaar.

The term is now wide-spread, and open source has become an ethos that captures more than software development. Without mapping out all of them, many producers are using the main idea and its licensing strategies to produce the goods. In the context of technological hardware, the Open Source Hardware Association (OSHWA) provides its definition and statement of principles in which hardware design is open source if the blueprints are available to the public. However, it also adds the specifics for hardware, including a plea to use readily available components, materials, and open infrastructure for the exchange of blueprints (Open Source Hardware Association 2022). While distributed version control—the visible bazaar of open source on the web— such as GitHub is known for the open-source software distribution, a website called Thingverse is an example of a hardware bazaar that provides the design blueprints.

The field of clinical engineering brings together both hardware and software with the emphasis on open source as valuable for our current needs (See, *e.g.*, De Maria *et al*. 2018, 2020). The situation, however, is complicated due to the function and nature of bionic and implantable medical devices. De Maria *et al*. (2020) nevertheless argue for the open-source medical devices (OSMD) because the current healthcare requires more accessible and affordable technologies for everyone—valued by Sustainable Development Goals and the World Health Organization. The challenge, however, is how to make open-source clinical engineering safe. The developers of OSMDs need to include and consider innovation

along the lines of all that makes up these technologies: software, hardware, safety regulations of a medical device. Thus, De Maria *et al*. propose the following definition. The first half consists of the requirements for open-source software as well as hardware. Such technologies should be built with interchangeable and interoperable parts and depend on more open e-infrastructures for information (De Maria *et al*. 2020, 9). It then proceeds by asking engineers to consider international safety standards. For the development of medical technologies, many legislation acts exist, and these are not hindering the commitments to open-source in the given area. The problem is rather about harmonizing different legislations across the globe so that those who are in need would benefit from the accessibility and affordability of OSMDs designs and products (De Maria *et al*. 2018).

Data Exchange Layer: A Bazaar for Cyborgs

Recently, some transhumanist scholars have started to see existing technological solutions and models as a viable option for certain ideas that we have about the future. Take, for example, Melanie Swan's idea on how blockchain technologies with smart contract applications could be able to support the idea of a cloudmind, defined as an individual's access to parts of his or her cognitive resources for the benefit of collaborative tasks (Swan 2019). In one commentary, Swan also explores the role of a distributed version control software and hosting site GitHub—essential for software developers to keep track of changes in code—to extrapolate how people who use cloudminds could be able to access earlier versions of their cognitive data, like memories in case of its loss (Swan 2015). GitHub, or Git-like version control software, in general, explicitly illustrates Raymond's idea of the bazaar for open-source development and its communities. Analogical to this well-controlled distribution database, others have been proposed, seen as beneficial for a safe progression towards a posthuman future. Stefan Sorgner promotes the second one by examining the usefulness of gene analysis, and the role of regulated distribution control on a governmental level. According to him, gene analysis—a promising path towards a posthuman

future—benefits from big data collection but is being troubled by the problem of privacy as many actors and companies have their interest in collecting and keeping the data to themselves (Sorgner 2017). In his later work (2021), however, Sorgner suggests a model that could be able to distribute gene data efficiently without letting interested parties centralize our data, and as a result, gene data can be used for greater innovation. Furthermore, in his vision we should think about the inescapable situation of total surveillance and reconsider the government as a democratic hosting service for distributing our data in a manner that we still have the freedom and control in this, *e.g.*, one can use data or other currencies to pay for medical necessities (Sorgner 2021, 42–46). In this scenario, since the government is the distributing host, private (pharmaceutical) companies lose the ability to overcharge for the goods as it then depends on whether people provide the data to the government to further distribute it to the company (Sorgner 2021, 44). While our investigation of human-technology interlinkages and open source differs from the ones above, the idea of regulating distribution models and technologies remains relevant. Two existing philosophical and technical explorations make use of distribution service support to mitigate the issues with data chaos and unorganized management. Considering our focus on the open-source devices and development, a technological infrastructure, providing similar benefits, should be asked for. In other words, what kind of "bazaar" architecture do we require so that we can have sociotechnical systems where open source is beneficial for:

- innovative and open clinical engineering so that certain parties do not make the device (meant to last for a long time) obsolete;
- people who use IMDs can take advantage of the freedom of relationality, *i.e.*, they have the option to choose among technicians and other professionals in hopes of avoiding uncertainties and unpleasant experiences;
- all the data required for this process to work across sociotechnical systems, nevertheless, can be safely managed, and used to improve the lives of those who either need IMDs or decide to wear any implantable device for any choice of cyborgization?

As these encompass both current and future difficulties, we must continue with the search for a "bazaar"—a distribution system able to accommodate the valuable world of open source: a decentralized system, rising from the need of human-technology interlinkage in a sociotechnical system with many actors involved. In this section, I will further provide a description of a technological model that could suit these needs, while in the next one, I will further analyze some related concerns.

Some countries are being recognized for taking the advantage of information and communication technologies (ICT) applied to government and other infrastructures. Estonia is considered a prime example having an e-governmental sociotechnical and information systems' solutions for running things. So, for now, I will consider a concrete example of the technology, called distributed data exchange layer (DXL), that runs behind the infrastructure of Estonia's e-Government (or e-Estonia, as it is sometimes called). A software-based ecosystem or a middleware, called the X-Road (officially "X-tee" in Estonian), is Estonia's open-source solution that is used to securely move data between different organizations (Estonian Information System Authority 2022)[2]. With this ecosystem, both the public and private sectors can exchange data without fragmenting collaboration and causing needless data duplication (Paide *et al*. 2018). Simply put, this means that all existing service providers of public and private sectors, who are members of X-Road, are linked together for easy data exchange by digitally signed contracts. Not going into all the technical details, I will lay out what it means for the existing residents within this ecosystem.

We could start with a concrete example of healthcare e-service. Residents can access all their data and histories concerning analyses, prescriptions, insurance, etc., on a patient web portal by authenticating themselves with a smartphone application or an ID-card. After some medical procedure, all the results go to the portal where both the patient and, for example, a family physician can access the data as it is. When the patient receives a prescription for medicine; the information goes to the database, and a pharmacist, entering the identification, can see it

[2] Available on GitHub: <https://github.com/nordic-institute>, last time accessed in September, 20, 2022.

together with other relevant information. While this forms interlinkage within healthcare infrastructure, the data exchange also reaches other institutions. To get a driver's license, one needs a medical certificate, and data goes from the patient portal to the road administration. To pay for the needed services, one exchanges data between the internet bank and, for example, the road administration. Suppose also that a car accident happens. After the phone call, an ambulance can quickly position the call, and the professionals can access, using a person's ID, all the real-time information, ranging from medicines currently used to allergies.

The open-source "cyborg" engineering and the users of these devices could benefit from a similar technological ecosystem that includes data exchange layer. This suggestion should be seriously and positively considered. The networked, software-based, yet at the same time health-related nature of implantable technologies, require better and quick-response connection within the complex sociotechnical systems to provide reliable and safe human-technology coupling. One could picture a working data exchange infrastructure that links together many bazaars or hubs of open-source developments. In case of a security issue or other bug, DXL would allow all the relevant parties to be aware of it, starting with patients, physicians, and technicians to engineers. The open-source code would be exchanged similarly in this case. Suppose, however, the update for the implant becomes available, the user and physicians would be able to consult about whether installation would be a good idea. The important factor in this case is also that all the other data in the system would give both the user and professionals a more personalized view over things to decide properly. A data exchange layer together with an ethos of open-source present a vision fit for cyborgs. A quicker and reliable data exchange and open innovation are both needed so that the technological part and the human part would continue to work in harmony. In short, this solution could help mitigate technical concerns we have with implantable technologies and would help us realize how open-source development could successfully work within sociotechnical systems, and further provide people the freedom and control over their health that is not limited by the pre-programmed device.

Open Source for Cyborgs as a Technical Problem

The solution is so far only partially examined. It might allow the experience of greater autonomy, a variety of choices and assurances regarding their device. However, the implementation of open source from a design perspective, and even more so in the context of cyborgization, is far from an easy decision-making. Both closed source and open source as methodologies for developing technologies are, besides the complicated economic decisions, also decided regarding assumptions about security and human factors susceptible for breaking the technology. Furthermore, while the previously mentioned idea could offer a technological solution to the people needing implantable devices for medical purposes to restore biological functioning and sustain their health, open source is also an ethos already applicable to various technological and scientific pursuits and needs, which can be used to reflect on wide-range cyborg developments. Beyond the development of computing devices, cyborgization also demands scientific materials, such as genetic information, body monitoring, and tools, to be opened to the public. As such, it is especially important for the question concerning a variety of cyborg practices and the current concerns with the latter. Here, such openness introduces a public concern about unregulated uses of information and technology. These are demanding challenges. Nevertheless, if technologies can potentially pose certain threats to cyborg or human lives and freedoms, and if open source could be seen as the remedy for these new vulnerabilities, which could also illuminate challenges for the future of cyborgs, it is worth addressing both technological and political issues. In this section, I will consider the technological ones because the implementation of a data exchange layer presupposes that open-source technologies are without fundamental errors or, in our case, life-threatening.

The problem of a company going bankrupt is a serious one for the wearers of implantable technologies. Nevertheless, while it could be a good reason for adopting open source, the proponents of closed source, especially in terms of software, might be right to argue that without the visibility of source we will have better security. Malicious actors are known to take advantage of bugs and overlooked instances of the

program to cause harm. The advocates of open source rely on the idea that a peer-review process of the community can assess and address these issues quickly (Payne 2002). One of Raymond's musings is that "[g]iven enough eyeballs, bugs are shallow" (2001, 19), emphasizing the role of community for greater security. Proponents of closed source might, nevertheless, claim that if bugs cannot be seen and known, there is no security problem (Payne 2002, 67). Furthermore, given that bugs in open-source software are always visible, it is a problem because it is not guaranteed that enough eyeballs mean enough skilled minds (Payne 2002, 70), or that good developers are quick enough to reach these vulnerabilities in an instance of a new version (Schryen & Kadura 2009, 2018). Besides removing bugs, security might be achieved with system integrity, which means that both hardware and software are designed as if meant for each other. Apple's closed source devices could be an excellent example of system integrity, having all the parts necessary for another so that the security cannot be compromised on any level of the system. While these are valid points for our computers, it is questionable if closed source in implantable technologies could be argued for from the same grounds. For example, is a perfect system integrity possible in the case of implantable technologies? It might be in the future. Nevertheless, it would require greater knowledge and data about humans, and it could mean that we still need open source on a level of biological science for a quicker innovation and access. This is the consideration for the next section. Here, we need to continue with the argument on open-source technology.

Where closed source advocacy might make a plausible input for implantable technologies against open source is about the latter being unable to reach bugs and errors quickly enough. But the same could be said about closed source, especially in the context of implantable technologies. On everyday computers, the user can use an alternative program until the problem is solved, upload the data elsewhere and switch the computer. To a living being, an overlooked bug could be fatal (Sandler *et al*. 2010), and there would be no quick alternative (shutting the device down with a magnet is a temporary solution, not beneficial for the whole being). Thus, any cyborg would need not only the largest set of eyes and skilled minds to access the information, but also the

quickest measures. The latter would require both existing data about the person, technology, and other information necessary. Having a data exchange layer technology over our open-source cyborgization could offer this. Moreover, a human wearer of technology would need a collaboration transcending one field of expertise: a bazaar for cyborgs.

My proposal is not without further difficulties. With a sort of data exchange layer technology, we would completely materialize our being within the Internet of Things (IoT). For people with wireless IMDs it already is the case. But with more and more data about biology, different technologies on top of the existing ones (especially with future technologies), and more actors could become involved. Such exploration shifts from the matter of accountability from technologies and cyborgs to sociotechnical systems: the accountability of cyborgs as networked beings. This would now mean a greater dependency on a single technological system to exchange everything. A malicious attempt is then an even greater risk. It is not only even about the attacker shutting down the data exchange layer system, or stealing a large amount of valuable data, but using the large-scale data exchange technology to literally get into a person's body, having now perhaps a greater variety of instances to attack. Open source is argued to be highly efficient for the IoT (Rayes & Salam 2022), and if we again use Estonia as an example, X-Road is also open source. If cyborgization means being more within the IoT than ever, then open source, based on the previous arguments on software, poses a serious concern. While distributed ledger technologies, such as blockchains and many other cryptographical means have proven to be safe enough, the solution in our case might not lie in the answer of adding more technology to make our existing and coming cyborg world a safer place. Rather, open source could perhaps benefit from social endeavors for the ideas proposed here. Besides technological problems, public acceptance tends to have differences regarding wide-spread open source as a valuable ethos.

Securing Biology, Freeing Information

Here we have concern for our security on two levels. One is security vulnerabilities reaching us through the technology, if we are to require a technological and open-source development solution. The second is the vulnerability of our very biological and bodily being at risk due to the ethos of open-source in the first place. Again, there is a concern about the possibility of malpractices when our biological information is freed. We need to address this, because information about biological matters and the development model for sharing the latter could prove to be useful for designers and other professionals working to contribute to the "cyborg bazaar".

As there is an "irresistible analogy between software and molecular biotechnology," (Hope 2008, 18) there is also a possible case for open source in bioscience where we move from software bazaar to a biobazaar, explored by Janet Hope with the hopes of innovation and faster progression. However, all the innovative work in open-source biology could be curtailed by malicious actors, unsafe practices outside the laboratory, and overall societal concerns (Evans and Selgelid 2015). While Evans and Selgelid argue that the ubiquitous spread of information by open-source biology is good for society and for advancements of our well-being, they nevertheless emphasize the role of regulations and biosecurity awareness in open-source biology communities. Their argument, however, is not simply a proposal for further enforcement of existing regulations on an innovative approach, but regulations themselves need to become innovative in the presence of potential malpractices. The threat lies greatly in the fact that, by allowing the radical openness of science and information, all the equipment and knowledge of genomics and synthetic biology gets thrown into the uncertain wilderness of actors with very different intentions. However, the future of open-source biology could benefit, if we do not conceptually or otherwise place the radical openness of science outside its initial context of software development but keep it in the successful model of open-source software (Hessel 2006). As Andrew Hessel says, if open-source software "works, and DNA is software, don't reinvent: adapt" to point out that genetic engineering should be done on

computers like software engineering, using open-source software for the task (Hessel 2006, 290). This re-thinking and re-making open-source biology back into open-source software, a collaborative online "bazaar" is found, for instance, in the project of Cambia.[3] More recently, SynBioHub[4] rose as a repository for engineered biological systems that allows downloading and uploading, for example, DNA and protein designs for sharing and storing within the repository (McLaughlin *et al*. 2018). Open source within biology and clinical engineering, although having ongoing challenges with security and regulations, are seen to benefit from the openness—its ethos, licensing, and technologies (CAD-programs, versioning repositories, free and open-source software, 3D-printers etc.)—that grant it its success. If taken separately, these various commitments could be seen to regulate their own sciences and devices, allowing them further success.

However, and in many ways, the concern with our genetic information and other scientific tools is about its uses for everyone, about the practitioners of the so-called "citizen science" (Irwin 1995). The body itself is something "hackable", and arguably open source, for body hacking "grinder" and biohacker communities in the hacker culture of the cyberworld.[5] Self-experimentation that goes beyond restoration, and thus explores with reconfiguration and enhancement through implantable RFID, magnetic, etc., technologies, is for some significant way to identify themselves as cyborgs. Instead of tackling the ethical issues about the safety, control, and freedom of these practices (See generally, Greguric 2014, 2022), I want to problematize these practices considering our data exchange layer proposal.

---

[3] Available on <https://cambia.org/> last time accessed December 17, 2022.

[4] Available on <https://synbiohub.org> last time accessed December 17, 2022.

[5] Nowadays, the term "hacker" has a negative meaning of a person engaging in malicious activities. However, in the context of open source and hacker culture, hackers have been technology-savvy people that enjoy challenges and solving problems and are opposite to those "crackers" (Raymond 2020). In other words, instead of malicious hackers, we should talk about malicious crackers.

Open-Source Politics for Cyborgs

The question is about inclusion of those using the information but wanting the freedom to seek their own way of cyborgization, and together with it, their own communities for the further development of source code or blueprint. Here is an alluring analogy with the practice known as *forking*, known to the users of the Linux operating system and GitHub. In the software world, forking means separating a copy from an existing community's source base to pursue developing it independently or with another community, often competing with the original (Fogel 2005, 69). It could be pursued if there are conflicts with the community or visions in development. However, the ultimate problem with forking is that it might generate a situation where two versions cannot "exchange code" (Raymond 2001, 72). It is a "technological equivalent" to how DNA and evolution works in case of "speciation in genetics" (Weber 2004, 157). Data exchange layer, our technological vision, could be taken as a solution for this situation in case of "cyborg forking" where people, unsatisfied with current actors, approaches, and data, seek to form their own communities to develop and sustain as cyborgs. Nevertheless, data exchange also fails in this sense: the original community consists of scientists, developers, medical experts, etc., who cannot, due to the lack of understanding of technologies of the new community, or are not willing to help these cyborgs. Not to mention the situation where people decide to copy and develop their own data exchange layer, incompatible with the one offered by the government. An optimal hope lies in a larger set of information, provided by open source ethos across many sciences, to cover for all possibilities of forking. I contest, however, that the "right to fork" should be removed from the data exchange technology for the safety of cyborgs.

Code—the crux of the philosophy of open-source development and information freedom—generates both the cyborg world, and combines both technology and politics, requiring us to take freedoms within the software world seriously, as Samir Chopra and Scott Dexter have said (Chopra & Dexter 2007, 2008). "In the cyborg world," Chopra and Dexter moreover claim, "closed software threatens individual autonomy; the advocacy for, and the provision of, closed software is a

form of paternalism, diminishing cyborg autonomy as it controls and regulates the nature of human-machine interaction" (Chopra & Dexter 2007, 47). The freedom of information, code-based or otherwise, and other rights have been emphasized for the sake of "cyborg citizenship" even earlier, needing political protection and political technologies for that matter (Gray 2000, 28–29). Nevertheless, while open source and other technologies, in principle, can grant these freedoms, it would be wise to account what open source already entails in other areas, and what other political and social concerns might arise if our cyborg world is not just any human-machine interlinkage, but dominantly constituted by information and computing devices (Chopra & Dexter 2008, 148). The issue of forking is, thus, also relevant.

Forking, at least in software development, is undesirable and under social pressure (Raymond 2001, 73), but it is also an essential right in open-source projects (Fogel 2005, 5) if the latter is about autonomy and community work. Among software developers, major or hard forks have not been frequent (Weber 2004, 134). If we take seriously the idea that cyborgs with implantable and other technologies need both open-source developments and all the help with the IoT where relevant data and free information exists, forking becomes a question about societies, culture, and politics. Because of changes in open source policies, such as forking, it could lead to paternalism, or if such questions are overlooked, could bring harm to cyborgs. Thus, there would be a need for addressing technological development in-depth. Furthermore, we must look at how new vulnerabilities benefit from technologies, but also what other questions might rise on a social and political level.

Conclusion

Next to other pathways to the posthuman future, cyborgization is often seen as an ongoing process, in which we tamper with biology and evolution through technologies. Ray Kurwzeil's vision is well-known: computational devices move from desks to pockets, from pockets to our bodies, until there is more machine than human (Kurzweil 2005). The

question here has not been about the likelihood of such an event, but an exploration of what cyborgization means advanced from the current, computationally ubiquitous, present. Looking at the possibility of designing human advancement with technologies, we should be interested in both the nature of technologies and the design practices. In terms of technologies, some transhumanist scholars, I believe, are on the right track when they inquire about the significance of data and internet, raising philosophical problems on the way to the future (Sorgner 2017; Sorgner 2021); or when they attempt to situate the utopian idea of mind uploading into a concrete technologically secure solution (Swan 2015, 2019).

This paper focused on the open-source development model, found in software engineering and elsewhere. It adds to the process of cyborgization the idea that we are not only designing humans with technologies, but have ways of developing the technologies that count as a part of the former. This is just one way of seeing how the internet, and different actors, from developers to users, constitute a complexity of cyborgization. Open-source development, often associated with the Linux operating system, has, from early on, taken advantage of the internet, and development has been described as co-evolution, binding together communities and structures of social and technical (Tuomi 2002). Nevertheless, the proposal to take this model into consideration should not be merely based on its current success in innovation. As I have shown, the reason for taking open source seriously arises from (new) vulnerabilities, related to implantable technologies. Decisions made by companies and professionals live within people or are deeply attached to them. The argued solution could be open-source technologies, freely modifiable by the other, giving the wearer a sense of freedom to choose experts, and autonomy. However, if this solution is to be applicable to more of the present and future cyborg technologies, it would require more and quicker minds and hands, at least for some time into the future. For this I have proposed how we could implement data exchange layer technologies that could bring together various public and private institutions, communities, and information to secure the future with cyborg technologies.

Acknowledgment:
I would like to thank Oliver Laas for providing valuable feedback on the early draft of this paper.

REFERENCES

Barfield, Woodrow; Williams, Alexander (2017). "Law, Cyborgs, and Technologically Enhanced Brains." *Philosophies* 2 (1): 6. https://doi.org/10.3390/philosophies 2010006.

Chopra, Samir; Dexter, Scott D. (2008). *Decoding Liberation: The Promise of Free and Open Source Software*. Routledge Studies in New Media and Cyberculture 4. New York London: Routledge, Taylor & Francis Group.

———. (2007). "Free Software and the Political Philosophy of the Cyborg World." *ACM SIGCAS Computers and Society* 37 (2): 41–52. https://doi.org/10.1145/1327325. 1327328.

Clynes, Manfred E.; Kline, Nathan S. (1960). "Cyborgs and Space." *Astronautice* 14 (9): 26–27, 74–76.

Coeckelbergh, Mark (2011). "Vulnerable Cyborgs: Learning to Live with Our Dragons." *Journal of Evolution and Technology* 22 (1): 1–9.

———. (2013). *Human Being @ Risk*. Philosophy of Engineering and Technology. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-6025-7.

Cooper, Robert; Foster, Michael (1971). "Sociotechnical Systems." *American Psychologist* 26 (5): 467–74. https://doi.org/10.1037/h0031539.

De Maria, Carmelo; Di Pietro, Licia; Díaz Lantada, Andrés; Madete, June; Makobore, Philippa Ngaju; Mridha, Mannan; Ravizza, Alice; Torop, Janno; Ahluwalia, Arti (2018). "Safe Innovation: On Medical Device Legislation in Europe and Africa." *Health Policy and Technology* 7 (2): 156–65. https://doi.org/10.1016/j.hlpt.2018.01.012.

De Maria, Carmelo; Di Pietro, Licia; Ravizza, Alice; Díaz Lantada, Andrés; Ahluwalia, Arti Devi (2020). "Chapter 2 – Open-Source Medical Devices: Healthcare Solutions for Low-, Middle-, and High-Resource Settings." In *Clinical Engineering Handbook (Second Edition)*, edited by Ernesto Iadanza, 7–14. Academic Press. https://doi.org/10.1016/B978-0-12-813467-2.00002-X.

Debian Project (2004). "Debian Social Contract." https://www.debian.org/social_contract #guidelines.

Denning, Tamara; Kramer, Daniel B.; Friedman, Batya; Reynolds, Matthew R.; Gill, Brian; Kohno; Tadayoshi (2014). "CPS: Beyond Usability: Applying Value Sensitive Design Based Methods to Investigate Domain Characteristics for Security for Implantable Cardiac Devices." In *Proceedings of the 30th Annual Computer Security Applications Conference*, 426–35. New Orleans Louisiana USA: ACM. https://doi.org/10.1145/2664243.2664289.

Estonian Information System Authority (2022). "X-Road® Data Exchange Layer." X-Road® Data Exchange Layer. 2022. https://x-road.global/xroad-introduction.

Evans, Nicholas G.; Selgelid, Michael J. (2015). "Biosecurity and Open-Source Biology: The Promise and Peril of Distributed Synthetic Biological Technologies." *Science and Engineering Ethics* 21 (4): 1065–83. https://doi.org/10.1007/s11948-014-9591-3.

Featherstone, Mike, Burrows, Roger (eds.) (1995). *Cyberspace/Cyberbodies/Cyberpunk: Cultures of Technological Embodiment*. Theory, Culture & Society. London: Sage.

Fogel, Karl (2005). *Producing Open Source Software: How to Run a Successful Free Software Project*. 1st ed. Beijing; Sebastopol, CA: O'Reilly.

Franssen, Maarten; Kroes, Peter (2009). "Sociotechnical Systems." In *A Companion to the Philosophy of Technology*, edited by Jan Kyrre Berg Olsen Friis, Stig Andur Pedersen, and Vincent F. Hendricks, 223–26. Blackwell Companions to Philosophy 43. Chichester, UK ; Malden, MA: Wiley-Blackwell.

Gray, Chris Hables (2000). *Cyborg Citizen: Politics in the Posthuman Age*. New York: Routledge.

Greguric, Ivana (2014). "Ethical Issues of Human Enhancement Technologies: Cyborg Technology as the Extension of Human Biology." *Journal of Information, Communication and Ethics in Society* 12 (2): 133–48. https://doi.org/10.1108/JICES-10-2013-0040.

———. (2022). *Philosophical Issues of Human Cyborgization and the Necessity of Prolegomena on Cyborg Ethics:* Advances in Human and Social Aspects of Technology. IGI Global. https://doi.org/10.4018/978-1-7998-9231-1.

Halperin, Daniel; Heydt-Benjamin, Thomas S.; Fu, Kevin; Kohno, Tadayoshi; Maisel, William H. (2008). "Security and Privacy for Implantable Medical Devices." *IEEE Pervasive Computing* 7 (1): 30–39. https://doi.org/10.1109/MPRV.2008.16.

Haraway, Donna Jeanne (1991). *Simians, Cyborgs, and Women: The Reinvention of Nature*. New York: Routledge.

Hessel, Andrew (2006). "Open Source Biology." In *Open Sources 2.0: The Continuing Evolution*, edited by Chris DiBona, Danese Cooper, and Mark Stone, 1st ed, 281–96. Beijing; Sebastopol, CA: O'Reilly.

Hope, Janet (2008). *Biobazaar: The Open Source Revolution and Biotechnology*. Cambridge, Mass: Harvard University Press.

Irwin, Alan (1995). *Citizen science: a study of people, expertise, and sustainable development*. 1. publ. Environment and society. London: Routledge.

Kurzweil, Ray (2005). *The Singularity Is Near: When Humans Transcend Biology*. New York: Viking.

McLaughlin, James Alastair; Myers, Chris J.; Zundel, Zach; Mısırlı, Göksel; Zhang, Michael; Ofiteru, Irina Dana; Goñi-Moreno, Angel; Wipat, Anil (2018). "SynBioHub: A Standards-Enabled Design Repository for Synthetic Biology." *ACS Synthetic Biology* 7 (2): 682–88. https://doi.org/10.1021/acssynbio.7b00403.

Open Source Hardware Association (2022). "Open Source Hardware Definition." Open Source Hardware Association. 2022. https://www.oshwa.org/definition/.

Open Source Initiative (2007). "The Open Source Definition." 2007. https://opensource.org/osd.

Oudshoorn, Nelly (2015). "Sustaining Cyborgs: Sensing and Tuning Agencies of Pacemakers and Implantable Cardioverter Defibrillators." *Social Studies of Science* 45 (1): 56–76. https://doi.org/10.1177/0306312714557377.

———. (2016). "The Vulnerability of Cyborgs: The Case of ICD Shocks." *Science, Technology, & Human Values* 41 (5): 767–92. https://doi.org/10.1177/0162243916633755.

Paide, Karoline; Pappel, Ingrid; Vainsalu, Heiko; Draheim, Dirk (2018). "On the Systematic Exploitation of the Estonian Data Exchange Layer X-Road for Strengthening Public-Private Partnerships." In *Proceedings of the 11th International Conference on Theory and Practice of Electronic Governance*, 34–41. Galway Ireland: ACM. https://doi.org/10.1145/3209415.3209441.

Payne, Christian (2002). "On the Security of Open Source Software." *Information Systems Journal* 12 (1): 61–78. https://doi.org/10.1046/j.1365-2575.2002.00118.x.

Pycroft, Laurie; Aziz, Tipu Z. (2018). "Security of Implantable Medical Devices with Wireless Connections: The Dangers of Cyber-Attacks." *Expert Review of Medical Devices* 15 (6): 403–6. https://doi.org/10.1080/17434440.2018.1483235.

Rayes, Ammar; Salam, Samer (2022). *Internet of Things from Hype to Reality: The Road to Digitization*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-90158-5.

Raymond, Eric Steven (2001). *The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary*. Rev. ed. Beijing ; Cambridge, Mass: O'Reilly.

———. (2020). "How To Become A Hacker." Thyrsus Enterprises. 2020. http://catb.org/~esr/faqs/hacker-howto.html.

Rodriguez-Blanco, Yiliam F.; Souki, Fouad; Tamayo, Evelyn; Candiotti, Keith (2013). "Magnets and Implantable Cardioverter Defibrillators: What's the Problem?" *Annals of Cardiac Anaesthesia* 16 (1): 54. https://doi.org/10.4103/0971-9784.105372.

Sandler, Karen; Moe, Marie (2021). "'Don't Mess With My Heart Device, I'll Do It Myself.' In Which Karen and Marie Interview Each Other." In *Modified: Living as a Cyborg*, edited by Chris Hables Gray, Heidi Figueroa-Sarriera, and Steven Mentor, First, 101–7. New York, NY: Routledge.

Sandler, Karen; Ohrstrom, Lysandra; Moy, Laura; McVay, Robert (2010). "Killed by Code: Software Transparency in Implantable Medical Devices." *Software Freedom Law Center*, 12.

Schryen, Guido; Kadura, Rouven (2009). "Open Source Vs. Closed Source Software: Towards Measuring Security." In *Proceedings of the 2009 ACM Symposium on Applied Computing*, 2016–23. SAC '09. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/1529282.1529731.

Sorgner, Stefan Lorenz (2017). "Genetic Privacy, Big Gene Data, and the Internet Panopticon." *Journal of Posthuman Studies* 1 (1): 87. https://doi.org/10.5325/jpoststud.1.1.0087.

Sorgner, Stefan Lorenz (2021). *We Have Always Been Cyborgs: Digital Data, Gene Technologies, and an Ethics of Transhumanism*. Bristol: Bristol University Press.

Strickland, Eliza; Harris, Mark (2022). "Their Bionic Eyes Are Now Obsolete and Unsupported." *IEEE Spectrum*. February 15, 2022. https://spectrum.ieee.org/bionic-eye-obsolete.

Swan, Melanie (2015). "Blockchain Thinking: The Brain as a Decentralized Autonomous Corporation [Commentary]." *IEEE Technology and Society Magazine* 34 (4): 41–52. https://doi.org/10.1109/MTS.2015.2494358.

———. (2019). "Transhuman Crypto Cloudminds." In *The Transhumanism Handbook*, edited by Newton Lee, 513–28. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-16920-6.

Tuomi, Ilkka (2002). *Networks of Innovation: Change and Meaning in the Age of the Internet*. Oxford; New York: Oxford University Press.

Weber, Steve (2004). *The Success of Open Source*. Cambridge, MA: Harvard University Press.

SEX AND TRANSHUMANISM

MARIUS MARKUCKAS[1]

Abstract

In this article, the transhumanist project is analyzed through the lens of one of the phenomena characterizing human existence, *i.e.* sex. The aim of the article is to reconstruct the ontological structure of transhumanism employing the reflection on the phenomenon of sex and, based on this reconstruction, to prepare the conceptual fundament for the most theoretically justified ethical treatment of the transhumanist project. The genealogical studies by Michel Foucault and historical-ontological research by Ian Hacking are used as the theoretical basis and methodological guideline for the analysis carried out in the article. The analysis demonstrates that due to the ontologically 'mobile' nature of the main object of transhumanism (the human), transhumanism itself unfolds as a project that is constantly shifting in terms of its purposes and goals. Therefore, transhumanism is a project that cannot be realized in practice in principle. The impossibility of realizing the transhumanist project, consequently provides an ethically engaged treatment thereof. The article is concluded by stating transhumanism to be a nihilistic ideology – or, paradoxically, an ideology without an ideal – which actually does not manifest as the improvement of the human, but only as their perpetual identity and anthropological remaking based on new ever emerging techno-scientific capabilities.

Key words: *transhumanism; sex; science; technology; Foucault; Hacking.*

Back in the middle of the 20th century, the well-known English scientist Julian Huxley put forward the idea that the existing human nature could potentially be radically remade while simultaneously improving the whole of humanity with the assistance of scientific and technological knowledge. Huxley's insights are currently attracting

[1] PhD in philosophy. E-mail: <marius.markuckas@gmail.com>. Orcid ID: 0000-0002-9067-7001.

increasing academic and public interest. The idea of transhumanism has been widely analyzed in scientific research, expressed in artistic pieces, and has also become a part of political agendas[2]. It is the latter which demonstrates that the idea of transhumanism has a strong project-oriented potential. Not only has this idea been treated as an object of philosophical considerations, there have also been intensive searches for technological and sociopolitical ways to realize it in practice as well.

It is noteworthy that opposite ethical camps have formed in relation to the phenomenon of transhumanism. One of those camps, represented by transhumanism apologists (many of them can be named simply as 'transhumanists'), emphasizes the advantages of transhumanism as a project of remaking the human. The other one, represented by the critics of transhumanism, the so-called 'bioconservatives', point out the threats to humans and to the whole of humanity that may arise from the transhumanist project. Despite these differences, none of these camps is inclined to doubt the possibility of the practical realization of the idea of transhumanism. The ethical issues of transhumanism are not directly touched upon in this article. Moreover, the aim of this article is not to find the answer to the question of which of these ethical positions – transhumanist or bioconservative – is more justified. The analysis carried out in this article has a different goal. It is to reconstruct the ontological structure of transhumanism as an analytical perspective using the phenomenon of sex, *i.e.* one of the existential elements, which is characteristic of the human, and based on this reconstruction, to open the possibility for the most theoretically grounded ethical treatment of transhumanism[3].

---

[2] Evidence for the readily increasing social and political significance of transhumanism can be clearly seen in the fact that an increasing number of political parties on a global scale are determined to present transhumanist ideas directly or at least respond to it in their political programs. The growing significance of these ideas and their intensifying development on a political level is clearly testified by the presence of studies that are devoted to the analysis of political parties representing the transhumanist worldview (*e.g.*, see Szabados 2018).

[3] This article relies on certain research insights and results of the doctoral dissertation prepared and defended by the author of the article, devoted to the analysis of the idea of transhumanism from the perspective of historical ontology (Markuckas 2022). Thus, the analysis carried out in the article can be treated as, on the one hand, the presentation of

First of all, this article is directed at the aforementioned goal because an ethically oriented research of transhumanism can only be conducted after a detailed analysis of ontological structure thereof. It is precisely for this reason that the analysis carried out in the article is focused on the ontological aspects of transhumanism rather than the ethical ones. As the theoretical basis and as a methodological guideline of the analysis, the genealogical studies by Michel Foucault and historical-ontological research by Ian Hacking have been chosen.

The Transhumanist Project: Its Idealistic and Realistic Aspects

In the rapidly growing field of transhumanism studies, it would indeed be difficult, if possible at all, to detect an aspect of transhumanism that has not yet received the attention of researchers. However, extensive research on transhumanism – including the analysis and assessment of both theoretical and practical aspects of the transhumanist project – hardly helps in finding an answer to the most fundamental question: what is transhumanism? Current works on transhumanism are lacking consensus not only on how transhumanism should be defined or elucidated, but also what essential practical goals should be ascribed to the transumanist project.

On the one hand, transhumanism is presented as a particularly ambitious project, which is focused on the radical remaking of the existing human, and which should actually manifest as the elimination of all physical and mental limitations in humans. This overcoming of current human anthropological characteristics would be so profound that it would ultimately enable the transformation of a human into a new kind of entity – the so-called 'posthuman'. Some theorists of transhumanism have pointed out that there are only different visions of what the future posthuman will be like or what form its existence will take in practice (Bostrom 2003, 5; Ranisch, Sorgner 2014, 8). However,

---

the main ideas of the research conducted in the aforementioned dissertation, and, on the other hand, as the further elaboration of some of its theoretical aspects, primarily in order to explicate in more detail the fundamental link between the necessity of cognizing the ontological structure of transhumanism and the justification of its ethical treatment.

these visions usually emphasize the invulnerability of the new future entity by comparing it with the existing human who is imperfect by their nature and accordingly is existentially vulnerable. Following the so-called attitude of 'epistemological certainty', transhumanists are convinced that virtually all natural human vulnerabilities along with their fundamental imperfection can be eliminated by means of scientific and technological knowledge (Ross 2020, 17). In many works of transhumanism apologists, the biological fact of human mortality is presented as the most fundamental limitation of the human being. Therefore, the most ambitious goal of the transhumanist project is to take control of the various biological processes that lead to human suffering (*e.g.*, arresting aging, eradicating various diseases, etc.), which, in turn, would eventually allow a human to become an immortal being.

On the other hand, even admitting that this image of the transhumanist project as directed at the transformation of the human into an almost 'divine', or absolutely invulnerable and flawless entity, really prevails in the presentations oriented to catch the attention of broader society, although it should be noted that there are also those apologists of transhumanism who tend to treat such visions just as a poor advertisement trick. They attempt to present the transhumanism project in a much more restrained or 'realistic'[4] manner, looking at it from a contemporary perspective. Rather than 'idealistically' proving, for example, that it is possible to turn a human into an immortal being, these transhumanism apologists orientate their theoretical efforts in another direction trying to demonstrate that the cutting-edge technology and the newest scientific knowledge can be successfully used to 'upgrade' the human, creating a more 'flourishing' personal and social existence. Therefore, in the works of such renowned researchers of transhumanism as Nick Bostrom or Stefan Sorgner, the question of whether a human can really become immortal receives much less attention than the question of if or how it is possible to extend a full-fledged existence of a human

[4] For example, Sorgner, who can be reasonably treated as one of the most prominent apologists of such 'realistic' transhumanism, characterizes the idea of immortality prevailing in transhumanist discourse just as an 'advertising tool' of transhumanism and, in turn, presents his own theoretical efforts as 'a defense of non-utopian version of transhumanism' (Sorgner 2022, 6–7).

who would remain healthy and enjoying the varied pleasures of life for a much longer time than the current conditions of their existence allow to do (*e.g.*, see Bostrom 2008; Sorgner 2022).

It would be reasonable to consider whether this 'realistic' presentation of transhumanism is not just another advertising strategy. Such a strategy is aimed at the consolidation of the idea that the remaking of the existing human into a less vulnerable being capable of achieving a better existential state is not only possible, but desired by most people as well. It is really hard to deny that one of the greatest fears of humans is the fear of death. Thomas Hobbes, one of the fathers of modern philosophy, had emphasized this fact. In his famous book *Leviathan*, Hobbes attributed the essential characteristic of fear of death to humans and grounded his own social and political theory on the basis of this characteristic. In its turn, it explains why some proponents of transhumanism ('idealists' of transhumanism) tend to present transhumanism as an ambitious effort to make humans immortal. However, Hobbes was not only talking about the fact that people fear death and therefore they are inclined to be social or even sacrifice some of their natural freedom in order to create and maintain a public and civic apparatus – the state. Hobbes also drew attention to the fact that people behave in this way because they pursue 'ease' and 'sensual delight' (Hobbes 1965, 76).

In a Hobbesian manner, the 'realists' of transhumanism emphasizes that in everyday life, people tend to reflect not on death, but how to avoid all kinds of discomfort and suffering, to experience as many joyous events in life as possible, to remain alive and healthy. As we go about our lives, we seek all sorts of ways to enrich our daily existence as much as possible—or, to put it in Hobbes' terms, to make life easier and more pleasant. This assumption is supported by research showing that aging people, whose lives have become more complex, less self-controlled, lonely, and increasingly dependent on the will and efforts of others, who lose some of the human life pleasures and needs, such as the ability to communicate, feeling needed, etc., tend to view death much more favorably (*e.g.* see Rurup, Pasman, *et al*. 2011). For reasons close to the latter, *i.e.* emphasizing the suffering and disabilities caused by various diseases, the loss of one's personal identity and the fear of the future, some people also ground their support for the option of ending life

voluntarily with the assistance of medical staff for anyone who has lost the joy of their existence and is drowning in constant misery (*e.g.* see Pearlman, Hsu, *et al*. 2005).

This suggests that it is not just 'life', but namely 'the good life' is what most people really value. As is well known, since at least the time of the ancient Greeks, the question on the essence of 'a good life' has been raised and intensively debated in philosophy. It is clear that the Platonic Socrates and Protagoras would have never fully agreed on that what constitutes the good. Therefore, it is not surprising that this debate in philosophy still exists. Truly, in the paradigm of transhumanist thinking, which, like Hobbes's philosophy, is based on a materialistic view of the human, the above-mentioned question on the 'a good life' does not take the form of such sharp discussions and accordingly presupposes a fairly clear answer. The old Hobbes and most of the transhumanists of our day share the common association – through turning their thinking in a materialistic direction all over – between relating the human well-being entirely with various parameters of human physical existence and the possibility of transforming these parameters. This principle is clearly revealed in the works of proponents of 'realistic' transhumanism. On the one hand, these works noticeably avoid considering, for example, the most fundamental human existential goals or the role of human mortality within the constitution of human existence. In this respect, Sorgner's book *We Have Always Been Cyborgs* is representative. There, Sorgner simply acknowledges the existing plurality of aspirations and ideals in human lives (Sorgner 2022, 135–138). On the other hand, in such 'realistically-oriented' works it is clearly emphasized that one of the most, if not the most, important goals of transhumanism is as aimed at the remaking of humans maximazing the extension of their health span, but not just life span.

Of course, like almost every concept in the field of philosophical reflection, the concept of 'health' is constituted by a complex discursive as well. Therein, the content of this concept is intensively contested and includes such broad currents of philosophical thinking as the philosophy of medicine and the philosophy of healthcare. So, it is not surprising that there are many different conceptions and definitions of health. However, without going into a more detailed analysis of this large field

of discussion of the concept of health, in this article it will be enough to note that in the aforementioned field prevails the general idea that health is much more than physical and psychological human wellness. This idea is clearly presented in the definition of health provided by the World Health Organization (1946 [1948]), where it is stated that health is the state of physical, mental and social well-being, and not just the absence of disease or infirmity. It should be noted that such a concept of health is clearly characteristic of the theoretical approaches of Bostrom and Sorgner. While presenting and discussing the transhumanist project through the perspective of extending the human health span, they refer, on the one hand, to those elements of human existence that testify to the fact of a person being physically undamaged by disease and infirmity, and, on the other hand, to those that clearly attests to their mental and social well-being as the ability to enjoy oneself and other persons, through a physically and mentally fulfilling coexistence with them. Looking from the transhumanist perspective, the phenomenon of sex can be treated as such an existential element, which combines these two strands of the comprehension of the phenomenon of health, and which accordingly helps transhumanists to present their notion of a healthy, life-fulfilling human condition in a 'realistic', or understandable way, even to those people, who by nature are not very imaginative or future-minded.

Sex as a Tool for the Representation of Transhumanism

Zoltan Istvan, a futurist and renowned researcher of transhumanism, who also makes great efforts to apply the ideas of transhumanism at the social and political level[5], claims that transhumanism has long focused on technology that ensures improved health and extended longevity (Istvan 2014). However, according to Istvan, we should not forget that in the future, transhumanism will make some of our most cherished

[5] Istvan is the founder of the world's most famous political party representing the transhumanist worldview, the *Transhumanist Party* in the United States of America. He also ran as a candidate in the election campaign for the President of the United States of America in 2016.

experiences – such as sex – much more fun and enjoyable. Here, Istvan discusses the specific technological devices that are already being developed, from microchip implants to exoskeleton suits, demonstrating that with the development of transhumanism the horizon of human sexual experiences will also expand. It should be noted that in the context of the implementation of the transhumanism project, the question of evolving sexual experiences is receiving attention not only from Istvan, but also from other researchers, who intensively consider how the novel technology will affect the phenomenon of sex, as well as the human existence (*e.g.* Haylock 2017; Schussler 2017; Hauskeller 2014).

Although it may appear that questions relating to the phenomenon of sex and raised in the context of transhumanism are only technologically engaged, this view would certainly be too narrow. It is true that the phenomenon of sex in studies of transhumanism is analyzed to better understand, for example, how the process of sex would look in the future or how through the technological development evolving sexual practices could transform our attitude towards sex itself. Nevertheless, such analyzes are also carried out in order to better elucidate and present the very idea of transhumanism. Moreover, through problematising the phenomenon of sex, ethically engaged evaluative positions in relation to transhumanism can be formed and expressed.

Reflection on sex thus enables both the apologetics of transhumanism and its critique. It should be conceded that there are still not many works where the phenomenon of sex is used as a tool for the analysis of transhumanism. However, in this growing specific thematic field it is already possible to see the existing difference between the bioconservative and transhumanist approaches, which also structurizes the entire transhumanist discourse. At the same time, it should be noted that although the aforementioned distinction can be observed, the analysis of the phenomenon of sex is still more intensely used to justify the relevance or necessity of the transhumanist project, but not to engage critically with it.

While representing their stance, the apologists of transhumanism exploit linguistic motifs and images associated with sex. This is not a simple coincidence. As mentioned before, by following materialistic

anthropology transhumanists perceive the human only as a corporeal being, who, just similarly like other beings of nature, desires a life that is as pleasant, long-lasting and joyful as possible. Of course, as John Stuart Mill has made it clear a few centuries ago, and as at least some proponents of transhumanism agree, the pleasures that make up human happiness are actually different in its quality (Mill 2009, 19). Based on the aforementioned, it could be argued that despite the tendency of cultural posthumanism to provide intensive critique of the dualistic and hierarchical approach to reality, which through the idea of hybridization also spreads in the paradigm of transhumanist thinking (for example, in rejecting the ontological distinctions between humans and technology, as well as humans and animals), we still live in a cultural context, where it is accepted, as in the times of Mill, that the satisfaction of a human can hardly be compared or equated with the satisfaction of a pig.

On the other hand, even if it is stated that there are indeed significant differences between the existence of humans and other, it should also be noted that in the works of the proponents of transhumanism, these are mentioned only in exceptional cases and not discussed in a detailed manner. While the human is sometimes recognized as possessing intellectual or moral qualities that distinguish them from other beings living in the world, this does not change the general anthropological direction of transhumanist thinking, which aims to reduce the human to a pure biological being, whose main existential goal is to experience as many physiologically-determined pleasures as intensely as possible. It is for this reason the references found in apologetically-oriented works of transhumanism to the limitedness of our current sexual experiences and the possibility to enhance it technologically or to grow sexual capacities in people (Bostrom 2003, 18; Sorgner 2022, 7, 12) cannot be considered as a mere advertising trick intended to make academic publications on transhumanism catchier for the readers or enticing them to become excited about the idea of transhumanism. Looking at these references from a broader perspective of transhumanist thinking, they should be considered as constituting the transhumanist anthropology, which assumes that only a person who is able of enjoying various pleasures (including the pleasure of sex) can be treated as a free and unlimited being, capable of experiencing the fullness and perfection of existence.

Sex and the Ontological Problematicity of Transhumanism

In the works of transhumanism apologists the phenomenon of sex is used in order to present the transhumanist project positively. There, sex acts as a tool to convey the message about the most important aim of transhumanism to transform the human state as blessed as possible. Likewise, it helps to establish the anthropological attitude that this state in particular, where people are able to enjoy pleasures unlimitedly, most adequately responses their existential interests and allows them to realize the existential human potential at a maximum scope. Without denying the element of rationality behind such an attitude or by agreeing in principle that human happiness may also be related to the experience of bodily structured pleasures, the analysis carried out in this article demonstrates that the phenomenon of sex may be used not only for transhumanism apologetics, but also for critique thereof.

Thinking about transhumanism from a critically-oriented stance and employing the phenomenon of sex as an analytical perspective, it is expedient to consider Foucault's works. Foucault skeptically looked at the idea rising from the paradigm of humanistic thinking that it is possible to acquire objective knowledge of the human and, and accordingly establish guidelines of ethical behavior and a sociopolitical state responding to 'true' human nature. Foucault was sure that there is no reason to talk about the 'true' human – or 'human in itself'. According to Foucault, such efforts to truly define the human only represent discourse prevailing at a specific historical time, or in his own terms, a certain 'regime of truth'. Therefore, in the work *The Order of Things*, which is devoted to the consideration of the ontological state of human sciences, Foucault declares that "[t]o all those who still wish to talk about man […] to all those who asks themselves questions about what man is in his essence […] to all those who refer all knowledge back to the truths of man himself […] to all these warped and twisted forms of reflection we can answer only with a philosophical laugh – which means, to a certain extent, a silent one." (Foucault 2004, 373).

In the context of transhumanism analysis, should be emphasized that this insight by Foucault can be applied not only to all classical philosophically-grounded anthropological views and visions, but also

to the modern scientific and expert knowledge constituting the comprehension of the current human on which the transhumanist project is based. Foucault began his own genealogical research explicating from the point that knowledge cannot be distinguished from the relations of power and that knowledge accordingly spreads as a process of subjective construction of reality rather than a process of objective cognition of it, by accepting only the truths that conform to the prevailing regime of truth. In this regard, Foucault's *The History of Sexuality* should be considered as an especially important work. Here, Foucault employs the phenomenon of sex as a core of analysis, and genealogically explicates the constitutive power of scientific knowledge, both to the reality and to the human as an integral part of it. By analyzing scientific works on human sexuality of the 18-19th centuries, Foucault convincingly demonstrated that from the time that experts took an interest in the phenomenon of sex, and our conception of sexuality became controlled by the prevailing desirable social 'ideal', which also can be presented in the name of 'norm' in accordance to the broadening tendency of modernity to assess all aspects of social reality through the lens of statistically-oriented thinking, our sexual practices are highly determined of certain knowledge intimately related with biopolitical goals (Foucault 1978, 115–133).

Through the use of the genealogical method in explicating the constitutive nature of the scientific and expert knowledge towards the human and reality, while also demonstrating how such knowledge forms 'appropriate' models of human behavior, Foucault clearly disclosed that historically evolving knowledge transforms the field of possibilities of conceptualization of various social phenomena and simultaneously reveals completely new horizons of 'ideality' (or 'normality') thereof. All these processes were also thoroughly analyzed in the historical-ontological research carried out by the philosopher of science Ian Hacking, a follower of Foucault's ideas. Hacking uses the conception of the 'moving target' to emphasize the dependence of peoples self-comprehension and their behavior on historically evolving scientific and expert knowledge (Hacking 2007). The aforementioned conception in the context of Hacking's philosophy is also employed to highlight epistemological and ontological differences between objects of

natural sciences and human sciences, which are accordingly named as 'natural kinds' and 'human kinds'. Hacking describes human kinds as 'interactive' and natural kinds as 'indifferent'. By stating these differences, Hacking emphasizes that scientifically-oriented language applied to humans has a very different ontological effect than, for instance, for a quark. Unlike quarks, people are conscious beings so they react to the scientific-expert language describing them (Hacking 1999, 59; 105). On the basis of these descriptions and the reactions sparked thereof, peoples' self-comprehension and behavior is changing. To put it in broader anthropological terms, it could be said that during this process even the conception of the human is transformed. This is why humans should be treated epistemologically and ontologically as 'moving targets' rather than 'stable objects'. Similarly to the views of Foucault, it accordingly enables to state that due to the characteristic of humans to be epistemologically and ontologically 'mobile', it is impossible in principle to provide their 'objective' scientific description.

These conceptual resources found in the fields of social philosophy and philosophy of science can be employed in order to reconstruct the ontological nature of transhumanism. As previously mentioned in this article, transhumanism is based on the fundamental premise that science and technology will eventually have the capability to overcome all vulnerabilities of humans and could even transform the human into a new kind of being altogether, enjoying one's existence without the yoke of any limitations. Such huge trust in science characterizing transhumanism, despite the fact of existing variety of its notions, for example, discussed comprehensively by Sorgner (2022, 129–138), enables us to consider transhumanism as the successor of Renaissance and Enlightenment humanism. Both humanism and transhumanism as ideological projects emerging from the aforementioned historical and cultural epochs, share the focus on the radical 'improvement' of the human, as well as the social state. These projects are grounded on the belief that this can be achieved by developing and applying scientific knowledge. On the other hand, humanism and transhumanism are clearly distinguished from each other by the 'implements' they offer to eliminate human imperfection. The most prominent humanists believed in the classical power of education, while the contemporary transhumanists place their greatest

hope in the use of the latest technology. It is the reason why the presentation of transhumanism as 'humanism on Viagra' – as described by Sorgner (2015) in one of his TEDx talks – is by no means just a catchy and easily remembered rhetorical figure helping to better understand the relationship between humanism and transhumanism. First of all, such a presentation of transhumanism conveys it essentially as a technologically mediated project aimed at increasing existential potential of the human.

However, when this fundamental goal of transhumanism is reflected on through the lens of Foucault and Hacking's explication of the enormous significance of science in constituting the phenomenon of humanness, it does not take long to expose the conceptual and practical problematicity of this project. The problematicity arises from the fact that scientific and technological knowledge substantiating transhumanism is not 'solid', but 'liquid'. Accordingly, the roots of the ontological 'fruits', such as the conception of the human, as well as the notions of their capacities and vulnerabilities, which 'grow' in the soil of this knowledge, are not strong. They just do not have any stable metaphysical ground where to root more firmly. This is due to the nature of modern knowledge itself. As is well known, the knowledge from the perspective of modernity is treated as continuously progressing, and the progress of knowledge is highly praised by all supporters of transhumanism. As the knowledge base changes in the constant process of history, even if it is conceptualized as 'progress', the roots ripening the aforementioned ontological fruits are inevitably cut off. Bearing all this in mind, we need to raise the question: if then is it possible at all to reach any 'ideal' or even to assess objectively the 'improvement' of the human?

Sex can help to explain the importance of the raised question in more detail, highlighting the grand ontological challenge confronting the transhumanist project. It is enough to take a look at how our attitude towards sex has changed in the course of history. There is no doubt that throughout the history of mankind, sex has been one of the greatest physical pleasures that a person can experience, and this is still the case in the contemporary world. Nevertheless, in spite of this historically stable fact of the pleasurability of sex, the treatment of the sex phenomenon itself – through the prism of the actualization of the fullness of human existence – has undoubtedly undergone historical

changes. To prove this, we can refer to the beginning of Plato's *Republic*, *i.e.*, the philosophical masterpiece that for many centuries ahead, reaching even modern philosophical thought, has established the notion of a human as a rational being. There, occurs a conversation between Cephalus and Socrates. Cephalus, who is on the 'age's threshold' during this conversation makes a point to Socrates and the other listeners that as old age sets a person free from unnecessary passions, that person acquires complete peace. Cephalus employs the words of poet Sophocles to express his attitude. According to him, when someone once asked Sophocles, "how is he as far as sex goes and can he still make love to a woman", Sophocles replied with the words that "he is very glad to have escaped from all that, like a slave who had escaped from a deranged and savage master" (Plato 2004, 329a).

It should be admitted that these words written by Plato sound rather strange, in the context of current culture structured by the growing social importance of the transhumanist project and the idea of 'humanism on Viagra' defining thereof. This is primarily because contemporary humans are constituted, *i.e.* look at themselves and the surrounding world through the prism of different knowledge and existential possibilities based on it. Today, people have access to all kinds of technology (including in the form of certain medications, such as the Viagra, if we are talking specifically about men), which enable them to remain sexually active and enjoy sex for an unlimited period of life. Perhaps many modern people would not agree with Cephalus' view that, in this way, they only drown in unnecessary passions and lose their peace. Even more so, it is wondered if anyone today would agree with the self-improvement actions allegedly performed by the famous Greek philosopher Democritus, who, as it was said in ancient times, has recognized his own weakness in the sense that he could not look at women without experiencing disturbing desires, and decided to destroy his vision (perhaps as a technology using a bronze sun-reflecting mirror directed at eyes) to correct it (Chitwood 2004, 119).

However, despite all these doubts, we should still debate whether by improving one aspect of our existence, we are at risk of impoverishing another. Just by observing such doubts, there is reason to consider whether the ideal of human 'improvement' intensely proclaimed by

the transhumanist project is only the 'servicing' of the specific 'regime of truth' on the human prevailing in a certain historical time. In relation to the human, such servicing manifests as a coercive techno-scientifically-oriented social practice, which, by inseparably combining the idea of human perfection with technological progress and the application of the latest technology in regard to the individual, focuses on their identity or even anthropological remaking. At the same time, it must be considered whether this inevitably restricts the field of possibilities of comprehension of human perfection, as the variety of which is an important, if not the main, condition of human freedom. All these considerations are based on the in this article already elucidated fact that the fundamentally different conceptions of the human and their powers and vulnerabilities dominate in different historical and cultural epochs. Accordingly, these anthropological visions cannot be reconsolidated each to another because they are grounded on the different scientific and technological capabilities (or 'knowledge') enabling to conceptualize the human.

A transhumanist paradigm of thinking that avoids such kind of considerations can reasonably be described as 'ontologically naïve', employing the phrase elucidating the ontological stance of transhumanism as proposed by Thomas D. Philbeck (2014, 180)[6]. The relevance of such description is determined by the fact that the transhumanist thinking tends to 'objectify' the human in the old humanistic manner and, on that

[6] It is noteworthy that in the works of at least some supporters of transhumanism, it is emphasized that transhumanism should be treated rather from a 'liquid' than 'stable' perspective of ontology. For example, Sorgner devotes considerable analytical attention to the relationship between transhumanist thinking and the ontology of becoming in presenting his conception of transhumanism (Sorgner 2022). Regardless, it in no way eliminates and, to the contrary, even reveals in a more vivid manner the contradictoriness of transhumanist thinking arising from, on the one hand, persistent attempts to suggest that the existing ontological boundaries of the human can indeed be 'transcended', remaking them into a 'higher' kind of being (which is precisely what the categories of 'transhuman' or 'posthuman' denote) and, on the other hand, the recognition that the human is a constantly ontologically changing (or 'becoming') being, without any clearly fixed ontological boundaries. In turn, the problem of the contradictoriness of the idea of transhumanism was analyzed in detail in the doctoral dissertation of the author of this article (Markuckas 2022).

basis, announce the possibility of fundamental human 'improvement'. Transhumanism apologists do not consider enough the circumstance which caught the attention not only of Foucault, Hacking or Philbeck, but also of Yoni Van Den Eede (2015) – *i.e.* that the ontological structure of the human is 'moving' historically, or undergoes profound transformations alongside changes in scientific and technological knowledge. Consequently, the lack of such consideration in in the perspective of transhumanist thinking determines the lack of reflection that transhumanism itself manifests as a project shifting in terms of its purposes and goals due to the ontological 'mobility' of the main object thereof (the human), and therefore, this project cannot be realized in practice in principle.

Conclusions

The analysis of transhumanism conducted in this article revealed the ontological problematicity of this project. The aforementioned problematicity is primarily related to the ontological 'mobility' of the human as an object of cognition. As it is demonstrated in this article, such a 'mobility' of the human grounding the ontological problematicity of transhumanism can be methodologically analyzed and explicated by employing philosophical resources of Foucault and Hacking. In its turn, the phenomenon of sex can be reasonably considered as the analytical starting point whereof it becomes possible to elucidate the conceptual and practical premises underlying the ontological challenges confronted by transhumanism. It should be emphasized that the research carried out in this article is still one of the few academic works in the rapidly growing field of transhumanism studies that considers the ontological status of transhumanism and reveals the ontological problematicity of this project. In the course of further research of transhumanism, it is necessary to analyze the ontological structure in more detail thereof. Providing such analyses is the main condition in seeking to adequately define the idea of transhumanism and to assess most reasonably the project based on this idea in both aspects, the possibilities of its practical realization and its ethical status. The conclusion of this article, namely, that due to the ontological 'mobility' of its main object (the human) transhumanism

reveals itself as a project that shifts in terms of its purposes and goals, and therefore, cannot be realized in practice in principle, consequently, establishes a conceptual fundament for an ethically engaged treatment thereof. It enables us to consider transhumanism as a nihilistic ideology – or, paradoxically, an ideology without an ideal – which actually manifests not as the improvement of the human, but as their perpetual identity and anthropological remaking based on ever new emerging techno-scientific capabilities. As an attempt to ground such a treatment of transhumanism more firmly, the forthcoming research on transhumanism should be directed at a more detailed analysis of the relationship between the types of ideological and transhumanist thinking as well as of the relationship of these types of thinking to nihilistic thinking.

REFERENCES

Bostrom, N. (2003). "The Transhumanist FAQ. A General Introduction (Version 2.1)". Available at <https://www.nickbostrom.com/views/transhumanist.pdf> last time accessed in 11-11-2022.

Bostrom, N. (2008). "Why I Want to be a Posthuman When I Grow Up." In: *Medical Enhancement and Posthumanity*, eds. B. Gordijn, R. Chadwik. Springer Science & Business Media, 107–137.

Chitwood, A. (2004). *Death by Philosophy: The Biographical Tradition in the Life and Death of the Archaic Philosophers Empedocles, Heraclitus, and Democritus*. Ann Arbor: The University of Michigan Press.

Eede, Y. V. D. (2015). "Where is the Human? Beyond the Enhancement Debate." In *Science, Technology, & Human Values* 40 (1): 149–162.

Hauskeller, M. (2014). *Sex and the Posthuman Condition*. London: Palgrave Pivot.

Haylock, S. (2017). "Sexual Liberation and the Emergence of Transhumanism." In *Crisis Magazine*. Available at <https://brownpelicanla.com/sexual-liberation-and-the-emergence-of-transhumanism/>, last time accessed in 13-11-2022.

Hacking, I. (1999). *The Social Construction of What?* Cambridge, Massachusetts; London, England: Harvard University Press.

Hacking, I. (2007). "Kinds of People: Moving Targets." In *Proceedings of the British Academy* 151: 285–318.

Hobbes, T. (1965). *Leviathan*. London: Oxford University Press.

Istvan, Z. (2014). "October 20. The Transhumanist Future of Sex." In *Vices.* Available at *<https://www.vice.com/en/article/ypwn4k/the-transhumanist-future-of-sex>*, last time accessed in 13-11-2022.

Markuckas, M. (2022). *Transhumanizmo idėja istorinės ontologijos perspektyvoje*. Daktaro disertacija. Vilnius: Vilniaus universiteto leidykla.

Mill, S., J. (2009). *Utilitarianism*. New Zealand, Auckland: The Floating Press.

Foucault, M. (1978). *The History of Sexuality (Volume 1: An Introduction)*. New York: Pantheon Books.

Foucault, M. (2004). *The Order of Things: An Archeology of the Human Sciences.* London and New York: Routledge.

Philbeck, T. D. (2014). "Ontology". In: *Post- and Transhumanism: An Introduction*, eds. R. Ranisch, S. L. Sorgner. Frankfurt am Main: Peter Lang, 173–184.

Ranisch, R., Sorgner, L. S. (2014). "Introducing Post- and Transhumanism." In: *Post- and Transhumanism: An Introduction*, eds. R. Ranisch, S. L. Sorgner. Frankfurt am Main: Peter Lang, 7–27.

Plato (2004). *Republic*. Indianapolis, Indiana: Hackett Publishing Company, Inc.

Schussler, A.-E. (2017). "Pornography in Transhumanism – Towards a Sexuality of Singularity." In *Postmodern Openings* 8(1): 41–56.

Sorgner, L., S. (2015). "Don't We All Wish to be Wonder Woman or Superman?" *Youtube*. Available at <https://www.youtube.com/watch?v=rFbz9zvAqqM>, last time accessed in 01-12-2022.

Sorgner, L., S. (2022). *We Have Always Been Cyborgs: Digital Data, Gene Technologies, and an Ethics of Transhumanism*. UK: Bristol University Press.

Szabados, K. (2018). "Transhumanist Parties as Niche Parties." In *Journal of Posthuman Studies* 2 (2): 213–237.

Pearlman, A. R., Hsu C., *et al*. (2005). "Motivations for Physician Assisted Suicide." In *Journal of General Internal Medicine* 20(3): 234–239.

Rurup, L. M., Pasman R. H., *et al*. (2011). "Understanding Why Older People Develop a Wish to Die: a Qualitative Interview Study." In *Crisis* 32(4): 204–216.

Ross, B. D. (2020). *The Philosophy of Transhumanism: A Critical Analysis*. UK: Emerald Publishing (*Kindle ed.*).

World Health Organization. (1946 [1948]). *Constitution.* Available at <https://www.who.int/about/governance/constitution>, last time accessed in 01-12-2022.

THE BORG – GLIMPSING INTO A NIETZSCHEAN TRANSHUMANIST UTOPIA

ANDREI NUTAȘ[1]

Abstract

Transhumanism is a cultural and philosophical movement that advocates for the use of technology to improve the human condition and overcome limitations as a species. The question of whether the transhumanist goal of a posthuman future is desirable remains open, as it depends on the values and goals that guide the process of technological enhancement. The essay inspects what it would mean to utilize nietzschean values and goals within the process of technological enhancement. Eventhough some consider Nietzche to be a proto-transhumanist, there is more then ample reason to believe that such an association is misguided. The argument presented here highlights why we need to avoid imbuing the goals of transhumanism with nietzschean values. To argue for the dangers inherent in nietzschean ideal the current work draws upon Losurdo and highlights the strongly Reactionary elements that can be found within Nietzsche's philosophy. Next the paper presents how by adding the Borg is a good representation of a society that has added technological enhancements upon a foundation of nietzschean ideal and encourages the abandonment of such utopian aspirations.

Key words: *Transhumanism, Nietzsche, Borg, Utopia, Reactionist*

"Transhumanism [is] the world's most dangerous idea" (Fukuyama 2004, 42). This is at least Francis Fukuyama's judgement concerning this cultural and philosophical movement. Transhumanism promotes an interdisciplinary approach to understanding and evaluating the opportunities of enhancing the human condition through the use of technology in order to improve human life (Bostrom 2005, 3). To achieve the transhumanist goal, we need to revolt against more than 2 millennia

[1] PhD Student. Doctoral School of Philosophy, Sociology and Political Science. West University of Timisoara. Email <andrei.nutas92@e-uvt.ro>.

of philosophical indoctrination to escape the current paradigm and to revolutionize the foundation of culture, society and most fundamentally of the self. In order to improve our current lot in accordance with the transhumanist dream, we need to take control over the course of evolution, not just as individuals but as a species, and as representatives of our planet. We need to "play god". Only by doing this can we meticulously address the problem of overcoming the limitations of our species and arrive at our destination of an improved human life.

Given the audacity of the transhumanist project it is no wander that Fukuyama sees it as a dangerous idea. The powers that transhumanist want to release in order to achieve their vision can under certain circumstance lead to catastrophes. While realizing the next step in the evolution of the Homo species is something which at first sounds as deserving of applause. We nevertheless have to think a bit more critically about what it actually implies. As a disclaimer, I am a believer in the basic transhumanist ideal that technology can better our lives. I also support the emergence of the next iteration of the homo species. What I don't support and think neither should any being with a modicum of wisdom is a world in which only one intrinsic good exists.

Having set the stage allow me to now get more specific by going through some terminology and the properties that I will be assigning to said terminology within the confines of this discussion. We need to differentiate between transhumanism and posthumanism. The reason we need to do this can be found within one of the points of contention present within this paper which regards Nietzsche's overhuman and its analogy to the posthuman (Sorgner 2022, 3). As the term posthuman is used both in transhumanist and posthumanist circles, clarification on this point is necessary.

Following Francesca Ferrando (Ferrando 2019, 27–60), we say that both transhumanism and posthumanism have as a goal the realization of the post-human. In both schools the post-human can be seen as an open-ended vision of the human. This however is where the similarity ends. Historically they are rooted in two completely different philosophical traditions: 1) transhumanism has its ancestry in the enlightenment and it shares its goals of progress and reason; 2) on the other hand posthumanism is rooted in post-modernism and it aims to deconstruct the human

which it sees as a cultural and historic notion. With these distinctions in mind, we can see that although the two schools of thought utilize the same nomenclature, namely post-human, they mean two distinct things when referring to it. Posthuman under the posthumanist tradition refers to a new definition of humanity that is post-anthropocentric and post-dualist. It is a sort of egalitarian ideal. On the other hand, for the transhumanist, it essentially means a more advanced human. Thus, there is no obligation to rid oneself of all the traditional conception about what a human is, but solely of those that block one's right to use technology and science as a means of evolution. The transhumanist view can allow for a post-dual and a post-anthropocentric approach as long as it does not come counter to the idea of enhancement via science and technology which is a core precept of the transhumanist post-human.

Due to his doctrine of the overhuman many would think that Nietzsche is a precursor of transhumanism.

> "I teach you the overman. Human being is something that must be overcome. What have you done to overcome him? All creatures so far created something beyond themselves; and you want to be the ebb of this great flood and would even rather go back to animals than overcome humans?" (Nietzsche 2006, 5)

While there is a raging debate in the field as to whether Nietzsche truly deserves to be considered a proto-transhumanist (where I side with Bostrom in saying he doesn't since he was not championing technology and science as an essential element for transcendence) I will leave this debate aside for now to focus on the implications of accepting the Nietzschean overhuman and the associated Nietzschean political ideal. In order to do this, I need to first outline what this program and ideal is. For this I derive most of what I will present from Losurdo's masterpiece "Nietzsche, the Aristocratic Rebel". Before engaging with Losurdo's interpretation of Nietzsche in particular, allow me to briefly outline what I perceive to be the three main ways of interpreting Nietzsche:

1. As a Reactionist:
    a. Good reception in reactionary Political Circles – Nazis, Fascism
    b. Bertrand Russell dismissed Nietzsche for a proto-Nazi (Russell 2004, 769 & 791)

2. Existential interpretation – points out many of Nietzsche's anti-Nazi views – Kaufmann
    a. Heidegger took up Nietzsche[2]
    b. Existential phenomenologists with egalitarian political and left wing French existentialist draw on Nietzsche[3]
3. Proto-Post-Structuralist
    a. Nietzsche's work is an analysis of power, aimed at liberating people form the strictures of coercive forms of morality
    b. Here Nietzsche is seen as somebody who wants to create healthy and creative personalities
    c. Genealogical critique of power and morality, that aimed to create spaces for an ethics of creation (Nietzsche 2001, 187)

What I will argue for is that after Losurdo it is unacceptable to consider Nietzsche in a benevolent way. Only an overly benevolent interpretation that closes an eye to the many regressive statements that can be found in Nietzsche's texts would allow us to continue to see Nietzsche as a progressive, instead of a regressive aristocratic thinker that has a superb, mesmerizing rhetoric which fibrates in sync with our frustrations and our desires to be the drivers of change, the creative, the Dionysian, the *ubermensch*, the agent.

One way to best summarize Losurdo's work comes actually from a quote from Rose:

> "It is a remarkable fact that Nietzsche, a self-professed decadent, nihilist, atheist, anti-Christ, opponent of academic philosophy, and scourge of socialism, egalitarianism, and 'the people'—who espoused aristocratic political and artistic views, insisted upon a rank of human beings, and went so far as to advise men to carry a whip when they visited the women's quarters—is today one of the highest authorities, if not the authority, for progressive liberals, existentialist theologians, professors, anarchist speculators, leftwing critics of the Enlightenment and bourgeois society, propounders of egalitarianism, and enemies of political and artistic elitism, the advance guard of women's liberationists and a multitude of contemporary movements, most if not all of which seem to have been castigated by Nietzsche's unparalleled rhetorical powers." (Rosen 1989, 189)

---

[2] Heidegger dedicated 3 books to Nietzsche.

[3] *I.e.* Derrida, Sartre, Beauvoir.

I cannot dare to state that I will give justice to the work of people like Losurdo, Rosen, Lukas, Bull, Beiner and others within the small space allotted here. All I can state is: given the constrains I face, I will offer a brief overview of what I believe to be their strongest arguments for a regressive reactionist Nietzsche.

The key issue with respects to political modernity is where one stands to the principles of the French Revolution. This is because it was a kind of egalitarian cultural revolution in the west – "1. The aim of society is the common welfare. Government is instituted in order to guarantee to man the enjoyment of his natural and imprescriptible rights. 2. These rights are equality, liberty, security, and property. […] Public relief is a sacred debt. Society owes maintenance to unfortunate [literally, unhappy: *malheureux*] citizens, either procuring work for them or providing the means of existence for those unable to labour" (Declaration of the Rights of Man and of the Citizen 1793, Article 1, 2 & 21).

Nietzsche regards the French Revolution as a sort of catastrophe for the west. He spells this out very clearly in the genealogy of morals – "the French Revolution, that gruesome and (on close consideration) pointless farce" (Nietzsche 2001, 37); "the last great slave revolt, which began with the French Revolution" (Nietzsche 2001, 45); "Judea once again triumphed over the classical ideal with the French Revolution: the last political nobility in Europe, that of the French seventeenth and eighteenth centuries, collapsed under the ressentiment-instincts of the rabble" (Nietzsche 2006, 33). He sees it as a disaster because it aims to abolish hierarchy, rank order, inherited privilege, etc. (Losurdo 2021, 55) Any civilization that aims to realize the ideals of the French revolution will be a faux civilization and culture and will destroy itself. By engaging in this sort of cultural mutation we will not be elevated to a higher moral plane but will instead become less human. (Losurdo 2021, 86–107)

While Nietzsche does support a revolution against the Ancient Regime, this revolution is just not in the direction in which liberals and socialists had hoped for. In Nietzsche's view, the problem with the Ancient Regime was that they were not tough enough to get the job done (putting liberalism and socialism back in their place). Thus, a more creative, vibrant, violent, and militant elite was needed in order to

accomplish the squashing of liberalism and socialism. Nietzsche wanted a revolution that would put the right hierarchy in place. This hierarchy would position a select few "aristocrats" at the top and everyone else at the bottom. (Losurdo 2021, 431–458) Nietzsche's project is an anti-modern project in the sense that we somewhat equate modernity to egalitarianism.

By reading Losurdo's very well documented book one can see that Nietzsche believed that unleashing the agency of certain kinds of people will lead to a more interesting and aesthetically gratifying kind of society. These people to whom agency is to be given are but a very thin elite at the top. Submission is thus what is warranted by everyone else. This view can be more attractive than the egalitarian project because here my agency and enjoyment of it is not contingent upon another's enjoyment of their agency. For the Nietzschean elite there is no limitation. Under the Nietzschean political system if an aristocrat decides that your life can be used in the pursuit of an artistic project or vision then so be it. The aristocracies artistic project trumps the submissive's right to life and right to the expression of their own artistic project.

With this in mind it follows that Nietzsche's attack on Christianity is an anti-progressist attack because under Christianity all people are equal before God, while under the Nietzschean political system some are more equal than others. Christianity, if taken to its logical conclusion will generate liberalism and socialism, a regime under which all are equal and in which no overhuman can develop. (Losurdo 2021, 459–490) Thus, there comes the necessity for the emergence of a morality, a regressive morality that looks back to the ancient master moralities. By taking this regressive stance one can revive the "true" aristocratic society. Hence, when Nietzsche is defending a European civilization that was free of the decadence of the soul of the French revolution, he is defending a world free of Christianity and its philosophical Greek fathers, he is defending a Homeric European Civilization.

Nietzsche's pessimism further underlines the need for such a strong hierarchy and limitation of agency. In his view, the masses and the elite can never both be satisfied indifferent of how much knowledge we obtain (counter to the Socratic ideal of knowledge as a universal panacea). Nietzsche holds that knowledge can grant us power over the world, but not in an optimistic way where all dissatisfaction is resolved.

Knowledge only grants us power over the world in order to build it in a way that all the dissatisfaction of an elite group of people gets resolved.

Thus, we come to Nietzsche's communitarianism (Sorgner 2022, 111–119): this consists of a few active artistic and philosophical creators and many passive workers who are there to provide for their daily needs and whims. In Nietzsche's view only by organizing society in such a way can we develop the overhuman. Here we have the problem. Because as long as we accept that there is a heritage between the Nietzschean post-human and the transhumanist posthuman, we will inescapably end up it the situation where we would develop a tyrannical polis in order to achieve our desired intrinsic good of the realization of the post-human/overhuman.

So, let us imagine a bit how such a polis could look like under the Nietzschean transhumanist normative system that would oblige us to use science and technology to develop the post-human. The first thing that comes to mind is the Borg from Star Trek. The Borg are cybernetic organisms (cyborgs) linked in a hive mind called "the Collective". They aim to continuously transcend themselves through the use of technology until they achieve perfection. The central nexus of the Borg Collective is the Borg Queen. "She brings order to the legions of voices within the Hive mind and provides a common direction — much like the queen of an insect colony" (intl.startrek.com 2011, Borg Queen). The Borg Queen has a unique personality and a sense of individuality that normal Borg drones are not allowed. She is usually the one who "speaks" for the Collective in situations where contact with outsiders is best conducted by an individual. The Queen defined herself as: "I am the beginning, the end, the one who is many. I am the Borg." Although this suggested she was an individual within the Collective, while addressing Borg drones as "my drones", she was not an individual in the sense in which we refer to it today. She is more like the module of agency or the consciousness of the Borg.

In their aim to continuously develop themselves the Borg are out on a mission to assimilate all species within the universe that could help them attain perfection. Their desire to achieve a post-Borg, or overborg state is the sole intrinsic good that matters. The Collective aims to perforate their current evolutionary equilibrium via organic, scientific,

and technological adaptation that is realized through the assimilation of other species. The Borg's dedication to a transhumanist developmental protocol can be seem starting with their way of replication. The Collective breeds their infants in maturation chambers and equip them with synthetic systems and organs from "birth". Beyond physical enhancements, the Borg are also equipped with mental enhancements. These technological enhancements rendered all Borg devoid of individual volition and caused them to work together while continuously exchanging information with the Collective. In addition, these enhancements make the individuals unable to feel emotion or to be autonomous. The cortical node ensured that emotions and individuality would be minimized amongst the Borg and, if necessary, terminates any "malfunctioning" Borgs to ensure complete obedience to the will of the Collective.

Now in Nietzschean terms, the Borg Queen is the absolute elite, the aristocrat to rule them all, the true agent and artistic soul of the species. Everything else falls under her will. All else are submissives whose role is to fulfill their ruler's desire. They exist only as objects while only the Queen, the aristocrat, exists as a subject. If we implement the Nietzschean political system, we may not end up being an intergalactic menace that destroys all life and will in the universe for the fulfillment of our ideal of creating the posthuman. But we will for sure subdue all humans, eliminate all those that disagree with us, all those that dare to express themselves outside of the confines of our established rulings, and limit people's right to procreate because certain genes should not be allowed to replicate within the species, force others to procreate because other genes are mandatory for the species, forbid personal connection beyond the depth and breadth necessary for the attainment of the posthuman ideal, limit speech by only allowing the expression of those statements that are in line with what we believe to be the truth, and many, many more horrible acts, all in the name of attaining the higher good.

If you think I am exaggerating with this, that this is just a slippery slope, I urge you to recognize that the left and the right can appropriate the Nietzschean political system. The Nietzschean political system does not discriminate between one's flavor of authoritarianism. As long as

we let one intrinsic ideal take hold of the rest, we can easily diverge into a system of tyranny the likes of which we have rarely seen.

Nietzschean Transhumanism is utopic as it believes that attaining the posthuman state will lead to a perfect future. Hence, it is defensible to engage in any sort of evil to attain this ideal because those with a "why" can endure any "how".

As a utilitarian at heart, I cannot help but disagree with such absurdity. Any utilitarian knows that predicting the outcome of an action is a computationally intractable problem. I mean that the outcome is, of course, in the utilitarian sense of calculating the net amount of wellbeing it would create. To render such a calculation, one would need to let the universe run its course following a particular action, replace that action with another and let the universe rerun its course. This process would get repeated until the set of all possible actions which can be done in context C at time T have been exhausted in the evaluation of their net wellbeing. Stating this out loud should allow one to understand how absurd it is to state that a vision, a complex structure, will lead us to the highest possible aesthetic fulfillment when we cannot even deduce the impact of a single action. Given that the epistemic limits on utilitarian doctrine also apply to the Nietzschean overhuman ideal, it is improbable that the enormous sacrifice would ever merit the reward. Thus, the Nietzschean Transhumanist Utopia has to be abandoned.

I propose that we head in the opposite direction. Given epistemic constrain, we must acknowledge that we cannot define the ideal goal or what the post-human or overhuman should genuinely look like. To follow in the transhumanist steps, we need to ensure that people are allowed to use technology in order to pursue their enhancement. However, how they decide to do this should be left to them. Within epistemic confines, rapid prototyping with unrestrained communication and documentation is the best way to generate knowledge about what may actually lead to a better future. We will probably discover that no one path is perfect. We may discover that some paths we started that looked promising will lead to dead ends. However, by allowing people to pursue a plurality of ideals, be they mundane, transcendental, or regressive ideals, we can maximize the chances of creating the best

version of the post-human. A version that probably no one has or will ever fully develop by themselves.

REFERENCES

Bostrom, Nick (2005). "Transhumanist Values." In Special Conference Supplement, *Journal of Philosophical Research*, *Ethical Issues for the Twenty-First Century*, Frederick Adams (ed.), 3 – 14.

Bostrom, Nick (2001). "What is Transhumanism", available at: <https://nickbostrom.com/old/transhumanism> last time accessed 18.12.2022.

Bostrom, Nick (2004). "Transhumanism: The World's Most Dangerous Idea?", available at: <https://nickbostrom.com/papers/dangerous> last time accessed 18.12.2022.

Bostrom, Nick (2008). "Letter from Utopia", available at: <https://nickbostrom.com/utopia.pdf> last time accessed 18.12.2022.

Beiner, Roland (2018). *Dangerous Minds: Nietzsche, Heidegger, and the Return of the Far Right.* Pennsylvania: University of Pennsylvania Press.

Bull, Malcom (2014). *Anti-Nietzsche*. London: Verso.

"Declaration of the Rights of Man and Citizen" (1793). ed. Frank Maloy Anderson, in *The Constitutions and Other Select Documents Illustrative of the History of France*, available at: <http://www.columbia.edu/~iw6/docs/dec1793.html> last time accessed 18.12.2022.

Domenico, Losurdo (2021). *Nietzsche, the Aristocratic Rebel: Intellectual Biography and Critical Balance-Sheet.* Chicago: Haymarket Books.

Ferrando, Francesca and Braidotti, Rosi (2019). *Philosophical Posthumanism*. London: Bloomsbury Academic.

Fukuyama, Francis (2004). "Transhumanism". *Foreign Policy*, no. 144: 42–43, available at: <https://doi.org/10.2307/4152980> last time accessed 18.12.2022.

György, Lukács (2021). *The Destruction of Reason*. London: Verso.

Hughes, James; Bostrum, Nick and Moreno, D. Jonathan (2016). "Human vs. Posthuman", In *The Hastings Center Report*, Vol. 37, No. 5. Wiley.

intl.startrek.com (2011). "Borg Queen", available at: <https://intl.startrek.com/database_article/borg-queen> last accessed 18.12.2022.

Nietzsche, Friedrich (2001). *Beyond Good and Evil*. Cambridge: Cambridge University Press.

Nietzsche, Friedrich (2001). *Gay Science*. Cambridge: Cambridge University Press.

Nietzsche, Friedrich (2006). *Genealogy of Morals.* Cambridge: Cambridge University Press.

Nietzsche, Friedrich (2006). *Thus Spoke Zarathustra*. Cambridge: Cambridge University Press.

Rosen, Stanley (1989). *The Ancients and Moderns: Rethinking Modernity.* New Haven: Yale University Press.

Russell, Bertrand (2004). *A History of Western Philosophy*. London: Routledge Classics.

Sorgner, L. Stefan (2022). *We have always been cyborgs: Digital Data, Gene Technologies, and an Ethics of Transhumanism*. Bristol: Bristol University Press.

Sorgner , L. Stefan (2020). *On Transhumanism.* The Pennsylvania State University Press.

The President`s Council on Bioethics (March 2008). *Human Dignity and Bioethics: Essays Commissioned by the President's Council on Bioethics*. Washington, D.C., available at: <https://bioethicsarchive.georgetown.edu/pcbe/reports/human_ dignity/> last accessed 18.12.2022

Filmography

Star Trek – Next Generation, Season 2, Episode 16, "Q Who", Directed by Rob Bowman, aired May 6, 1989

Star Trek – Next Generation, Season 3, Episode 26, "The Best of Both Worlds, Part I", Directed by Cliff Bole, aired Jun 16, 1990

Star Trek – Next Generation, Season 4, Episode 1, "The Best of Both Worlds, Part II", Directed by Cliff Bole, aired Jun 16, 1990

Star Trek – Next Generation, Season 5, Episode 23, "I, Borg", Directed by Robert Lederman, aired May 9, 1992

Star Trek – Next Generation, Season 6, Episode 26, "Descent, Part I", Directed by Alexander Singer, aired Jun 21, 1993

Star Trek – Next Generation, Season 7, Episode 1, "Descent Part II", Directed by Alexander Singer, aired Sep 18, 1993

Star Trek – First Contact, Directed by Jonathan Frakes, 1996

Star Trek – Voyager, Season 3, Episode 17, "Unity", Directed by Robert Duncan McNeill, aired Feb 12, 1997

Star Trek – Voyager, Season 3, Episode 26, "Scorpion, Part I", Directed by David Livingston, aired 21, 1997

Star Trek – Voyager, Season 4, Episode 1, "Scorpion, Part II", Directed by Winrich Kolbe, aired Sep 3, 1997

Star Trek – Voyager, Season 4, Episode 6, "The Raven", Directed by LeVar Burton, aired Feb 12, 1997

Star Trek – Voyager, Season 5, Episode 2, "Drone", Directed by Les Landau, aired Oct 21, 1998

Star Trek – Voyager, Season 5, Episode 15, "Dark Frontier, Part I", Directed by Cliff Bole, aired Feb 17, 1999

Star Trek – Voyager, Season 5, Episode 16, "Dark Frontier, Part II", Directed by Terry Windell, aired Feb 17, 1999

Star Trek – Voyager, Season 6, Episode 2, "Survival Instinct", Directed by Terry Windell, aired Sep 29, 1999

Star Trek – Voyager, Season 6, Episode 16, "Collective", Directed by Allison Liddi-Brown, aired Feb 16, 2000

Star Trek – Voyager, Season 6, Episode 19, "Child's Play", Directed by Michael Vejar, aired Mar 8, 2000

Star Trek – Voyager, Season 6, Episode 26, "Unimatrix Zero, Part I", Directed by Allan Kroeker, aired May 24, 2000
Star Trek – Voyager, Season 7, Episode 1, "Unimatrix Zero, Part II", Directed by Michael Vejar, aired Oct 4, 2000
Star Trek – Voyager, Season 7, Episode 2, "Imperfection", Directed by David Livingston, aired Oct 11, 2000
Star Trek – Voyager, Season 7, Episode 25, "Endgame", Directed by Allan Kroeker, aired May 23, 2001

TWISTING TRANSHUMANISM: SOME SCOPE AND LIMITS OF SORGNER'S PHILOSOPHY

NICOLÁS ANTONIO ROJAS CORTÉS[1]

Abstract

Acknowledging an ontology of permanent becoming has ethical and epistemological consequences for the presentation of a thought system. Ontological presuppositions have consequences for our understanding and our ways of inhabiting the world. These statements are particularly true when we think of transhumanism. The problem is that when these assumptions are not recognized, the proposals of transhumanism are understood as contradictory. In this sense, we analyze Stefan Lorenz Sorgner's proposal to "twist" transhumanism in order to understand the difference between classical transhumanism and weak Nietzschean transhumanism or metahumanism. We will show that classical transhumanism, based on enlightened presuppositions, has ethical consequences that are too culturally specific to be shared beyond the global north. We will then point out how the ontology of permanent becoming, through ethical and alethic nihilism, makes it possible to distance oneself from classical transhumanism and open oneself to further dialogue. However, we will conclude by pointing out that Sorgner's concept of negative freedom is not sufficient to sustain a democratic transhumanism beyond Europe. It will be necessary to expand on positive freedom, for the question remains: what can a cyborg, who does not have the minimum material conditions for self-enhancement, do about it?

Keywords*: Transhumanism, Metahumanism, Negative Freedom, Twisting, War for the Values, Cyborg*

---

[1] PhD Candidate in Philosophy at Universidad de Chile. Member of INTEMPESTIVA, Chilean Network of Nietzsche Studies & ALECT, Asociación Latinoamericana de Estudios Críticos sobre Transhumanismo. Research Associate of the Centro de Estudios de Ética Aplicada (CEDEA). Email: <nicolas.rojas.c@ug.uchile.cl>.

1. Introduction

Stefan L. Sorgner's original work will be, undeniably – if it is not already–, a minimum and fundamental reference point for the historiographical study concerning transhumanist philosophy. We say "transhumanist philosophy" for a reason that should be emphasized: not every transhumanist expression is versed in philosophy, even if its exponents have a background in the discipline[2]. In this sense, Sorgner's works may abandon certain discussions of first-generation transhumanism that make the movement seem like a utopian illusion exclusive to the first world. Or at least, we consider that as one of the aims of the metahuman[3] philosopher's work, namely, expressing the proposals of transhumanism in a philosophically coherent way.

In a previous article I have argued that in the Spanish-speaking world the discussion of Sorgner's thought has been relegated to the dispute over whether Nietzsche is actually a precursor of transhumanism (Rojas 2022b). Apart from the fact that Sils-Maria's philosophical influence may be or not effective in transhumanism, it is certain he is an influence on the metahuman's thinking development. I consider that this discussion does not contend with the metahuman thought development and overlooks the fact that the author's original proposals, while influenced by Nietzsche, discuss quite interesting contemporary

[2] In a previous article I have sketched a preliminary difference between different types of transhumanism. This difference consists mainly in the ability to respond to criticisms in a philosophically coherent way or to articulate ideas in a philosophically coherent way. In both cases, thinkers like Kurzweil, Bostrom, More or others – who can be easily refuted – do not meet these requirements, so we call them first-generation transhumanists. Thinkers such as James Hughes or Stefan L. Sorgner explain their thinking in a much more argumentative and philosophically developed way than the previous ones. We call these second-generation transhumanists (Rojas 2022a). One basis for this difference lies in the exposition of contents presented in the Posthuman Reader edited by Jan Santiesko and Evi D. Sampanikou. In my review of this text, I insist on the radical difference between transhumanist thought that is not versed in philosophy and that which can coherently argue its proposals (Rojas 2021).

[3] For his effort to establish metahumanism as a serious line of academic discussion, I will refer to Sorgner as metahuman so as not to repeat his name as often.

topics, such as, the democratic use of collected data, the paternalism of bio-enhancement, among others.

This is a significant opportunity to discuss one of the topics that, in our opinion, is one of the most interesting in Sorgner's thinking development: the idea of twisting transhumanism and its consequences for the exposition of this thought. Although the proposal to "twist transhumanism" is original to *Philosophy of Posthuman Art* (2022b), I believe that both *On Transhumanism* (2021b) and *We have always been Cyborgs* (2022c) already contain outlines for understanding this idea. In this sense, we can affirm that Sorgner's metahumanism[4], in its effort to represent a weakened trans- and post-humanism, is also one of the possible consequences of this "twisting". However, it is this strategy that also enables him to position himself beyond naïve transhumanism and to have reasons to distance himself from it.

However, despite their originality, we can identify a limit to Sorgner's original proposals: particular social and historical contexts. In this case, this refers to the possible application of his proposals in Latin America. In particular, the idea of negative freedom as a more valuable achievement than privacy is a proposition that can be questioned. The key point of this critique lies in the radical economic and social differences that exist in Latin America.

2. War for The Values: Against Transhumanism as Humanism

On Transhumanism begins by stating that the "Transhumanists simply share one basic view, which they continually adapt to the latest state of philosophical insight, scientific research, and technological capacity" (2021b 1). However, it is highly complex to determine what is meant by "latest state of philosophical insight, scientific research, and technological capacity" here. Also, and I am not sure that all transhumanists share this statement.

[4] In what follows we will use metahumanism and weak Nietzschean transhumanism as synonyms.

To illustrate my statement, it seems relevant to mention some examples of classical transhumanism. First of all, we can recall the first principle offered by the founder of the Extropy Institute, Max More, in his *The Principles of Extropy*, namely "Perpetual Progress":

> "Extropy means seeking more intelligence, wisdom, and effectiveness, an open-ended lifespan, and the removal of political, cultural, biological, and psychological limits to continuing development. Perpetually overcoming constraints on our progress and possibilities as individuals, as organizations, and as a species. Growing in healthy directions without bound." (More, 2003)

While in the unfolded version, it reads:

> "Continual improvement will involve economic growth. We can continue to find resources to enable growth, and we can combine mindful growth with environmental quality. This means affirming a rational, non-coercive environmentalism aimed at sustaining and enhancing the conditions for flourishing. Individuals enjoying vastly extended life spans and greater wealth will be better positioned to intelligently manage resources and environment. An effective economic system encourages conservation, substitution, and innovation, preventing any need for a brake on growth and progress. Migration into space will immensely enlarge the energy and resources accessible to civilization. Extended life spans may foster wisdom and foresight, while restraining recklessness and profligacy. We can pursue continued individual and social improvement carefully and intelligently." (More, 2003)

Likewise, in article *Why I Want to be a Posthuman When I Grow Up* (2008), the famous Nick Bostrom tries to explain by means of an analogy why it could be very good for us to become posthuman:

> "A life might be good or bad because of its causal consequences for other people, or for the contribution it makes to the overall value of a society or a world. But here I shall focus on the value that a life has for the person whose life it is: how good (or bad) it is for the subject to have this life. The term 'well-being' is often used in this sense.
> When I speak of the value of a life here, I do not refer to the moral status of the person whose life it is. It is a separate question what the moral status would be of human and posthuman beings. We can assume for present purposes that human and posthuman persons would have the same moral status. The value of a life refers, rather, to how well a life goes for its subject. Different human lives go differently well, and in this sense their lives have different values. The life of a person who dies from a painful illness at age 15 after having lived in extreme poverty and social isolation is typically worse and has less value than that of

> a person who has an 80-yearlong life full of joy, creativity, worthwhile achievements, friendships, and love." (Bostrom 2008, 10).

These statements, rather than representing the latest state of some form of knowledge – philosophical, scientific, technological – are statements worthy of utopian futurists poorly versed in rational argumentation (we suppose that is why they want to improve their reason). Sorgner would not be against our statement as he is the one positioning himself beyond what Bostrom naively put forward in his *Letter from Utopia* (2008a) and defends himself against Hauskeller's (2012) criticisms in his article *Transhumanism and 'The Land of Cockaygne* (2018; 2022c). If we admit with the metahumanist a naturalist worldview then removing political, biological, cultural and psychological boundaries in order to achieve a good life, as More and Bostrom propose, cannot be realized in a literal sense. Unlimited health (not to mention immortality) is not a realistic option where we understand ourselves from a naturalistic perspective. It is not only the end of our existence but also the end of the universe that should be taken into consideration when declaring such an affirmation; since, how could it be possible to have unlimited health once the universe has collapsed? (Sorgner 2022c, 164).

In this sense, the proposal to overcome limits by assuming economic growth based on rational conditions for the flourishing of individuals still seems to be a terribly humanist practice, hence both hypercolonialist and extractivist (del Val, 2021). Bostrom's comparison between a poor individual dead by the age of 15 and another one that lives "happily" until the age of 80 does not seem rational. There is no doubt that this is a very Northwest comparison and worse than that, it reduces "the value of life" to hedonistic happiness. It is necessary to point out that the problem is not, of course, hedonism itself, but rather the desire to justify an existence beyond human existence as a privilege granted only to a certain type of human being. If so, then it is not about having political and cultural limits overcome, but rather progressively maintaining ideas that have allowed the first world to sustain its living standard. This is not surprising for a proposal that presupposes that the market is free and aspires to go beyond ideology (More 2003).

James Hughes has explained precisely what the matrix underlies proposals such as those mentioned above, namely the enlightened and

too humanist roots of transhumanism (2010). In this sense, it is key to consider Bostrom's gesture in his article *A History of Transhumanist Thought* (2005) of dismissing Nietzsche as a precursor of transhumanism. The conjunction between confidence in reason and confidence in scientific progress is the reason why projects such as those of Ray Kurzweil (2005) or Elon Musk are considered transhumanist. But in a humanist sense, *i.e.*, as a paternalistic (Sorgner 2021b, 22) way of imposing conceptions of subjectivity (Sloterdijk 1999; Coeckelbergh 2013). It would not be problematic to identify these projects alongside those from a first-generation transhumanism (Rojas 2022a), a movement from the global North that does not consider the specific cultural differences of other localities. Moreover, the enlightened roots generate a profound contradiction when it comes to promoting human enhancement, since this enhancement would be possible only for the few, as Bostrom rightly says, for those who manage to live a healthy and joyful 80-year life, but not for the many whose bodies-territories[5] suffer from capitalist extractivism. That is because the material for technological progress is not exclusively available in North America or Europe.

In this sense, what is at stake is a War for Values. We can define this notion as the dispute regarding the different axiological consequences of the intellectual presuppositions from cultural or philosophical movements. It is known that the metahuman has an academic war with New Realism and the Frankfurt School, as well as with Humanism in general (Sorgner 2017; 2022b). However, if values are the ones in dispute, then it is also possible to argue that Sorgner's proposal outlined in *We have always been Cyborg* (2022c) represents an effort to move away from this all too humanistic transhumanism.

If someone wants to be a posthuman when he or she grows up, and his or her reasons are confidence in reason and technological progress that will allow him or her to live a happily hedonistic life, then we are talking with a very American perspective. Thus, the argument of beneficence (Bostrom 2008) is not democratically conceivable[6]. Finally,

[5] To learn more about this concept, its relation to the academic feminism developed in Latin America, technology and indigenous spiritualities, see Balcarce 2021, 446–458; Torrano & Fischetti 2018, 267-279.

[6] See the arguments developed by Vaccari 2019, 192–219.

this ideology remains paternalistic in relation to other cultures. Unfortunately, from the global south, we know that "we" are not considered among the "we" who can voluntarily accede to a posthuman existence. So, the question is: how do you fight against the overly humanistic transhumanist assessments? The answer we get from Sorgner's efforts is the following: with other presuppositions.

3. Twisting Transhumanism: Coining a New Concept

If the presuppositions are different, so are the consequences. Having this in mind, Sorgner's strategy is relatively straightforward: it is to offer a philosophical description of transhumanism. While Bostrom dismissed Nietzsche's influence on transhumanism (2005), the metahumanist began a long discussion about the positive influence of the philosopher from Sils-Maria on this movement. I am not interested in discussing here whether this influence is effective or not, as much ink has been spilled on this topic already (Tuncel 2017). What is certain is that Sorgner's most original thought has indeed been influenced by Nietzsche's philosophy (Rojas 2022b), and in this case, Nietzsche's influence cannot be denied for the weak Nietzschean transhumanismus (*schwachen Nietzscheanischen Transhumanismus*), or metahumanism, proposed by Sorgner.

Thus, admitting Nietzsche's influence on transhumanism implies embracing Heraclitus and abandoning — at least in principle — the Enlightened rationalism that structures the transhumanism we described in the previous section. Here we encounter an ontological view that understands the world as "continual becoming" (Sorgner 2022c, 12; 138). The immediate consequence of this, following in our view an updated reading of *Über Wahrheit und Lüge im aussermoralischen Sinne* (1873), is alethic nihilism and ethical nihilism (Sorgner 2022c, 17 and following). Regarding the former it is affirmed that "Nothing positive can be said which corresponds with the world with certainty" (2022c, 18), while for the latter it is stated: "that no non-formal judgement of the good is plausible for all people" (2022c, 19). From our perspective, these consequences can also be understood as useful strategies for positioning another kind of transhumanism.

The question "why would it be a good idea to be a posthuman?" could not be answered by addressing Bostrom's comparison between the Third Worldist short life and the First Worldist long life, because such a perspective is violent towards the multiple conceptions of good that other people may have. A short, intense and healthy but tragic life might be even more valuable to someone than a long and comfortable one. In this sense, More's fantasy of removing the political limits of our condition would constitute a nonsense. If nothing can be established as a single, universal truth, politics is not exempt from that either. A first-world transhumanist utopia, is just one possible perspective in the war for values.

Furthermore, Sorgner's naturalistic perspective would not allow him to blindly trust proposals such as mind uploading or cryogenics based on faith in perpetual scientific progress[7]. In fact, if we admit that

> "A naturalist account of the word implies that all entities can in principle be accessed empirically. This does not mean that all entities can already by [sic] investigated empirically, but it implies at least that in principle an empirical analysis of all entities is possible" (2022b, 50-51)

Then, the relationship between transhumanism and the goal of extending life should change in order to follow a coherent argument. It would not be acceptable to keep faith with enhancement projects that aim to spend material resources and productive efforts on achieving the singularity in its mind-machine fusion version or the cryogenisation of the richest. That is not because we believe it impossible to achieve, but because its success depends on a series of empirical investigations that have not yet occurred. Rather than contributing to a state of cryogenic "half-life" suspension as the one accessed by Glen Runciter in *Ubik* (1969), it would be more coherent to focus efforts on pressing contemporary problems — why freeze people when we can save glaciers? From an empirical analysis it is possible to explain the death of human entities. However, the desire for the overcoming of the limits of the human condition does not seem explicable. Such a goal is only a possible perspective on the question of what is good.

---

[7] Sorgner should accept, alongside the ontology of permanent becoming, Nietzsche's critique in the third essay of *On the Genealogy of Morality*. See Nietzsche 2007, 109-110.

It is also important to note that the question of why it might be a good idea to be a posthuman loses meaning from an ontology of permanent becoming. Because *We have always been cyborgs* (2022c). This statement is important insofar as it represents a break from rationalist and enlightened transhumanist assumptions. Ultimately, what underlies this corresponds to posthumanist premises from which we can understand ourselves,

> "Cyborgs are therefore controlled organism. Control already happens with us becoming human. In philosophy, human beings were usually defined by their ability to speak. Learning language is our first upgrade, which our parents provide us with. Our cyborgization continues with the acquisition of new skills, such as learning mathematics, history, and so on. However, a new dynamic is currently emerging. […]. With the integration of digital technologies intro our humanity, new possibilities as well as serious challenges arise." (Sorgner 2022c, 32).

Sorgner's definition is consistent with the denial of an essential, fixed and immutable identity justified by the appeal to a conception of human nature (Fukuyama 2002; Sandel 2009; Habermas 2003) and also aware that the understandings we offer of ourselves depend on discourses of power that steer the human herd, to use Sloterdijk's expression in his *Regeln für den Menschenpark* (1999). Transhumanism, then, is also inserted in the disputes over the definition of the entity that we are. However, Sorgner's gesture prevents this dispute from continuing in essentialist terms. If there is no natural essence, neither is it possible to predicate a difference between technology and nature. Undergoing enhancement to abandon the human condition makes no sense if we have never been human. To exist as cyborgs implies an existence beyond the human if we admit that technology plays a crucial role in our self-understanding. This is why Sorgner speaks of "cyborization", as the entity that we are must be subject to changes that also include the agency of technology. In this sense, despite technology not being new to our self-understanding, we do find ourselves in an age where technological agency has global and cellular impact — the possibilities for political manipulation or medical research through the production of digital data have never been more real.

More specific consequences of embracing an ontology of permanent becoming could be discussed, but in our view the most important is the weakening of transhumanism (*indebolimento*) (2021b, 53; 2022c, 138). Briefly, this weakening is referred to the fundamental premises of trans- and post-humanism, admitting a degree of confidence in technology like the transhumanists and discarding the claim of a special existence for the human with the posthumanists. The outcome of this operation would result in the original metahumanism of Sorgner and del Val (2021). However, we consider that this operation responds to a philosophical strategy on Sorgner's part, since a transhumanism such as we discussed in the previous section can be easily refuted[8]. The strategy, as mentioned in the introduction, is the "twist",

> "The concept of the twist corresponds to the German notion of *Verwindung*. *Verwindung* (twist) differs from *Überwindung* (overcoming) in so far as it is not a leaving behind of something. Overcoming leaves behind and separates itself categorically from the past, whereas a twist develops the past further in an inclusive manner" (Sorgner 2022b, 120-121).

The strategy is particularly interesting and, in fact, could be reformulated in the following statement: transhumanism needs a coherent philosophical foundation, and this means re-interpreting their presuppositions. An example of this "twisting" is what happens with the notions of soul and body,

> "The immaterial mind and the material body get woven into a psychophysiological unity. What used to be the divine spark in us gets interpreted as a technological steering of an organism. We turn into cyborgs, i.e., cybernetic organism." (Sorgner 2022c, 21).

This strategy is particularly useful, for example, when they are criticized as dualists[9] and hegemonic hyperhumanists (del Val 2021). In Sorgner's words, the problem is that classical transhumanism, philosophically misinformed, represents a form of violent and morally problematic paternalism (Sorgner 2022c, 139). However, the solution

---

[8] See, for example, the arguments developed by Gabriel 2015. Also, Rojas 2022, a5-27. In particular the pages 9–14.

[9] However, it is necessary to note that the description of "psychophysiological unity" does not make the examination of who we are any simpler either. Rather, it moves it into a realm of Nietzschean or psychoanalytic analysis that must also be discussed.

could not be the eradication of human beings from existence to solve the problems we have created for the earth and the other beings that inhabit it. That is why even posthumanist influence must be weakened (2022c, 185) to respect and safeguard the concept of negative freedom that was achieved with modernity. In this sense, neo-Luddism would be as paternalistic an alternative as extreme technophilia. Consequently, we speak of conceiving new concepts since "twisting" is a methodological strategy that serves to explain the difference between metahumanism or weak Nietzschean transhumanism and classical transhumanism.

Therefore, "twisting" functions as a strategy or methodology that attempts to base on coherent presuppositions the ideas that are worth rescuing from a movement such as the transhumanist one and discarding them in order to respect both ethical nihilism and the achievement of negative freedom. However, we could say that even by twisting transhumanism and generating a weak version of it, there are still clear limits to its general acceptance.

4. A Limit of Sorgner's Transhumanism: Beyond the Global North

One of the most controversial theses in *We have always been Cyborgs* is the distinction between freedom and privacy. This distinction takes place in the context of the argument for a democratic use of data (Sorgner 2022c, 30-60). Elsewhere I have called this distinction "Sorgner's thesis"[10] and my understanding of it is based on the following statement:

> "We have no reason to regret giving up on privacy. We do not cherish privacy, but we cherish freedom. We can have health and freedom, but not health and privacy. Why do we think that we need privacy?" (Sorgner 2022c, 54)

Sorgner's thesis conjugates at least two premises: (a) privacy and freedom are not equivalent and (b) they do not need each other, *i.e.*, we can be free without privacy. We agree with the first premise, since

[10] I have developed part of this argument in an article called "Could Cyborgs' Freedom Enhance Philosophically Transhumanism? A response to Sorgner's *We Have Always Been Cyborgs*" which is expected to be published in the coming months.

freedom is said in many ways and, of course, depending on the meaning, we can affirm non-equivalence. Obviously, this depends on the ontological presuppositions at stake. From a mechanistic ontology we could not speak of freedom, whereas from an existentialist perspective freedom would be the foundation of our actions. In the case of Sorgner's metahumanism one could — and we do so explicitly — ask the philosopher to clarify what he means by freedom and in what sense.

At least if we trace this idea back to *On Transhumanism* (2021b) and *We have always been Cyborgs* (2022c) we only find freedom understood as the negative freedom of democratic and liberal systems. In this sense, the most precise definition we found is: "Freedom includes that one must not limit the freedom of others and that one must not harm other people" (2021b, 46). The definition is further expanded to include "that people have the right to morphological freedom, including the right to decide for themselves whether or not they want to be modified by genetic engineering measures (2021b, 67). Finally, the reason for affirming this kind of freedom is that this "enables each of us to conceive our own lives according to our own ideas of the good" (Sorgner 2021b, 83).

The problem with this understanding of freedom is that, in our view, it remains culturally specific and presupposes living in a society where, at least, democracy functions, political structures allow for morphological freedom and, above all, minimum conditions for sustaining and maintaining a healthy life. Are we not talking here about an almost utopian society? It is true that Sorgner refers to the data usage that Europe should employ (Sorgner 2022c, 30-46). However, the minimum conditions do not exist everywhere in the world to reduce freedom only to its negative form.

Sorgner's proposal for Europe seems interesting. A democratic use of data in terms of sustaining a health system with data is an idea that could be realized. In the global south, however, not even the minimum conditions of access to basic resources exist for negative freedom to be a sufficient condition for a democratic existence.

It is widely known that access to safe drinking water is not for everyone. In fact, in 2019 there was also insufficient data to measure drinking water coverage in schools in either Latin America and the Caribbean or Eastern and South-Eastern Asia. It seems to us that access

to water is a necessary, but not sufficient, condition to start talking about the possibility of health insurance paid for with our data.

Furthermore, working with data in a democratic way implies access to quality education. From Chile it can be stated that we do not have equitable[11] access to education (World Bank Group 2021, 4-5). How would it be possible to develop systems to manage this data? Possibly the option would be to buy it from companies in the global north or to rely on private production. It is important to note that at the beginning of the covid-19 pandemic, education depended on internet access. However, internet access is not for everyone in Latin America either. Even worse, according to UNICEF: "2.2 billion – or 2 in 3 children and young people aged 25 years or less – do not have internet access at home" (2020).

While one could argue with Sorgner that human rights are a case of a useful fictional ethic that reduces violence (2022c, 169), unfortunately we cannot claim that human rights are secured in Latin America either. The effects of the dictatorships promoted by the US on this continent are still being felt. Worse still, deniers of human rights violations in the dictatorships such as José Antonio Kast or Jair Bolsonaro have managed to gain significant political visibility and have even tried — and succeeded in the case of the latter — to gain high political positions. Authoritarianism, xenophobia, machismo, denial of the climate crisis and, in general, conservative extreme right-wing ideas are elements that constitute the discourses of these characters. If we add to this their good relations with extractivist businessmen — as was the case of former president Sebastián Piñera, remembered for his government's disastrous organisation of COP25 — then this part of the world seems more like a victim of neoliberal colonialism than a place where people can live freely.

With such decadent social and political structures, one might ask: what place could transhumanist ideas have in this part of the world?

[11] A year in the PhD programme in Philosophy at the University of Chile, currently ranked 401-500 in the Shanghai Ranking, currently costs about 4 million Chilean pesos (about 4,000 USD). The minimum salary in 2022 is 365,000 Chilean pesos (less than 400 dollars approximately). It is reasonable that a large part of society is dissatisfied with access to education. Scholarships are not enough to live on.

We see two options: The first is simple and involves sticking to classical transhumanism, namely that transhumanist enhancement projects would only be accessible to the few, privately and at the expense of the many. Extractivists, corrupt politicians, drug traffickers and heirs of dictatorships could benefit from this. The rest would continue to beg for the minimum. The second option is more complex and involves thinking about freedom in its positive version: how could we improve our lives through technological means if there are value, economic, social, and political structures that prevent it? Our answer, for now only an intuition to be developed, is: through social revolution.

The problem with discussing only negative freedom in this context is simple: we are still slaves. It doesn't matter if we are ontologically understood as cyborgs, because we don't have the minimum conditions to subsist or to have access to improvements such as the universal health system paid for by our data. We are a place of matter extraction. But if we have affirmed an ontology of permanent becoming, this should be able to change, shouldn't it? The question is: What can I do with my freedom understood as autonomy and as self-rule? (Coeckelbergh 2021, 59). In other words, what can I do with my positive freedom?

We offer a preliminary intuition for the development of the answer to the question. There are at least three desires at stake in this sketch of an answer: (1) if we want morphological freedom, (2) if we want to live according to our own convictions of the good, and (3) if we want to overcome humanism, then we must be agents of change. Cyborgs that can destroy extractivist colonialism, neoliberal structures and the moral conservatism that still plagues us. But this will not be possible without violence and social revolution. Because nobody wants to lose their privileges.

5. Conclusion

There is no doubt that one of Sorgner's most original contributions to transhumanism is to twist it. Twisting it involves examining its presuppositions, redirecting, and modifying some of them. The achievement of modern negative freedom is a basic element for

sustaining transhumanism beyond the global north, but it is not a sufficient condition for transhumanism to be globally accepted.

It is true that Sorgner is already working on the impact of transhumanism in Africa (2022a).[12] However, as we have already pointed out, in the Spanish academic discussion, we are still very reluctant to reflect on how to effectively implement the proposals of transhumanism in our social and historical context. In our view, the reason is quite simple: material conditions here are not equivalent to those in the global north. If we follow Sorgner and still claim that "one's own freedom ends where the freedom of others begins" (2022c, 18), then the question remains the same: What can a cyborg, who does not have the minimum material conditions for self-enhancement, do about it? Our intuition is that Sorgner's transhumanism needs to think about positive freedom.

I would rephrase the statement that a large majority wants a healthy life. We would propose that a large majority wants the possibility of access to the minimum material conditions that make it possible to conceive a healthy life. We keep on twisting transhumanism.

Acknowledgement:

Agencia Nacional de Investigación y Desarrollo / Subdirección de Capital Humano / Beca de Doctorado Nacional 21210804.

REFERENCES

Blackford, Russell (2017). "Nietzsche, the *Übermensch*, and Transhumanism: Philosophical Reflections." In *Nietzsche and Transhumanism: Precursor or Enemy?*, Yunus Tuncel (ed.). UK: Cambridge Scholars Publishing, 191–204.

Balcarce, Gabriela (2021). "Políticas de bom viver: Olhares sobre o terricídio e espiritualidades não modernas." *Abatirá – Revista De Ciências Humanas E Linguagens*, 2(4): 446–458.

[12] In fact, the metahuman along with thinkers such as Agbolade Omowole, Chogwu Abdul, Leo Igwe, among others, held the symposium "Transhumanism and Africa" in February 2022.

Bostrom, Nick (2005). "A history of transhumanist thought." In *Journal of Evolution and Technology*, 14(1): 1–25.

Bostrom, Nick (2008a). "Letter from Utopia." In *Studies in Ethics, Law, and Technology*, 2(1): 1–7.

Bostrom, Nick (2005). "The fable of the dragon tyrant." In *Journal of Medical Ethics*, 31(5): 273–277.

Bostrom, Nick (2014). *Superintelligence: Paths, Dangers, Strategies*. UK: Oxford University Press.

Bostrom, Nick (2008). "Why I Want to be a Posthuman when I Grow Up." In *Medical Enhancement and Posthumanity*, Bert Gordijn & Ruth Chadwick (eds.), 107-136. Springer.

Coeckelbergh, Mark (2021). *Green Leviathan or the Poetics of Political Liberty Navigating Freedom in the Age of Climate Change and Artificial Intelligence*. New York: Routledge.

Coeckelbergh, Mark (2013). "The Transhumanist Challenge." In *Human Being @ Risk. Enhancement, Technology, and the Evaluation of Vulnerability Transformations*. Philosophy of Engineering and Technology, Vol 12. Dordrecht: Springer, 19–36.

Dick, Phillip K. (1969). *Ubik*. USA: Doubleday.

Fukuyama, Francis (2002). *Our Posthuman Future. Consequences of the Biotechnology Revolution*. New York: Farrar, Straus and Giroux.

Gabriel, Markus (2015). *Ich ist nicht Gehirn: Philosophie des Geistes für das 21. Jahrhundert*. Berlin: Ullstein Buchverlage.

Gabriel, Markus (2020). *Moralischer Fortschritt in dunklen Zeiten : universale Werte für das 21. Jahrhundert*. Berlin: Ullstein Buchverlage.

Habermas, Jürgen (2003). *The Future of Human Nature*. UK: Polity Press.

Hauskeller, Michael (2012). "Reinventing Cockaigne. Utopian Themes in Transhumanist Thought." In *Hastings Center Report,* 42(2): 39–47.

Hughes, James (2010). "Contradictions from the enlightenment roots of transhumanism." In *The Journal of medicine and philosophy*, 35(6): 622–640.

Kurzweil, Ray (2005). *The Singularity is Near*. USA: Viking.

More, Max (2013). "The Philosophy of Transhumanism." In *The Transhumanist Reader*, M. More & N. Vita-More (eds.). UK: Wiley, 3–17.

More, Max (2003). "Principles of Extropy." *Extropy Institute*. Available at <https://web.archive.org/web/20131015142449/http://extropy.org/principles.htm> last time accessed inNovember 2, 2022.

Nietzsche, Friedrich (2007). *On the Genealogy of Morality*. UK: Cambridge University Press.

Reality Check (2019). "Chile protests: Is inequality becoming worse?." *BBC News*, 21 October.

Rojas Cortés, Nicolás (2022a). "Contra Bostrom y a favor de un transhumanismo de segunda generación." In *Revista OtroSiglo*, 6(1): 5–27.

Rojas Cortés, Nicolás (2022b). "El Metahumanismo de Stefan Lorenz Sorgner." In *El Banquete de los Dioses*, (10): 75–106.

Rojas Cortés, Nicolás (2021). "Thinking Ourselves Beyond Humanism." In *Journal of Posthuman Studies*, 5(1): 78–86.

Sampanikou, Evi D. & Jan Stasieńko (eds.). (2021). *Posthuman Studies Reader. Core readings on Transhumanism, Posthumanism and Metahumanism*. Schwabe Verlag.

Sandel, M. (2009). *The Case against Perfection: Ethics in the Age of Genetic Engineering*. USA: Belknap Press of Harvard University Press.

Sandberg, Anders & Nick Bostrom (2008). "Whole Brain Emulation: A Roadmap." In *Technical Report #2008-3*. Oxford University: Future of Humanity Institute.

Shulman, Carl & Nick Bostrom (2021). "Sharing the World with Digital Minds." In *Rethinking Moral Status*, S. Clarke, H. Zohny, & J. Savulescu (eds.). Oxford; New York: Oxford University Press.

Sloterdijk, Peter (1999). *Regeln für den Menschenpark. Ein Antwortschreiben zu Heideggers Brief über den Humanismus*. Frankfurt am Main: Suhrkamp.

Sloterdijk, Peter (2017). "Rules for the Human Park." In *Not Saved. Essays after Heidegger*. Malden, MA: Polity Press, 193–216.

Sorgner, Stefan Lorenz (2022b). *We Have Always Been Cyborgs – Digital Data, Gene Technologies, and an Ethics of Transhumanism*. Bristol, UK: Bristol University Press.

Sorgner, Stefan Lorenz (2021a). "Adage of Metahumanism (2020)." In *Posthuman Studies Reader: Core readings on Transhumanism, Posthumanism and Metahumanism,* Evi D. Sampanikou / Jan Stasieńko (eds.). Basel, Schweiz: Schwabe Verlag, Schwabe Verlagsgruppe AG, 305–308.

Sorgner, Stefan Lorenz (2022a). "Africa, China, and the Meaning of Digital Data." In *Revista Ethika+*, (5): 139–147.

Sorgner, Stefan Lorenz (2021b). *On Transhumanism*. University Park, Pennsylvania: The Pennsylvania State University Press.

Sorgner, Stefan Lorenz (2017). "Philosophy as "Intellectual War of Values." In *Philosophy's Future*, R. Blackford & D. Broderick (eds.). Hoboken/NY: Wiley Blackwell, 193–200.

Sorgner, Stefan Lorenz (2022a). *Philosophy of Posthuman Art*. Basel, Schweiz: Schwabe Verlag.

Sorgner, Stefan Lorenz (2016). *Transhumanismus – "die gefährlichste Idee der Welt!?W"*. In *Die Gefahrlichste Idee Der Welt!?* Freiburg im Breisgau: Herder Verlag.

Sorgner, Stefan Lorenz (2018). "Transhumanism and the Land of Cockaygne." In *탈경계인문학 TRANS-HUMANITIES*, 11(1): 165–188.

Sorgner, Stefan Lorenz (2017). "Zarathustra 2.0 and Beyond: Further Remarks on the Complex Relationship between Nietzsche and Transhumanism." In *Nietzsche and Transhumanism: Precursor or Enemy?*, Y. Tuncel (ed.). UK: Cambridge Scholars Publishing, 133–169.

Torrano, Andrea & Fischetti, Natalia (2018). "Apuestas del feminismo: Ciencia/Técnica/Latinoamérica. Nuevas urdimbres desde el Sur." In *RevIISE – Revista De Ciencias Sociales Y Humanas*, 11(11): 267-279.

Treder, Mike (2004). "Emancipation from Death." In *The Scientific Conquest of Death. Essays on Infinite Lifespans*, 187–196. Buenos Aires: LibrosenRed.

UNICEF (2020). *How many children and young people have internet access at home? Estimating digital connectivity during the COVID-19 pandemic.* Available at <https://data.unicef.org/resources/children-and-young-people-internet-access-at-home-during-covid19/≥ last time accessed in December, 2022.

Vaccari, Andres (2019). "Why Should We Become Posthuman? The Beneficence Argument Questioned." In *Journal of Medicine and Philosophy*, 44(2): 192–219. doi:10.1093/jmp/jhy041.

del Val, Jaime (2021). "The Dances of Becoming and the Metahumanist Manifesto. Its genealogy, evolution and relevance 10 years after (2020)." In *Posthuman Studies Reader Core readings on Transhumanism, Posthumanism and Metahumanism*, E. D. Sampanikou & J. Stasieńko (eds.). Basel, Schweiz: Schwabe Verlag, Schwabe Verlagsgruppe AG, 299–303.

del Val, Jaime & Sorgner, Stefan Lorenz (2021). "A Metahumanist Manifesto (2010)." In *Posthuman Studies Reader Core readings on Transhumanism, Posthumanism and Metahumanism*, E. D. Sampanikou & J. Stasieńko (eds.). Basel, Schweiz: Schwabe Verlag, Schwabe Verlagsgruppe AG, 295–298.

World Bank Group (2021). *Pieces for Development: How to solve the puzzle for a more inclusive and sustainable Chile?* Washington, DC: World Bank. Available at <https://openknowledge.worldbank.org/handle/10986/36466>, last time accessed in December 20, 2022.

PALADINS WITH MUSCLE CRAMPS: IMMERSION, MMORPG'S, AND NOZICK'S EXPERIENCE MACHINE

RADU USZKAI[1], EMANUEL SOCACIU[2], MIHAI-VALENTIN CERNEA[3]

Abstract

When he first formulated the famous 'Experience Machine' thought experiment in 1974, Robert Nozick could only give his readers one analogy with a potential real-life 'device': psychoactive drugs. Now, decades later, MMORPG's like *World of Warcraft* and EVE Online, with their promise of delivering immersive pleasure to their users, have been dubbed as experience machines. However, some have argued that MMORPG's cannot be what Nozick had in mind when he developed his thought experiment for two reasons: because of the interaction between users mediated by avatars and because they lack sensory-immersion. We discuss how both arguments fail to address the necessary conditions for a truly Nozickian experience machine: (i) guaranteed immunity from any stimuli generated in the circumstances of the "real", physical world and (ii) lack of awareness, while being plugged in an experience machine, of the actual source of an experience. Even though MMORPG's cannot be experience machines in the sense of the original thought experiments, they might still provide an opportunity to address the original concerns that Nozick had regarding hedonism, or other, more recent (but closely related) as escapism through technology.

Keywords*: Robert Nozick, Experience Machine, MMORPG's, immersion, digital games, escapism.*

---

[1] Assistant Professor, Department of Philosophy and Social Sciences, Bucharest University of Economic Studies, Research Institute of the University of Bucharest (ICUB), University of Bucharest. Email: <radu.uszkai@cadi.ro>.
Radu Uszkai's contribution has benefited from the support of the Romanian Young Academy, which is funded by Stiftung Mercator and the Alexander von Humboldt Foundation for the period 2020-2022.
[2] Associate Professor, Faculty of Philosophy, University of Bucharest. Email: <emanuel.socaciu@filosofie.unibuc.ro>.
[3] Assistant Professor, Department of Philosophy and Social Sciences, Bucharest University of Economic Studies. Email: <mihai.cernea@man.ase.ro >

1. When Nozick Meets Blizzard

Have you ever been so caught up in a novel you were reading while commuting in a subway train that you realized in shock to have reached the end of the line, way past your destination point? Did you ever "warn" the character in a movie that she was unknowingly heading into an ambush? Have you ever experienced playing a video game with such an intensity that the line between the game world and the real world became blurred? Maybe none of these ever happened to you. But if you think such occurrences are plausible for some (other) people, then there is room for at least the hint of a worry.

What if the technology were available to make persistent this sort of experience? What if, at the limit, one could live one's entire life in such a state, and it would all be about playing and having fun, shielded from any unpleasantness? What if, to raise the bar even higher, one wouldn't be alone in this, but interacting with others, in something very similar to a society?

We aim here to look at one particular present-day technology that might deliver on the promise. Our question, in more technical terms, is: can MMORPG's be experience machines of the sort described by Robert Nozick in his famous thought experiment[4]? Of course, video games have already been around for a long time, becoming part of our culture and, for many individuals, part of their lives. But many feel that there's something special about MMORPG's that raises difficult challenges. In Taylor's words, "what gives Nozick and contemporary philosophers cause for consideration is that MMORPG's are closer in essence to the experience machine than they are to traditional games. These games offer persistent, immersive (though not sensory-immersive) worlds, and people sign onto them specifically for pleasure." (2007, 71)

We think this is not just an abstract philosophical question. The stakes can be quite high, and they touch both a conceptual and a more practical level. A positive answer could become the source of legitimate

[4] MMORPG stands for "massively multiplayer online role-playing games". They are our main focus here, but the question might be equally valid for other types of "synthetic worlds" (Castronova 2005), such as MUDs (multi-user dungeons) or MMOWs (massive multi-user online worlds).

concerns about online computer games: it would appear that a huge number of individuals are willing or even eager to give up authentic experiences in favor of artificially induced simulacra. Games might turn out to provide the perfect meeting ground between popular expressions of such worries (how many people, if presented with a real choice, would take the blue pill from *The Matrix*?) and the refined accounts of philosophers deploring the on-going "murder of the real". Even if, they contend, the crime is already taking place in our ("real", in a gamer's terms) world, it can become complete only in a virtualized one: "this perfect extermination could only be achieved if the process of virtualization were fully realized." (Baudrillard 2001, 63).

Our thesis here is that, at least when it comes to MMORPG's, such intense worries are not warranted. Although the state of "being plugged" presupposed by the experience machine falls in the same category as the immersion discussed both by scholars and the gaming community, we will argue that it is of a special and extremely demanding type, one that MMORPGs cannot replicate yet, and it seems unlikely they will in the foreseeable future.[5] We shall describe this special kind of immersion and distinguish it from other types. In the last sections of the paper, we shall discuss other attempts to answer the same question and explore some possible implications of the argument.

2. Nozick's Garage. The Experience Machine

Developed by Nozick (1974, 42-45) and alluded to in his subsequent works (1981; 1989), the 'Experience Machine' (EM) has elicited in the past decades a wide array of acclaim and critical responses. It is located in what might seem a strange context, in a chapter dedicated to a discussion regarding moral constraints and the state. More precisely, it follows his normative stance of how libertarian side constraints work and an assessment of the possible ways in which we should approach

[5] We remain agnostic though about the possibility that other future technologies might provide more accurate approximations of the conditions in the original thought experiment.

animal ethics. Would utilitarianism, Nozick wonders (1974, 39-41), be an adequate ethical theory to address the issue of how humans should treat animals? Couldn't there be something else relevant for animals, other than their felt experiences? What about our case, as human beings: is there something else that should matter to us, besides our felt experiences?

This is how Nozick gets to the EM, through a "digression within a digression within a digression" (Feldman 2011, 63). Before analyzing the implications of the thought experiment, let's see how it unravels: "[s]uppose there were an experience machine that would give you any experience you desired. Superduper neuropsychologists could stimulate your brain so that you would think and feel you were writing a great novel, or making a friend, or reading an interesting book. All the time you would be floating in a tank, with electrodes attached to your brain. Should you plug into this machine for life, preprogramming your life's experiences?" (Nozick 1974, 42). There is no cap on the possibilities you have while inside the EM, as you could pick whatever you want from an open library of potential experiences and, after two years, choose a different scenario. Moreover, if you fear losing your relatives or loved ones, rest assured: they would also have the opportunity of plugging into the EM. Should you, Nozick asks his readers, live all your life inside the EM? Is there something else that matters to us, besides "how our lives feel from the inside?" (Nozick 1974, 43).

The short answer, Nozick posits, is that you shouldn't. His intuition is that, by imagining such a hypothetical scenario, we learn that there is something else that matters to us besides sheer pleasurable experiences while being plugged into such a device. First, "we want to *do* certain things, and not just have the experience of doing them. […] A second reason for not plugging in is that we want to be a certain way, to *be* a certain sort of person. Someone floating in a tank is an indeterminate blob. There is no answer to the question of what a person is like who has been in the tank. Is he courageous, kind, intelligent, witty, loving? [...] Plugging in is a kind of suicide […] Thirdly, plugging into an experience machine limits us to a man-made reality, to a world no deeper or more important than that which people can construct. There is no *actual*

contact with any deeper reality, though the experience of it can be simulated"(ibid.).

Later on, while exploring the issues of happiness and meaning of life in *Philosophical Explanations*, Nozick addresses again the third reason for not plugging into the EM: such a life would be enclosed "within the circle of just your own experiences" (1981, 595). He arrives at the same point one more time, asserting that the question of plugging in is a question of value. The issue is not whether or not it would be preferable to enter the EM if you live a miserable life, but whether being connected to the machine is the best life you could have. While in the EM we are opaque to how reality really is: "what we want and value is an actual connection with reality […]. To focus on external reality with your beliefs, evaluations, and emotions, is valuable *in itself*, not just as a means to more pleasure or happiness" (Nozick 1989, 106).

The question of what Nozick intended to prove with this thought experiment has puzzled philosophers and psychologists for the last couple of decades. Feldman (2011, 64) hypothesizes that there are four main possible ways in which we could interpret the EM:

(1) a critique to utilitarianism;

(2) a critique of ethical hedonism as a theory of individual welfare;

(3) a potential knockdown argument against all mental state theories of welfare;

(4) a critique of psychological hedonism regarding human motivation.

A tentative overlapping consensus between scholars is that the EM tried (more or less successfully) to provide a fatal blow to hedonism, broadly understood (Crisp 2006; Barber 2011; van der Deijl 2019; Vitrano 2020) or to desire satisfactionism (Lowe and Stenberg 2017). Still, some openly criticised Nozick's design and use of the EM. For Pietrucha (2017), not wanting to plug might not have anything to do with whether or not we see any intrinsic value in living in reality, but with our imaginative failure. Weijers sees the EM as what Dennett called an 'intuition pump': the thought experiment is designed so as to lead the respondents to a particular response based on intuitions. What might explain Nozick's intuition that people would not want to live their lives plugged into the EM is the status quo bias (Weijers 2011; Belshaw 2014). The objection is vindicated by some experimental evidence with testing

the EM (De Brigard 2010; Weijers 2014). Löhr (2018) examines whether there are any differences between how laypersons and people with a background in philosophy respond to EM-like scenarios and concludes that most people without a degree in philosophy do not reject plugging in because they reject hedonism, but due to other reasons, like the possibility that the machine could malfunction. All such experiments have been, in turn, criticized because the subjects (most of them college students) cannot be "in the state of confronted agents" (Smith 2011) or because a real empirical test would be with regards to something that Rowland (2017) dubbed the "Undistorted Experience Machine".

Whether or not Nozick wanted to fight the 'utility monster' or against hedonism and whether or not he was successful in his venture is an issue on which we take no stance here. But our argument might have some indirect implications for the idea that we could empirically test people's reaction to EM's or EM-type of situations. Nozick himself makes an analogy between the imagined device and psychoactive drugs, "which some view as mere local experience machines" (1974, 44). Could MMORPG's also be EMs in Nozickian terms? If yes, then experimental philosophers would have endless avenues open for empirically testing our intuitions regarding utilitarianism, ethical/psychological hedonism or mental state theories of welfare. For that to happen, we need to take a closer look at what immersion in gaming entails and whether agents are in a similar state then the ones who would plug in to Nozick's EM.

3. Gaming and Immersion

Defining immersion in video games is a difficult task, for a number of reasons. Video games have a dual nature, at the same time being objects of art and of digital engineering. The quality of the experiences they provide is inherently subjective and, as such, difficult to pin down in a definite classification and/or measurement. No matter how thorough a scientific investigation we may provide, one will always find a transgressive example that defies what seemed obvious and hard to contradict. A gaming experience works differently than a user experience for an operating system or text editor, even if the means of interaction

are seemingly the same. A confusing and convoluted user interface may work for and not against the success of a game (an example is the *Dark Souls* series).

Immersion, what some in the industry have called the "Holy Grail" of gaming, is an ever-present term in the conceptual space of video games (Rolling and Adams 2003). Game developers praise their creation as "the most immersive game ever made", reviewers around the Web acclaim games that provide immersive experiences, and deplore "immersion-breaking" elements. There is even a whole genre of video games that get its name from this moniker: the immersive sim (the *Deus Ex* series, the *Thief* series, the *BioShock* series and many other milestone games). Given its widespread use and importance, it is no accident that immersion has prompted many scholarly attempts to catch exactly what gamers mean when they call their gaming experiences immersive.

These studies can be grouped into more empirical approaches, which take the form of sociological or psychological studies into the experience of games, or more personal approaches (for lack of a better word) which use the authors' own experience with the medium as the cornerstone of the study.[6] For the purposes of this section, we will concentrate on empirical approaches, given the highly subjective nature of the gaming experience. Providing a comprehensive definition for immersion is outside our purpose. Instead, we aim to extract from the literature the traits of immersion in games most people experience, and then compare those with the kind of immersion presupposed by an EM.

We do acknowledge, however, some caveats of the empirical studies: small and skewed samples of the population (game players come from all walks of life, not only computer science departments) and a sometimes problematic choice of the games. The gaming industry moves at an astonishing pace, as well as the preferences of its fans: what used to be a state of the art photo-realistic experience ten years ago looks rather ridiculous today to the same player. This is why using an old game like Tetris (Sigailov-Lanfranchi 2019) to explore immersion in contemporary times might not yield warranted empirical results. MMORPG's come

[6] A third, but less common, possibility involves a more conceptual approach. Calleja (2011) is an example.

with challenges of their own for these studies – they are conceived by their developers as a service to be continuously improved, not a standalone game. The *World of Warcraft* of today is not, in many ways, the same game that it was ten or five years ago.

One of the first qualitative studies to provide a theory of immersion was published by Brown and Cairns (2004). They found that gamers distinguish between different degrees of immersion in the games they enjoy playing. Brown and Cairns went to crystalize three levels of immersion: (i) engagement – gamers simply learn to play the game investing time and effort; (ii) engrossment – gamers dedicate a significant part of their attention to the game and also become emotionally involved in it; (iii) total immersion – when the player is completely involved in the game and cut off from reality. Brown and Cairns use the word "presence" in the case of total immersion, but further literature on this subject seems to indicate something else by "presence", a spatio-temporal location of the player "in the game". Not all immersive games need something like this. For our purposes, though, MMORPG's usually involve some form of presence through player-avatars exploring and interacting in and with extensive and complex virtual worlds, so Brown and Cairns' evaluation can apply.

Building on this earlier work, Jennet *et al.* (2008) tried to obtain a way to measure the extent to which players feel immersed while engaged in a video game. The result was a questionnaire built from 31 Likert scale questions that was validated through a large-scale study (Jennett *et al.* 2008, 658). Jennett *et al*. also maintained that "immersion is a concept which we argue is important to gaming, transcending existing cognitive theories of flow, CA [cognitive absorption – n. red.] and presence. Immersion is an experience in one moment in time and graded (*i.e*. engagement, engrossment, total immersion). Immersion involves a lack of awareness of time, a loss of awareness of the real world, involvement and a sense of being in the task environment" (Jennett *et al.* 2008, 657).

The main result of the study is that the concept of immersion was irreducible to other manners of characterizing the experience of playing digital games. The theories of flow, cognitive absorption and presence that Jennett *et al*. mention in the quote above are not chosen

accidentally, as they may be used to describe the phenomena involved by immersion. Thus an appeal to a specialized study of immersion would provide fertile research. More importantly, their study suggested that five "clear factors" could be involved in achieving immersion – three personal factors (cognitive involvement, real world dissociation and emotional involvement) and two game factors (challenge and control).

As excited as gamers, developers and journalists could be about calling the latest hit in the video gaming industry "immersive", what they may actually be saying is that they've entered a state of flow while they were playing the game, the game-world managed to overtake most of their cognitive processes or that the sense of presence in the game was masterfully achieved by the game's creators. Jennet *et al*. and other studies (Cairns *et al.* 2014, 343-345) argue that the conflation of immersion with these concepts is mistaken.

We'll briefly look at each of them in turn. "Flow" is a concept used in positive psychology that was developed by Mihaly Csikszentmihalyi (1975) and it represents "the feeling of complete and energized focus in an activity, with a high level of enjoyment and fulfillment" (Chen 2007, 31). The idea is to be fully concentrated on performing an activity (which can be anything, from washing clothes to particle physics research) and for it to give a sense of achievement and gratification. In his later work, Csikszentmihalyi identified the eight components of being in a state of flow: the activity must require some skill, a merging of action and awareness in the subject, clear goals, direct feedback, concentration on the task at hand, a sense of control, a loss of self-consciousness and an altered sense of time (Csikszentmihalyi, 1990). The role of flow in video games has been the focus of some research. Jenova Chen (2007), for example, has argued that games should be designed in such a way that the experience of flow is encouraged and has provided a methodology to help developers achieve that.

Despite superficial similarities, flow and immersion are distinct, even if related states of mind. Firstly, "immersion is more of a graded experience whereas flow is an all-or-nothing sense of being <in the zone>" (Cairns *et al.* 2014, 344). Secondly, being in a state of flow may not be necessary for achieving immersion. Some games, especially horror games like *Amnesia: The Dark Descent* or *Alien: Isolation*, deliberately

disrupt the flow of a player engaged in gaming activity like gathering resources and solving puzzles to increase tension and increase immersion. As Cairns *et al*. also note, a game does not need to provide a clear goal to foster the immersion of the player – sometimes great graphics, sound and story are enough (Cairns *et al.* 2014, 344).

Cognitive absorption is a personality trait that speaks about a subject's willingness to become involved with their general use of technology, while the state of immersion results from the active interaction between player and video game (Agarwal and Karahana 2000). One can possess the characteristic of cognitive absorption without being immersed in the game-world, thus, the two concepts should be accounted for by different means.

Lastly, having a sense of presence is not required for a player to be immersed in a game. The simple reason is that not all immersive games provide the player with an avatar that interacts in a meaningful way with the digital play space. Some puzzle games or strategy games may be immersive even if they don't look the part and don't give the player the actual sensation of "being" in the game. Presence would surely help immersion, but is not required.

Ermi and Mayra (2005) have proposed a model of immersion that provides three types of immersive experience for the player: sensory, challenge based and imaginative. Sensory immersion is the type of immersive experience that arises when a player is involved in a graphically intensive game. Visual and auditory stimuli afforded by the game are the main ways in which the player is immersed in the game world. Challenge based immersion is related to the difficulty of the game and skill required from the player to achieve success in the game (the *Dark Souls* series is a good example of immersion through high difficulty). Imaginative immersion is achieved by the player's imagination being stimulated by a game's narrative or emotional content. Rather than giving an account of the different levels of immersion that can be provided by video games (like Brown and Cairns 2004 and Jennett *et al.* 2008), this model tries to account for the different ways a game may "trick" the player into being immersed into its digital world (Ermi and Mayra 2005). The model was refined by Arsenault, who replaces the notion of challenge based immersion with systemic immersion – the

rules of interaction provided by the game replace the general rules of interaction with the real world in the mind of a player immersed in a game experience. The game needs not be difficult, but have game mechanics doing a good enough job so that the player has the illusion of performing a "real" activity, despite the fact that she is holding a game controller or just pressing buttons on a keyboard. Arsenault also proposed the replacement of imaginative immersion with fictional immersion, as it is the game's fiction that stimulates the imagination of the player (Arsenault 2005).

Another conceptualization of immersion was proposed by Adams. Here, we also have three types of immersion, but they look at immersion from the perspective of the different ways a player may interact with the game. These types are tactical, strategic and narrative immersion. Tactical immersion occurs in moment-to-moment gameplay and is related to the decisions required from the player in negotiating the immediate challenges in the game. Strategic immersion can appear in games that provide long-term goals to their players (most contemporary games do that through various forms of upgrade systems for their avatars, for example). Narrative immersion, as discussed regarding the previous model, is the type of immersion that takes place when the player becomes emotionally involved with the game's story and setting (Adams 2004).

Haggis-Burridge (2020) provides yet another range of classifications for immersion: (i) systems immersion (players are engaged by the game mechanics and are in a state of flow); (ii) spatial immersion (players feel like they are actually present in the virtual world); (iii) emphatic/social immersion (players develop an emotional connection towards the characters in a virtual world – the characters can be artificial or actual humans); (iv) narrative/sequential immersions (players feel a compulsion to see how a game's story unfolds or to explore new regions of the game's world and mechanics).

MMORPG's are video games that also provide a strong social aspect to gameplay: players engage in the game's challenges together, form communities and interact with one another in all aspects of gameplay. To narrow down immersion for MMORPG's, we need to look at how the presence of other players in the game-world can affect

immersion. There is some evidence that a social setting actually increases the immersion of the players (Cairns *et al.* 2013), but we will return to this point in Section 5.

It is worth noting that not all research agrees that immersion should be conceptually distinct from flow. Michaillidis *et al*. (2018, 4) argue that the studies cited in this section fail to conceptually differentiate between flow and immersion in a convincing way as there is no research to prove that some games are better than others at triggering flow over immersion or the other way around.

To sum up, research into immersion is not ready to provide a one size fits all definition of the concept. We can however note some common themes that appear in all the models we have reviewed here. One important way in which video games achieve immersion for their players is given by the means through which the player interacts with the games: Jennett *et al*. speak about challenge and control, Ermi and Mayra (and Arsenault as well) underline the importance of good game mechanics in getting the players immersed and Adams' model has two main components that relate immersion to the tactical and strategic aims a player's actions may have in game. Even if immersion were not distinct from flow (Michaillidis *et al.* 2018), the same conclusion would plausibly hold: to be immersed, players need to feel that they are actually doing something, that they are getting rewards based on actions they have achieved in the game world. As such, it seems safe to say that providing the player with a sense of agency in her interaction with a digital game is capital to achieve immersion. The player needs to have her skills tested and her actions be meaningful to truly feel immersed. She needs to feel like she is the one who is living a digital life and that her decisions do indeed carry some weight and consequences. While other mediums (like film or literature) may get to some form of immersion through visual or imaginative stimuli, video games' specific type of immersion require a sense of achievement through doing and not only feeling. MMORPG's with their complex systems (crafting, combat, base-building, etc.) manage to provide that along with social aspects that enhance immersion. Given that this concept of immersion can be plausibly tied to the notion of player agency, it remains to see whether one could actually be immersed in an EM.

4. Being Immersed in the Experience Machine

If the EM had been the product of a major game developer, instead of a philosopher's figment of theoretical imagination, it would have been surely marketed as the device for the ultimate immersive experience. Gaming companies are not shy as a rule when bragging about the immersion their products are able to elicit from users, and if a fully EM-like technology were available, they would probably not hesitate to seize the marketing opportunity.

At least at a first glance, such counterfactual boastfulness seems justified by a certain overlap or similarity between the state of "being plugged" into the EM, described by Nozick, and the state of immersion sought by the gaming community and studied in the literature we reviewed above. In both cases, the subject is somehow cut off from the reality outside the machine or the game world, and thus opaque to the inputs it generates. But this perceived overlap deserves a more in-depth scrutiny.

One can describe the Nozickian state of being plugged as a special kind of immersion; it wouldn't be a particularly unfortunate choice of words. What are the specific features of this type, however, can make quite a difference when set against its gaming counterpart.

Two main conditions, in terms of immersion, should be met by the EM in order to be able to do its purported job:

(1) There is a guaranteed and complete immunity of the subject from any stimuli generated by changes in the circumstances of the "real", physical world

(2) While being plugged in, the subject cannot be aware of the actual source of her experiences ('while in the tank you won't know that you're there; you'll think it's all actually happening' (Nozick 1974, 43), even if the initial decision to connect to the EM were voluntary.

These two conditions could arguably be collapsed into a single and more general formulation ("guaranteed immunity from any stimuli or thoughts not generated by the machine through the simulation of

preprogrammed events and experiences")[7]. The choice to draw an analytical distinction between them is still justified by the fact that they are meant to capture two different but equally necessary sets of requirements. The first relates to blocking the possibility to receive stimuli from the world outside the machine, the other to the subject's complete lack of awareness as to the source of her experiences.

We encounter quite similar, though not identical, presuppositions in other famous philosophical thought experiments. The prisoners in Plato's Cave allegory (*The Republic*, 514a-520a), for example, only see the shadows projected on the wall in front of them by the puppet masters and, moreover, they mistake the shadows for real objects. But, as John Finnis once insightfully observed, choosing to plug into the EM "would be to bury oneself in a tomb much deeper than Plato's cave; the experience would, *ex hypothesi*, be more veridical than that of Plato's prisoners, but the actual divorce from reality would be more total than theirs, since they could *actually* communicate with, for example, those who, having ascended to a clear view of reality, had returned to tell them about it" (Finnis 1983, 41). Nozick's 'indeterminate blobs' are not presented with any such opportunity.

The EM is such an enormously troubling thought-experimental setup precisely because it presupposes (1) and (2). If one or both were absent, the machine would be just a mere pleasure machine that can be turned on and off at will or can be easily overridden by changes in circumstances (either exterior or in the subject's volition).

Imagine the superduper neuropsychologist offering you a similar deal to the one in the initial scenario (electrodes attached to you brain, simulating the experiences of your choice and all that), but with some further provisions written in small print: (1) your ride can be interrupted at any time by external events for which the machine operator takes no responsibility (like your phone ringing, or somebody entering the room) and (2) at every point while being plugged in, you are aware that your experiences, though extremely vivid and intense, are artificially generated by the EM. It would be like an experience rollercoaster: the excitement is

[7] These two conditions have been previously discussed by Tavinor (2009; 2017). We elaborate on his insight by explicitly identifying them as conditions for immersion.

real, and the Luna Park around fades almost completely into oblivion, but you are still well aware that it's the rollercoaster ride giving the excitement. Such a scenario might raise some interesting philosophical questions about hedonism and valuable pleasures on its own, but they would be different, and probably less dramatic, than those springing from Nozick's EM. And it definitely wouldn't be about "preprogramming your life's experiences"; nor would the plugging in amount to the sort of "suicide" that Nozick mentions, after which "[t]here is no *actual* contact with any deeper reality", but only with the artificial one simulated by the machine.

Let's call this special type of immersion presupposed by the EM "full immersion" and set it against the most comprehensive kind achievable through the technology currently available and deployed in MMORPG's.

Brown and Cairns (2004) and then Jennett *et al.* (2008) distinguish between three different grades or levels of immersion: engagement, engrossment and total immersion. At the highest level of total immersion, "gamers described a sense of presence, being cut off from reality to such an extent that the game was all that mattered." (Jennett *et al.* 2008. 642). It is acknowledged by the authors to be a rather short-term, rare and unusual occurrence, in which gamers feel part of the gameworld and forget that they are just playing the game.

However, even this highest level of immersion in games (admittedly, the closest to an EM-like situation) fails to meet the requirements that are presupposed by Nozick's counterfactual scenario. True, in total immersion players detach themselves from the reality and the game is all that matters; this description might at a first glance be taken as an approximation of (2). But the wording here is indicative of an important nuance: the game is "all that matters", not "all there is". The way gamers describe their experience points to a state in which they are so much 'into' the game that they lose any interest in the outside world and do not register events that happen in it (*e.g.* they don't hear sounds, or realize they are hungry, etc). But it does not seem to be a state in which all awareness of the fact that they are actually playing somehow ceases. Immersion in MMORPG's, no matter how deep, is based on the players retaining some agency (if only for gameplay decisions), while the EM scenario is built on the opposite assumption.

Let's assume, however, that gamers descriptions are inaccurate and it does cease, even if for a short time only (but long enough to count – or we can disregard the time span requirement altogether for the sake of the argument). So, in this highly far-fetched scenario, the second condition would be met. What about the first though?

Many gamers know the frustration of experiencing a muscle cramp in the real world, just as the avatar's battle against the last boss in the dungeon reaches its climax in the virtual one[8]. Such physiological events are as much outside the control of the technology as the changes in the physical reality are. Sensations of pain, hunger, cold, etc are prone to break immersion or flow. The player can, of course, choose to keep playing regardless, in the same way that a professional athlete can choose to play an important game through a painful injury. But the simple fact that such a decision is taken indicates the end of, or at least a pause in the immersive experience. Similar effects are produced by disruptive events in the physical world, like noises, or by sudden and guilty realizations of the "oops, I was supposed to submit a school project today" kind.

This lack of a complete and guaranteed immunity with respect to other external stimuli, as well as the sometimes unpleasant awareness of the fact that you're just playing and not actually *in* the game world, would be fatal malfunctions for an EM. If Nozick's super scientist sold a subscription to one of the MMORPG's currently available on the market as the ticket for an EM ride, maybe the gullible client should ask for a refund.

To phrase it differently, playing MMORPG's, although immersive in many significant ways, is not quite the same as having the preprogrammed experience of playing MMORPG's while fully immersed into the EM. As Tavinor (2017, 100) argues, "the experiences afforded by the virtual worlds of video games are not indistinguishable from those of the real world, because games provide only limited sensory channels to their

[8] The example may sound cartoonish, but there are a number of injuries and diseases with very high incidence among professional or heavy gamers (carpal tunnel syndrome, stenosing tenosynovitis or the so-called tennis elbow) and that can become severe enough to endanger or even end their gaming careers. Cf, for example, dr. Levi Harrison, interviewed by Emanuel Maiberg (2015) for *Vice Magazine*.

artificial worlds, comprising audio-visual depictions, and sometimes very limited tactile or haptic feedback".

But then, it is not the MMORPG technology functioning as an EM, but the machine itself. We'll return to this point, and make some further considerations about the implications of our position, after a brief critical discussion of two other arguments in the literature exploring the same issue.

5. Assessing Other Objections: From Sociality to Sensory-Immersion

Several other objections (Silcox and Cogburn 2009; 2014; Taylor 2007) have been advanced in relation to the possibility of MMORPG's acquiring the status of EM in the sense that Nozick discussed. While both capture something relevant, they do not go far enough. Our purpose in this final section of the paper is to show why both arguments fall short..

Silcox and Cogburn see the emergence and evolution of MMORPG's like *World of Warcraft* and *EVE Online* as a potential avenue for testing our intuitions regarding Nozick's machine. With the technology employed in video games that we have at our disposal we are in a much better position to do this, but "the person who chooses to inhabit Azeroth for even a very substantial portion of his life would of course be in a somewhat different position from that of the hypothetical person'' (2009, 57) who chooses a wide array of experiences from the digital library of the superduper neuropsychologists. Their main argument is that an MMORPG cannot be an EM because of the particularities of the interactions between users, mediated by avatars: "[t]he massive multiplayer aspect of WoW is precisely why it is not an experience machine in Nozick's sense. When characters are in guilds, their players in "the real world" have formed a community to assist one another, and one's play does affect the lives of other players' characters

as well as the people playing them'' (Silcox and Cogburn 2009, 59-60). A device like the EM, they add, would only simulate such interactions[9].

In a later paper, while arguing against Nozick's claim that no person would want to plug into the EM, Silcox and Cogburn distinguish between Virtual Reality technologies and the Virtual Reality while you're plugged into the EM, highlighting a fundamental ontological presupposition of Nozick's thought experiment. While inside the EM you would live in a solipsistic world, in MMORPG's like *World of Warcraft* "players play in groups and people rightfully judge one another by the behavior of their avatars'' (2014, 570).

In our opinion, this objection, while valid, cannot be a knockdown argument against the possibility of MMORPG's becoming EMs. If immersion is a subjective experience, would it really matter that the interaction between users is mediated by online avatars? As long as the stimuli are 'felt' inside the gameworld, their source is irrelevant. They might as well be generated by the AI of the game, as there are a series of interactions with bots inside the worlds of different MMORPG's.

The social aspect of gaming has also been an aspect discussed in the scholarly literature more specialized in video game studies. Along the same lines as Silcox and Cogburn, Sweetser and Wyeth have considered sociality as detrimental to immersion, as being aware of other real persons involved in the game should break the illusion of the virtual (Sweetser and Wyeth 2005). Other authors (like the previously cited Ermi and Mayra) have concluded that social interaction is a marginal element in relation to the degree in which players may become immersed with a digital gaming experience.

The apparent tension between immersion and sociality has been called into doubt by three experiments devised by Cairns *et al.* (2013). Using the games *WoodPong*, *Midtown Madness 2* and *MarioKart Wii*, the researchers tested whether players felt more immersed when playing against a computer, against another player through online multiplayer

[9] It might be worth noting that our argument does not rest on a reading of Nozick's EM that is different from the one employed by Silcox and Cogburn. Our aim is just to show that there are stronger reasons supporting the claim that both ourselves and Silcox and Cogburn put forward.

or against a co-located player. While there wasn't enough evidence to support a meaningful difference between playing against a player at a distance or in the same room, it was found the playing against the computer was a much less immersive activity rather than playing against other people, contradicting previous views on this subject (Cairns *et al.* 2013, 1072-1076). The results of this study support our reply to Silcox and Cogburn's objection, as they show that playing games in a social environment actually works for immersion rather than against it. As MMORPG's are essentially social experiences, if done right, they may prove even more effective tools for immersion than other single player experiences.

A different objection was raised by Taylor (2007). Assessing the implications of MMORPG's for philosophy, he conjectures that games like *World of Warcraft* are almost EMs, but not quite. They do offer immersive, pleasurable and persistent experiences to the players who log in, but there is one major difference: MMORPG's are not sensory-immersive and hence they do not manage "to allow the players to perceive the fictional reality as if it were real'' (Taylor 2007, 71). For a game like *World of Warcraft* to acquire the status of an EM it would need to make us ''forget anything we know about the real world ... [and] sever our link with anything that we have been or done in the real world'' (Taylor 2007, 74).

In saying this, Taylor addresses in his argument one of the philosophical presuppositions behind the thought experiment: for an MMORPG to be an EM, it should be sensory-immersive in the sense that the player cannot be aware that the fictional reality is just, well, fictional. It is in this sense that the sensory-immersion Taylor writes about is similar to the standard reading of the Nozickian setup of the thought experiment: while floating in the tank but being plugged in you cannot be aware of the actual source of your experience, namely the AI of the computer your brain is connected to. The same should happen while playing MMORPG's if they were to be considered EM.

While provoking and in line with some of the philosophical assumptions of Nozick's thought experiment, sensory-immersion seems to also fall short of the role of best candidate for explaining the difference between MMORPG's and the EM, though it might shed some

light on the issue. Taylor wrote his article in 2007, when MMORPG game design and technology were still in their post-infancy period. Since then, both game design and technology evolved at an amazing pace in the direction of sensory-immersion of MMORPG players, with a big plus for technology. To take just one example, think about the endless possibilities that can be made available by the Oculus Rift, Valve Index and other VR headsets, though the costs are still pretty high for average young users.

For $400, gamers can buy an Oculus Quest headset that ''has the inherent potential of manipulating people's mind with a superlative 3D experience'' (Desai *et al.* 2014, 175). For an additional $249.99 you can also get a VR Sensory-Immersion Generator, which will bring the full sensory-immersion Taylor wrote about. The product is actually marketed taking into account and highlighting the demand on the market for such sensory-immersive devices, triggering our senses of touch, smell and taste into the game. It basically aims at enhancing virtual reality experiences through sensory-immersion. Valve's Index headset goes even further, as its motion controllers can detect and simulate the action of each finger in the real world, thus allowing players to actually grab, point, push or pull in the virtual world with the same hand movements they use in their daily lives.

Game critics are also enthusiastic about the potential of developing VR MMORPG's, though their costs might be prohibitive for the time being (Ilano 2016). We have shown that sensory-immersion is not impossible. However, MMORPG's, even incorporating this new and emerging VR technologies, would still fall short of being considered EMs. First of all, they would still fail to guarantee the user/player immunity from external factors like pain or hunger. Secondly, playing a VR MMORPG, regardless of how immersive it might be, would still rely on the "attitude of make-believe and the player's acknowledgement of the artificiality and fictionality of the world in which he or she plays" (Tavinor 2017, 100). Furthermore, many of the contemporary improvements to VR gaming aim to provide a greater and more precise interactivity with VR game worlds. The aim of VR developers seems to be to increase the ability of the player to achieve greater agency rather than just provide a better passive experience of the virtual space.

6. Implications and Further Discussion

Up to this point we have argued that MMORPG's, at least in their current technological stage of development, are not EMs in the sense of Nozick's thought experiment. They fail to satisfy the necessary conditions and, for the predictable future, it is rather unlikely that things will change dramatically. There are two happy and straightforward implications of our argument. First, we still have quite some time left to think through the possible answers to the difficult questions raised by Nozick's hypothetical scenario, before MMORPG technology turns them into actual life choices. Second, the worry that gamers are in any clear and present dangers of becoming the 'indeterminate blobs' Nozick mentions, as a result of MMORPG technologies, is unwarranted.

Both implications, however, are in need of some qualifications. Let us start with the second. Maintaining that a particular worry is unwarranted has no bearing upon the way we assess other related sources of concern. We are not trying to say that MMORPG's, and digital games in general, are fully proofed against a large array of ethical, psychological and social concerns, some of them possibly related to the immersive qualities of games. Addiction, to name but one, is already a well-documented threat. There is a bounty of evidence available about its effects, both anecdotal (in media coverages and cases from almost everybody's circle of friends and acquaintances) and scientific (we follow the extended overview available in Klemm and Pieters 2017). Neglect of health and hygiene, failures in professional and family lives due to the amount of time and energy spent for playing are among the possible outcomes listed. One notable difference they mention between addiction to games and other addictions (to drugs, alcohol, smoking etc) is the lack of the physiological dimension encountered in the latter. But we may reasonably conjecture that immersion is what contributes instead to the addiction in the case of games. It might not be totally meaningless to ask how far an analogy between the addicted gamer and Nozick's indeterminate blob floating in a vat can go; but we hope to have shown that it is only partial, and that the missing components are significant. Moreover, the worrisome self-destructive effects of addiction seem to be quite similar across the

board, irrespective of its object (addiction to sports betting might be more widespread, for example, while having neither a physiological, nor an immersion component).

Turning to the first implication, it might be worth observing that, although MMORPG's do not offer us the option to plug into an EM, they do offer an opportunity to experience some artificially-induced pleasures and to break away for a while from the troubles and constraints of everyday, 'real' life, especially if the agent does not regard it as fulfilling enough. Inasmuch, they could serve as a more vivid illustration of the high stakes behind the philosophical questions about what pleasures are valuable.

The related topic of 'escapism' has been also widely discussed recently in the literature on computer games, with MMORPG's being seen sometimes as a setting for active escapism (Kuo *et al.* 2016). Moreover, escapism has been directly linked with immersion. An interesting result in the qualitative study by Warmelink *et al.* (2009) is that gamers equate one of the eight identified forms of escapism with hardcore immersion. In this sense, casual gaming cannot be escapist (but we should notice that at least four other forms mentioned by the authors might be related to the immersive qualities of the game). We do not wish to deny the importance of the issue for today's societies, nor take a particular stance in the debate. Still, it seems to us that when assessing escapist behavior, questions about the *escaper* and about the reality she is trying to escape *from* are *prima facie* just as important, epistemologically prior and maybe even more worrisome, than the destination she is trying to escape *to*.

As a further suggestion, it might be that Plato's cave allegory provides a more adequate analogy to MMORPG's than Nozick's machine. The gamer's route back to reality, just like the characters' in the cave, is not blocked *ex hypothesi* and remains available to her, through effort or accident. One of the important differences though, pointing back to the problem of escapism, is that gamers are voluntary, or indeed eager prisoners, unlike their speluncean counterparts.

Tavinor highlights another difference between playing an MMORPG and being plugged into an MMORPG's world: "If these games provide a version of Nozick's machine, it is one in which the moral nature of the

world is almost the opposite envisaged by Nozick: life is nasty, brutish, and short (it is a puzzle then, why players choose to play in these worlds!)" (2017, 100). This is related to the subjective sense of safety the player would have in these two scenarios. The fantasy worlds of MMORPGs would be horrific if they were actual real worlds - realms of perpetual conflict where the possibility of death and torture is always around the corner. The experience of being a heroic soldier involves much more trauma than the experience of being a successful and famous writer. The whole selling point of violent video games (and MMORPG's by extension) is the caricaturization of violence to the point where it feels non-threatening to the player. Violence in the EM, given the standard reading of the experiment, would have to completely resemble that in real life. While we have not found any empirical studies to prove this, we can safely assume that no one plays video games to experience actual world violence with all its very real and painful consequences. Even the most decorated and heroic fighter has seen her share of wounds and fallen comrades.

Could future MMORPG's become EMs? We can imagine a scenario in which the two conditions we argued for were almost met. Maybe the player will be in a coma-like induced state, fed through tubes and perfusions, with all receptors for pain, cold etc, as well as her previous memories, chemically blocked. The brain would only be stimulated, such that the only reality accessible for her would be the magical and amazing world of the game. Would that qualify as a close-enough approximation of the EM? One way to counter the objection is to address its feasibility. No technology, no matter how intensely tested, is 100% proof against hardware failure or software bugs. Even in such a case, the immunity with respect to non-machine generated stimuli is not guaranteed and complete. But a more philosophically interesting answer would point to the fact that it would not be the MMORPG playing the role of EMs in such a scenario, but the entire setup. Worse, the device takes away the main source of enjoyment in games. Taylor (2007, 71) is right to say that people sign onto MMORPG's "specifically for pleasure". But we sometimes tend to forget that it is the pleasure of actively *playing* the game, together with others, not of being sucked in as mere passive receptors of an elaborate illusion.

REFERENCES

Adams, Ernest (2004). "Postmodernism and the Three Types of Immersion", (internet) available at <http://www.gamasutra.com/view/feature/130531/the_designers_notebook_.php> (accesed August 11th, 2019).

Agarwal, Ritu & Elena Karahana (2000). "Time Flies when You're Having Fun: Cognitive Absorption and Beliefs about Information Technology Usage", *MIS Quarterly* 24(4): 665-694.

Arsenault, Dominic (2005). "Dark Waters: Spotlight on Immersion", *Proceedings of Game-On North America 2005*, 50-52.

Barber, Alex (2011). "Hedonism and the Experience Machine", *Philosophical Papers* 40: 257-278.

Baudrillard, Jean (2001). *The Vital Illusion*. NY: Columbia University Press.

Belshaw, Cristopher (2014). "What's Wrong with the Experience Machine", *European Journal of Philosophy* 22: 573-592.

Brown, Emily and Paul Cairns (2004). "A Grounded Investigation of Game Immersion", *Proceedings of ACM CHI 2004*, 1297-1300.

Cairns, Paul, Anna L. Cox, Matthew Day, Hayley Martin & Thomas Perryman (2013). "Not Winning but Beating: The Effect of Social Play on Immersion in Videogames", *International Journal of Human-Computer Studies* 71(11), 1069-1077.

Cairns, Paul, Anna L. Cox and A. Imran Nordin (2014). "Immersion in Digital Games: Review of Gaming Experience Research", in *Handbook of Digital Games*, edited by Marios C. Angelides and Harry Agius, Hoboken: John Wiley & Sons, pp. 339-361.

Calleja, Gordon (2011). *In-Game: From Immersion to Incorporation*. Boston: MIT Press.

Castronova, Edward (2005). *Synthetic Worlds: The Business and Culture of Online Games*. Chicago: Chicago University Press.

Chen, Jenova (2007). "Flow in Games (and Everything Else)", *Communications of the ACM* 50(4): 31-34.

Cogburn Jon & Mark Silcox (2014). "Against brain-in-a-vatism: on the value of virtual reality", *Philosophy & Technology* 27: 561–579.

Crisp, Roger (2006). "Hedonism Reconsidered", *Philosophy and Phenomenological Research* 73: 619-645.

Csikszentmihalyi, Mihaly (1975). *Beyond Boredom and Anxiety: Experiencing Flow in Work and Play*. San Francisco: Jossey-Bass.

Csikszentmihalyi, Mihaly (1990). *Flow: The Psychology of Optimal Experience*. London: Harper Perennial.

De Brigard, Felipe (2010). "If you like it, does it matter if it's real?", *Philosophical Psychology* 23: 43-57.

Desai, Parth Rajesh, Pooja Nikhil Desai, Komal Deepak Ajmera & Khushbu Mehta (2014). "A Review Paper on Oculus Rift-A Virtual Reality Headset", *International Journal of Engineering Trends and Technology* 4: 175-179.

Ermi, Laura & Frans Mayra (2005). "Fundamental Components of the Gameplay Experience: Analysing Immersion", *Worlds in Play: International Perspectives on Digital Games Research* 37: 37–53.

Feldman, Fred (2011). "What we learn from the experience machine" in *The Cambridge Companion to Nozick's Anarchy, State and Utopia*, edited by Ralf M. Bader and John Meadowcroft, Cambridge, MA: Cambridge University Press, pp. 59-86.

Finnis, John (1983). *Fundamentals of Ethics*. Washington DC: Georgetown University Press.

Ilano, Ian (2016). "What's really needed to make VR MMOs a reality?", (internet) available at: https://www.gameskinny.com/m49u9/whats-really-needed-to-make-vr-mmos-a-reality (accesed June 9th, 2019).

Haggis-Burridge, Mata (2020). "Four categories for meaningful discussion of immersion in video games", (internet) available at <https://www.researchgate.net/publication/340686774_Four_categories_for_meaningful_discussion_of_immersion_in_video_games> (accesed July 7th, 2020).

Jennett, Charlene, Anna L. Cox, Paul Cairns, Samira Dhoparee, Andrew Epps, Tim Tijs and Alison Walton (2008). "Measuring and Defining the Experience of Immersion in Games", *International Journal of Human-Computer Studies* 66: 641-661.

Klemm, Christian & Wolter Pieters (2017). "Game mechanics and technological mediation: an ethical perspective on the effects of MMORPG's", *Ethics and Information Technology* 19(2): 81-93.

Kuo, Andrew, Richard J. Lutz & Jacob L. Hiler (2016). "Brave new World of Warcraft: a conceptual framework for active escapism", *Journal of Consumer Marketing* 33(7): 489-506.

Löhr, Guido (2018). "The experience machine and the expertise defense", *Philosophical Psychology* 32(2): 257-273.

Lowe, Dand & Joseph Stenberg (2017). "The Experience Machine Objection to Desire Satisfaction", *Journal of the American Philosophical Association* 3(2): 247-263.

Maiberg, Emanuel (2015). "Achilles' Wrists: Meet the Doctor Who Is Saving eSports Careers", *Vice*, (internet) available at: https://motherboard.vice.com/en_us/article/ypw7pm/achilles-wrists-meet-the-doctor-who-is-saving-esports-careers (accesed August 7th, 2019).

Michailidis, Lazaros, Emili Balaguer-Ballester and Xun He (2018). "Flow and immersion in video games: The aftermath of a conceptual challenge", *Frontiers in Psychology* 9: 1682.

Nozick, Robert (1974). *Anarchy, State and Utopia*. New York: Basic Books.

Nozick, Robert (1981). *Philosophical Explanations*. Cambridge, MA: Harvard University Press

Nozick, Robert (1989). *The Examined Life. Philosophical Meditations*. New York: Simon & Schuster Inc.

Pietrucha, David (2017). "Intuition and Imaginative Failure" in *Experience Machines. The Philosophy of Virtual Worlds*, edited by Marx Silcox, New York: Rowman & Littlefield, pp. 33-43.

Rollings, Andrew & Ernest Adams (2003). *Andrew Rollings and Ernest Adams on Game Design*. Indianapolis: New Riders.

Rowland, Rirchard (2017). "Our Intuitions about the Experience Machine", *Journal of Ethics and Social Philosophy* 12(1): 110-117.

Silcox, Mark & Jon Cogburn (2009). "Azeroth versus the Experience Machine" in *World of Warcraft and Philosophy. Wrath of the Philosophy King*, edited by Luke Cuddy and John Nordlinger, Chicago: Open Court, pp. 55-67.

Sigailov-Lanfranchi, Jim (2019). "Impact of Different Levels of Difficulty on Immersion in Video Games", *AICS*, 248-259.

Smith, Basil (2011). "Can We Test the Experience Machine", *Ethical Perspectives* 18: 29-51.

Sweetser, Penelope and Peta Wyeth (2005), "GameFlow: A Model for Evaluating Player Enjoyment in Games", *ACM Computers in Entertainment*, 3(3): 1-24.

Tavinor, Grant (2009). *The Art of Videogames*. Hoboken: Wiley-Blackwell

Tavinor, Grant (2017). "Welcome to the Achievement Machine: Or, How to Value and Enjoy Pointless Things" in *Experience Machines. The Philosophy of Virtual Worlds*, edited by Marx Silcox, New York: Rowman & Littlefield, pp. 99-113.

Taylor, Josh (2007). "Why Robert Nozick Should Have Played More Video Games", *Comparative Humanities Review* 1: 70-76.

van der Deijl, Willem (2019). "Is pleasure all that is good about experience", *PhilosophicalStudies* 176(7): 1769-1787.

Vitrano, Christine (2020). "Hedonism and the Good Life", *The Journal of Value Inquiry* 54: 21–40.

Warmelink, Harald, Casper Harteveld and Igor Mayer (2009). "Press enter or escape to play: deconstructing escapism in multiplayer gaming", in *Proceedings of DiGRA Breaking New Ground: Innovation in Games, Play, Practice and Theory*, pp. 1-9.

Weijers, Dan (2011). "Intuitive biases in judgments about thought experiments: the experience machine revisited", *Philosophical Writings* 50&51: 1-18.

Weijers, Dan (2014). "Nozick's experience machine is dead, long live the experience machine", *Philosophical Psychology* 27: 513-535.

WHAT SOCIETY FOR OUR POSTHUMAN FUTURE? A TASK FOR THE SERIOUS TRANSHUMANIST THINKER

LORIN GHIMAN[1]

Abstract

What is still lacking in the expanding research field of transhumanism is a serious investigation of the social existence and its role in the shaping of the future. Taking Sorgner's recent book, *We have always been cyborgs*, as a guide, the paper investigates the roots and possible reasons of the societal blindness that seems to be so prevalent in the transhumanist community. The analysis shows that the principled positions that are by and large accepted in the community, notably Sorgner's alethic and ethical nihilism, which inform the political-philosophical liberalism of most if not all of its members are untenable without an underlying historical reasoning that sees the evolution of the societies as virtually complete. Since this historical argument is always nonfactual, it results that the question about the future society can and must be addressed if transhumanism is to keep with its philosophical programme.

Keywords: *Sorgner, Stefan Lorenz, transhumanism, society, utopianism, critique.*

On the Seriousness of Post- and Transhumanist Studies

Seriousness has always been a trait of the continental philosophical tradition (Lu 2013), without which both its ambition and rigour are unthinkable. Indeed, to be accused of lacking seriousness is one of the most damning in the philosophical profession. On the other hand, being "tierisch ernst" means, in a certain German tradition, being a bit too serious, being too narrowly focused on an issue at the expense of the larger context, and has to do perhaps with either immaturity or a failure

[1] Lecturer, PhD, University of Art and Design Cluj-Napoca.
Email: <lorin.ghiman@uad.ro>.

to come to terms with one's own contingency (Dries 2014, 459; Buck-Morss 1977, 193).

Rumour has it that transhumanism (and posthumanism for that matter) is not a serious academic or even intellectual preoccupation. And this might be the main reason why they haven't become what some of their long-standing proponents hoped for (an academic field of study in their own right, that is), and might never receive the recognition and respectability philosophy enjoyed for so many centuries. I maintain however that the post- and transhumanist studies do represent a serious field of investigation, and the problem, in fact, is that the post- and transhumanist thought is still immature, or, dare I say, "tierisch ernst".

Of course, this seriousness must be demonstrated by recourse to the history of thought and the pressing matters of the present. And secondly, it must be met accordingly by serious reflection on the crucial issues – and this is the bone I've got to pick with most if not all transhumanists.

At the end of a brilliant retake of the core concepts of Enlightenment, Foucault called for a radical shift in the main philosophical programme of criticism, away from the concern with the necessary limits of the human to the investigation of their possible transgressions, pledging for an "ontology of the present and of ourselves", a critical interrogation that is meant to open up the door to the future (Foucault 1986). If this Foucauldian ontology of the present is indeed a serious philosophical proposition, then so is, mutatis mutandis, the whole spectrum of post- and transhumanist studies (Ghiman 2021). The two are further linked by their shared origin in the project of Enlightenment understood historical-philosophically as transgression (Passerin D'Entreves 1996, 12). These few lines are only gesturing in the direction of the historical link that binds contemporary post- and transhumanist fundamental problems with the core issues of modernity as a whole, a problematic that, for different reasons, is being explored more and more in the recent years (See, for instance, Hughes 2010; 2012; Saage 2013). Furthermore, I think it's fair to say that transhumanism is at its core preoccupied with our technological present and future. This is perhaps the most pressing concern for any serious attempt to understand what kind of a world we have made for ourselves. It is easy to see, then, that both in

historical-philosophical perspective and in the light of contemporary concerns, transhumanism (and posthumanism) do seem to be serious preoccupations.

One Recent Book as A Case in Point: Sorgner's *We Have Always Been Cyborgs*

I am happy to report that I am not alone in stressing the philosophical seriousness of the field, and in calling for an appropriate seriousness in addressing the task of developing an "ontology of the present and of ourselves". In his latest book, *We have always been cyborgs,* Sorgner (2022) makes a point of underlining that the core value of negative freedom, central for the transhumanist intellectual community, is one of the main achievements of the Enlightenment (Sorgner 2022, 31, 48, 59) without which the giant leap forward in the natural sciences and engineering would have been, perhaps, impossible (Sorgner 2022, 80). This insistence on the link between the Enlightenment and transhumanist thought, reductive and pro domo as it may well be, is worthy of a thorough examination; the simple fact that the emancipation project of Enlightenment is still unfinished (Honneth *et al*. 1992) warrants it. Sorgner also is of the opinion that transhumanism has a lot in common with postmodernism and the subsequent posthumanism and even advocates for "metahumanism" as an inclusive alternative to both trans- and posthumanism (Sorgner 2022, 138). He also calls repeatedly for seriousness - and not for lacking other rhetorical means.

I find myself in agreement with Sorgner on all these matters of principle, and regard his latest book as a promising starting point. Nonetheless, I am dissatisfied with it in at least one crucial point that is still not considered seriously: the question of the future outlook of human society. I would argue that this matter is symptomatically shunned not only by Sorgner, but by the entire transhumanist literature, albeit in Sorgner's case it is somewhat easier to point out the importance of it within the logic of his book.

This is why I would endeavour to develop my point through a discussion of Sorgner's recent book, *We have always been cyborgs.* It's not

going to be a book discussion, since I am primarily concerned not with what is being said in this book, but rather with what it is not said, and why it goes unsaid. I agree with Sorgner where he calls for seriousness, but maintain that his book falls short of the mark of seriousness in a crucial point: that of accounting for the social reality. I argue that this blind spot is a symptom for the whole transhumanist community. Sorgner's book however seems to be more welcoming towards the kind of criticism I would formulate here, that is not intended as a dismissal of the author's efforts, but rather as a way of orienting efforts to (again) serious efforts to tackle acute issues.

Being Cyborg Is Being Part of a Society

Sorgner wastes no time to explain why we, as humans, have always been cyborgs. Etymologically, cyborg refers back to the old *kybernaetes*, helmsman of a ship. Since each member of the *Homo sapiens sapiens* species is steered in a non-metaphorical fashion by its educators, we have always been and will be, for the foreseeable future, steered organisms (Sorgner 2022, 22). Now, of course, education eventually "upgrades" humans, imbues them with the tools of reason, and, if one remains on this side of the Enlightenment project, enables them to steer themselves and, eventually, to leave behind the need to adapt to natural environment and move towards creating an environment thought better suited for human life. This also allows for further hybridization projects to take shape. Being cyborgs, then, means that we, as individuals, need the prosthetic of society to even survive, let alone thrive.

If one admits that humans have always been steered organism in the sense described by Sorgner, one has to admit that humans are, as the old Aristotelian definition goes, "political animals", or, better said, beings that only exist as such as long as they are part of society. Unless the emancipation sought by the Enlightenment (and its presumptive heir, transhumanism) is the emancipation from society, the progress of human beings must be put in the context of society, even in the context of the progress of society as such. To put it bluntly: there is no serious

transhumanism without a critical investigation of whether and how would we come about a society for our post-human future.

Societal Blindness

It seems that any discussion about human improvement or enhancement must naturally lead to a discussion about society. But the transhumanist community skews it by resorting to "species" and "individual" instead. If indeed society is one major avenue for human improvement, then transhumanists bear the burden of explaining why it is not worth discussing it. The reader is entitled to know why society does not matter. One would expect that such an attitude comes as a result of careful reasoning and consequent decisions. If there is a thorough investigation of this issue in the transhumanist camp, I failed to notice it. Instead, I will try to put together the bits and pieces from Sorgner's book that might amount to such an argument against discussing about the future of human societies. I hope this might serve as an invitation to further work that might bring to light the ideas about society that are harboured by the members of this community.

First, I would discuss the potential consequences of Sorgner's professed alethic and ethical nihilism. Secondly, I would analyse his rejection of utopianism. And thirdly, I would look for the type of societal organization that he seems to privilege, before turning to a critique of his position.

Alethic and Ethical Nihilism Signals the End of Societal Evolution of Humanity

"Alethic nihilism" is the term Sorgner uses instead of the more common "relativism": an epistemological position that "implies that all philosophical judgements are interpretations, whereby the concept of interpretation does not imply that it has to be false, but merely that it can be false" (Sorgner 2022, 11). Ethical nihilism is a further consequence of this position in the realm of value judgements: "the ethical

perspective that no non- formal judgement of the good is plausible for all people." (Sorgner 2022, 19) For Sorgner, this position precludes "serious transhumanists" from arguing in favour of political authoritarianism and makes them naturally uphold liberal views. Indeed, it seems evident that in absence of absolute truth and universal value, we are bound to "contingent nodal points" of stability of our worldviews, convictions and beliefs that should not be imposed on others simply because sooner or later they might be rendered obsolete or flat out wrong. Negative freedom is a right borne from the humility of reason, and it must be maintained for as long as possible as a historical achievement of human societies against any type of fundamentalism as an expression of its hubris. (Sorgner 2022, 19, 21)

The history of humanity, and the evolution of human society, seems to have reached for Sorgner an endpoint in which all the struggles to impose a specific idea of good onto others on the basis of its universal relevance are exposed as being logical and epistemological fallacies: there is no common good, there is no common goal other than as a contingent nodal point subjected to change. A potential consequence of this survey of human history is that, indeed, the liberal, democratic organisation of society is simply the last logical step in the evolution of society: there is no going beyond that. All that is to do within society is to expand the boundaries of generally accepted freedoms and rights the individuals should enjoy, and the task of transhumanism is to contribute towards that end by advocating for the social and, perhaps, legal recognition of extended concepts of morphological, educational and reproductive freedoms as human rights. Since the societal becoming of humanity has reached its logical endpoint, the role of standard bearer of progress must now be assumed by the individual.

I can see no other way of understanding the societal blindness in Sorgner's book than as a consequence of such sort of argument that I formulated here. For if we are indeed "steered organisms", the only way of avoiding talk of rudder is by maintaining that, by and large, the direction has been already set for good. This would imply a subterraneous historical reasoning I would address later, as it seems to bear the full weight of the argument. The second observation that I would make here for later use is that the concept of "contingent nodal

point" is conveniently vague from a sociological perspective: it only seems to obscure the way societies work.

Utopianism as Taboo

Sorgner has a decidedly non-utopian take on transhumanism (Sorgner 2022, 62, 110, 162), and he has good reasons for that. Utopias stem from a reflection on the good life in general. Once they enter the political field of ideas they threaten pluralism and the diversity of the concepts of good his presentist relativism rests upon. For Sorgner utopias must be dangerous in principle: they are the thoughts that should not be thought in our present by a consequent alethic and ethical nihilist (compare with Sorgner 2022, 157). Utopias, then, are a mental taboo. It follows that the search for a rational society is also taboo, at least because it rests upon the premise that reason can indeed assume and complete the task of realizing such a thing that would be, in a way, less contingent.

My first observation here is that this position, as principled and consequent as it is, tends to treat society more as a natural object, as a fact, and not as an artefact, which it certainly is. (It reminds me of a phenomenon much discussed in a different philosophical tradition, called reification.)

The second observation I will return to later is that, in fact, our current societal model in the western world would not have come into existence without a certain type of rationality, and is still inhabited by one. It is something that continental social philosophers call ideology or grand narrative, and Foucauldians analyse as governmentality.

The third observation is a bit more technical (and perhaps should have come first). As Saage (2013) shows, there are marked differences between the modern utopias and the type of transhumanist techno-visions Sorgner himself discusses (and partly rejects) under the same term. They appear to rest on different anthropological premises. Modern utopianism differentiates between a factual nature of human beings, that shares information and evolves biogenetically, and an artificial, artefactual, second nature of human beings, *i.e.* the social world they

created, with its own "tradigenic" evolutionary path (Saage 2013, 3), that Sorgner would call education (compare with 2022, 85). For modern utopianism progress involves the construction of a new tradition, of a new society, leaving the natural human being and its biogenetic evolution more or less untouched. The transhumanist utopianism, however, borrowing from the sci-fi imagination and literature, extended the reach of its interventions to the biogenetic processes as well, abolishing the classical utopian distinction between the two. And this precisely is what Sorgner himself tries to substantiate when he argues at length against Habermas: that the biogenetic (genetic enhancement) and the tradigenetic (enhancement through education) are interlinked and structurally analogous. One must conclude that Sorgner actually shares the fundamental anthropological assumption at the core of transhumanist utopianism, and only argues against it when the latter calls for interventions at the social-political level. It is now clear that Sorgner's own transhumanist anti-utopianism is utopian when it comes to genetic enhancement, which would explain the head-on collision with Habermas' position, which maintains the potential of tradigenetic utopianism while refusing the biogenetic utopianism. One could ask whether a presentist relativist like Sorgner is best equipped to even enter this polemic, which touches on more profound issues: the reversibility or irreversibility of the changes each enhancement technology presupposes, and the availability of tools to change the society if need be (utopias are indeed such a tool, and treating them as a taboo does not seem appropriate if one looks beyond the current contingent nodal point of political consensus).

What Kind of Political Thinking Inhabits Sorgner's Book?

So far, I have tried to make Sorgner's text speak, against itself, about society, and I've done so by looking for the consequences of his own principled positions. I would now try and complete his views about society and politics (as the form of steering the evolution of society) by picking out what he does say about this subject matter in his latest book. I am only interested in getting a more complete and accurate form of the

missing discussion about society in his book and in the transhumanist community as a whole.

It is important to note that the core principle of social and political life is, for Sorgner, the negative freedom. He goes as far as to enlist the whole Enlightenment in its service, a hyperbole that isn't historically accurate (see, for instance, Israel 2011), but doesn't betray its meaning either. Negative freedom goes hand in hand with the alethic and ethical nihilism, which in the field of political thought might be translated as a liberal egalitarianism, since it negates the relevance of the moral value of individuals for the political or legal order. Instead of an universal common good, Sorgner uses the weak normative universalism of probable consensus: beyond all cultural and historical differences, he sees the "prolongation of human health span" (Sorgner 2022, 6, 7, 160) as a suitable candidate. This is just another facet of the marked individualism of his thought. His anti-utopianism and presentism alluding perhaps to the old discussion of the constructivist fallacy, and the meliorist convictions illustrated in the opening sections complete the picture of a rather classical liberal world view (compare with Gray 1986).

A Critique from Without

The fact that Sorgner is a liberal is hardly surprising: all transhumanists are (Sorgner 2022, 19). Neither is his stance in favour of a (still limited) internet panopticon (Sorgner 2022, 37): panopticon is simply an instrument of governmentality that fits rather well historically in the liberal paradigm (Foucault 1995). What is surprising is how little else of substance is to be found in the book beyond these liberal convictions. One would really struggle to treat all its basic assumptions as anything else than another neat "philosophical" repackaging of the dominant world view of the current historical "contingent nodal point". The argumentation strategy of the book seems to be that the current state of affairs should allow for this or that particular evolution because it doesn't seem to threaten the status quo. Sorgner's thought is hopelessly locked in the prison of the philosophical image of the contemporary social world that he himself meticulously constructed.

A Critique from Within

It might well be however that such a critique from outside would achieve little for those on the other side. This makes it as pointless as fruitless. Moreover, it will be of little help for thinkers like Sorgner, who is actively involved in a field that changes and evolves constantly. Therefore I go back to the roots of the self-inflicted societal blindness of transhumanism with the intention of showing that a serious transhumanist, an enlightened transhumanist, must and indeed can address the social world as part of a broader program of the "ontology of the present and of ourselves" that I presented as its maximalist programme.

This requires a return to the principle of alethic and ethical nihilism. As we have seen already, from this principled position, since there are no facts in a world of constant becoming, all judgements should be regarded as interpretations. This is the case even, or perhaps especially, for our judgements about the world, our philosophical judgements, including this form of perspectivism: they are all "contingent nodal points" (Sorgner 2022, 18). Now, the reconstruction of Sorgner's ideas about the social world I attempted here seems to present us with a contradiction: on the one hand our ideas about society are "contingent nodal points", subject to change just as our social reality is, but, on the other hand, it appears that the changes in the organization and structure of society must happen through means beyond our reach (Sorgner's anti-utopianism sees the modern utopianism as taboo). Society, then, is simultaneously something humans create, an artefact, a second nature, and something humans should not try and rationally transform, but leave it somehow to its own devices to evolve, as if it was nature itself. Leaving aside the fact that there is hardly an aspect of natural reality left untouched by human technologies, Sorgner is ready to accept technical (*i.e.* rational) interventions at the most profound levels of our natural being, even influencing the natural reproduction of human individuals, but regards every intervention in the mechanisms of social reproduction of human individuals as an overreach of human reason.

I can see only one way one could argue for such a position: if one declares the evolution of human society as already consumed down to

its last moment. If the human society has reached its peak evolutionary point in this particular form, then every attempt to change it further could only lead us backwards. But this ideological and, indeed, evolutionary end of history described by Fukuyama (2006) is still not a fact, but an interpretation.

The point of my critique from within of the transhumanist thought and of Sorgner's book is exactly this: the nature of the contingent nodal points and their role in the larger mechanism of society. These contingent nodal points are not that contingent: every social organisation uses a wide range of tools to prolong the 'life' and to expand the reach of some contingent nodal points that are decisive for the preservation and reproduction of the same social organization and the forms of life within it. Education is just one of these tools I mention here because Sorgner seems to understand well its importance. Every dominant power structure can be said to try and maintain a certain ephemeral interpretative system beyond its logical or historical 'due date' because it serves the interests of its own self-preservation. I think this is also an idea Sorgner would be familiar with as a diligent reader of Nietzsche: one cannot treat ideas as something separate from power relations and interests.

Every transhumanist, then, every thinker preoccupied with our present and its potentialities, and interested in furthering the cause of freedom must start anew with this simple question: *cui prodest*? This opens up a hard but inevitable discussion about the responsibility for what is being thought and said.

What Society for Our Post-Human Future?

To be sure, if there's something that we can count on it that our current world view and social organization will disappear. The utopian taboo is but a delay tactic, part of the wide-raging strategies developed in our society to ensure its reproduction. Nothing good can come out of it, only stagnation. We can, and we must ask ourselves how to build a new society. Because, again, society is something that we build, and it preserves a certain rationality, like any other artefact, in every moment of its history. There is nothing inherently wrong with the fact that

societies from all epochs have developed mechanisms to ensure their survival over generations. But everywhere, in every tradition, there were ways to think critically about it, and to think beyond it, to develop new traditions and new societies. So, the question becomes: if all thinking is linked with interests, what interests does transhumanism align with, and what interest should it serve?

Is it possible to think outside power and interest structures? Is it possible to imagine and enact a different distribution of power, or to find another, more reasonable distribution matrix of interests? How about doing away with social power entirely by designing ways to sublimate it? What governmentality would enable us to go beyond the limitations of our current human condition?

There's so much to do, there are so many questions that won't go away if one ignores them, so much to understand about a society and a world that is already changing because of us, yet already without us (Anders 2002). It is incumbent upon all serious thinkers to break their allegiances and start working for the society of the future.

REFERENCES

Anders, Günther (2002). *Die Antiquiertheit des Menschen. Band II: Über die Zerstörung des Lebens im Zeitalter der dritten industriellen Revolution*. München: C. H. Beck.

Buck-Morss, Susan (1977). *The Origin of Negative Dialectics: Theodor W. Adorno, Walter Benjamin and the Frankfurt Institute*. New York: Free Press.

Dries, Christian (2014). *Die Welt als Vernichtungslager: Eine kritische Theorie der Moderne im Anschluss an Günther Anders, Hannah Arendt und Hans Jonas*. transcript Verlag.

Foucault, Michel (1986). "Kant on Enlightenment and Revolution." Translated by C. Gordon. *Economy and Society* 15 (1): 88–96.

———. (1995). *Discipline & Punish: The Birth of the Prison*. Translated by Alan Sheridan. New York: Vintage Books.

Fukuyama, Francis (2006). *The End of History and the Last Man*. Reissue edition. New York: Free Press.

Ghiman, Lorin (2021). "Necessary Transgressions and Possible Limitations of the Human Condition. A Plea for a Critical, Posthumanist Approach to Transhumanist Issues." *Annals of the University of Bucharest - Philosophy Series* 69 (2): 103–13.

Gray, John (1986). *Liberalism*. 2nd edition. Minneapolis: University of Minnesota Press.

Honneth, Axel, Thomas McCarthy, Claus Offe, and Albrecht Wellmer (eds.) (1992). *Philosophical Interventions in the Unfinished Project of Enlightenment*. MIT Press.

Hughes, James J. (2010). "Contradictions from the Enlightenment Roots of Transhumanism." *Journal of Medicine and Philosophy* 35 (6): 622–40.

———. (2012). "The Politics of Transhumanism and the Techno-Millennial Imagination, 1626–2030." *Zygon®* 47 (4): 757–76. https://doi.org/10.1111/j.1467-9744.2012.01289.x.

Israel, Jonathan (2011). *A Revolution of the Mind: Radical Enlightenment and the Intellectual Origins of Modern Democracy*. 8.7.2011 edition. Princeton: Princeton University Press.

Lu, Mathew (2013). "Getting Serious about Seriousness: On the Meaning of Spoudaios in Aristotle's Ethics." *Proceedings of the American Catholic Philosophical Association* 87 (July): 285–93. https://doi.org/10.5840/acpaproc201441412.

Passerin D'Entreves, Maurizio (1996). "Critique and Enlightenment. Michel Foucault on 'Was ist Aufklärung?'" Barcelona.

Saage, Richard (2013). "New Man in Utopian and Transhumanist Perspective." *European Journal of Futures Research* 1 (1): 1–7. https://doi.org/10.1007/s40309-013-0014-5.

Sorgner, Stefan Lorenz (2022). *We Have Always Been Cyborgs: Digital Data, Gene Technologies, and an Ethics of Transhumanism*. 1st ed. Bristol University Press. https://doi.org/10.2307/j.ctv21wj5ng.

REVIEW

***We Have Always Been Cyborgs: Digital Data, Gene Technologies, and an Ethics of Transhumanism.* By Stefan L. Sorgner. Bristol, UK: Bristol University Press, 2021. 240 pages. ISBN: 978-1529219210. $149.95, Hardcover.**

TRANSHUMANISM GETS AN UPGRADE

NATASHA BERANEK[1]

Transhumanism affirms the use of emerging gene and cyborg technologies as a way to promote our collective survival as well as our most fulfilling individual fruition. Due to the latter, it is frequently associated with notions of perfection.

As a childhood gymnast, I became very well acquainted with its pursuit. The pointed toe, the straight leg, the split leap locked into a 180 degree angle – ostensibly gymnastics is an extended dance in distancing oneself from the fallibility of flesh and bone existence. This is not the sort of perfection that most gymnasts are pursuing though. An intensely psychological endeavor, to be a gymnast is to be engaged in an ongoing mastery of one's hopes and fears, one's impulses and drives, all while repeatedly being yanked back down into a gritty sensorium of ripped blisters and chalk dust. Under the stewardship of coaches, gymnasts track their progress – the acquisition of each new skill and the triumph

[1] Series Editor, *Anthropology and Transhumanist Philosophy*, Trivent Publishing, Budapest, Hungary. Email: <nbberanek@uclmail.net>

over every nagging fear a sort of 'upgrade' in their development. Pushing ever further against the laws of physics, gymnasts say "yes," again and again, to the pain of everyday practice. *Amor fati*, love of one's fate, is the mantra these athletes murmur as they chase fleeting moments of aerial freedom and corporeal power.

A gymnastics routine is ostensibly based upon an established dynamic between an athlete, a piece of equipment, and gravity, with the balance beam or uneven bars serving simply as a piece of technology over which the gymnast temporarily exerts control. Yet an apparatus is much more than a means to a stuck landing. When the 2020 Olympic gold medalist Nina Derwael (see Olivera 2021) explains that "the [uneven] bars speak to me, and in return I listen," she is describing a relation in flux, ceaselessly becoming anew in the time and space between. The philosopher Tobias Rees is speaking to such moments of possibility when he defines *poetry* as something that attempts to "capture that which cannot be captured" (Rees 2018, 26). Gymnasts in flight crack open and make visible our human defined order of things – an ontological landscape that has been dominant since the Enlightenment.

Forthcoming innovations in AI, synthetic biology, and cyborgization are similarly unfolding outside of our taken-for-granted understandings of human, nature, and technology. Microbiome research throws into question where microorganisms end and the individual 'self' begins. Like Simone Biles' highly publicized case of the "twisties," (see Reeve 2021) a condition that renders a gymnast completely disoriented in the air, the philosophical vertigo that these 21st century inventions and scientific discoveries induce has the potential to be debilitating, leaving us feeling rudderless, unsure of our identity and our relation to the world around us.

They also bring to the fore monumental pragmatic challenges for how we may most authentically and empathetically live amid these breakthroughs. For example, as gene technologies continue to expand, what limits, if any, should be placed upon parents who wish to 'enhance' their children's intelligence or athletic aptitude?

Stefan Lorenz Sorgner, a German philosopher based at John Cabot University in Rome, has been broadly engaged with these questions for

the past decade, and his new book, *We Have Always Been Cyborgs: Digital Data, Gene Technologies, and an Ethics of Transhumanism* tackles their intricacies head on. The title encapsulates Sorgner's fundamental appeal to his readers. In order to most smoothly navigate and benefit from the next act of the Digital Age, when these technologies will begin to enter the human body, we must realize that since the dawn of *Homo sapiens* we have been "steered" organisms (*cyborgs*) involved in a series of "upgrades," including language acquisition, education, and vaccination. Brain-computer interfaces like Elon Musk's Neuralink will simply be the next step in our long history of cyborgization – albeit one, Sorgner admits, for which our capacity to guide our own evolution is exponentiated. Dualistic modes of thinking that cast nature as 'pure' (and 'good') while dismissing technology as 'artificial' (and 'bad') are not only flawed; they unnecessarily hamper our potential to thrive.

A self-identified "weak" transhumanist, Sorgner's vision is significantly influenced by the writings of Friedrich Nietzsche – a factor, he admits, that makes his already minefield-laden philosophical trek even more volatile, given Nietzsche's unfortunate appropriation by the Third Reich and eugenicists. As it already stands, transhumanists tend to receive one-sided portrayals in the media, being cast, Sorgner says, as "cold-hearted, blood-sucking, Silicon Valley billionaires" (Sorgner 2021, 5).

Straight from the jump, Sorgner thus distances his vision of transhumanism from these more specious portrayals. He also takes pains to differentiate his weak approach from those of other notable transhumanists, such as Oxford's Nick Bostrom, who upholds a Renaissance ideal of perfection that Sorgner finds implausible, as well as a Simulation Argument (see Bostrom 2003, 243-255) that he regards as largely irrelevant. Rather than focusing on more fantastical forms of "silicon-based" humanity, such as mind uploading, his concept of the good life is one that stresses a "carbon-based" practical relevance and fosters a greater plurality of avenues towards human flourishing.

Even so, as a "positive pessimist," Sorgner aligns himself with most transhumanists' optimism about technological innovation. Yes, our lives are full of struggle, and with moments of pleasure (espressos, sexual euphoria, met work deadlines) too brief and too few and

far between. Considering how much vaccination, antibiotics, and anesthetics -– all recent developments in the grand scheme of things – have improved our lives though, we should be hopeful that newly emerging technologies will continue to alleviate our suffering.

How receptive one is to Sorgner's proposals is likely to be a reflection of how well one perceives liberalism to be currently working as a dominant political model. This cannot be overstated. In *We Have Always Been Cyborgs*, underlying each of Sorgner's proposals about parenting, education, morphological modification, sexuality, and reproductive technologies is his allegiance to the norm of negative freedom, a core Enlightenment value that aims to safeguard personal freedom from outside interference. "I am not claiming that it is always morally and legally acceptable to act in accord with one's own authentic wish," he writes. "A person's freedom ends where the freedom of another person begins" (Sorgner 2021, 137). Sorgner's ethical nihilism would not condemn an adult brother and sister who mutually consent to sexual relations, but it would strongly reject the actions of a pedophile, who in pursuing his or her idiosyncratic drives, is harming others. Certainly this is a stark example. Anthropologists have long observed incest – of varying degrees of closeness – as a universal cultural taboo. Yet Sorgner's view of negative freedom would override this taboo, rendering it paternalistic and violent.

With all of this 'freedom' talk in mind, I wouldn't be surprised if the eyebrows of readers who are new to Sorgner will raise when they arrive at *Cyborg*'s section on digital privacy. I can hear people asking, "How is it not an utter contradiction in terms for Sorgner to cherish a radically pluralistic concept of the good on the one hand, while on the other concluding that we must inevitably accept total surveillance via RFID chipping? How could we ever truly safeguard our freedom, our individual power, if there's no longer any privacy? If our biometric data is constantly being entered into an Internet of Bodily Things?"

Key to understanding Sorgner's consent to an Internet Panopticon is his perspective that it is not freedom and *privacy* that are synonymous, but freedom and *health*. Aging, he writes, is the worst mass murderer in the world. In order to prolong our healthspan – the number of *vital* (not total) years we enjoy on this planet – we must take full advantage of the

scientific knowledge and medical innovations that will arise from the collection of a wide range of our personal data. This is because most human beings identify an increase of the healthspan with a higher likelihood of living a good life. The predictive maintenance of our health will inevitably be improved the more data is collected concerning correlations between aging, genes, and lifestyle choices. Aging should be approached as disease rather than an undoable process, but unlike many Silicon Valley transhumanists who strive for immortality via cryonics or mind-uploading, Sorgner's pragmatism leads him to conclude that our entrance into Big Data is our most realistic chance to promote the flourishing of all individuals.

It is a reluctant conclusion on Sorgner's part. "I wish to stress very much that my analysis is not one about which I am happy, as I am aware that the risks and dangers for a liberal system are enormous" (Sorgner 2021, 38). Yet most of us are only one or two clicks away from an abundance of social media evidence that indicates we don't cherish our privacy as much as we like to think, and so Sorgner is fairly convincing in his conclusion that what we actually fear about total surveillance is not a loss of privacy, but that we might be *unfairly sanctioned* for our genetics, disease states, personal preferences, or behaviors.

All the more reason, he argues, that the norm of negative freedom must be culturally promoted and legally protected. In an Internet Panopticon, no one holds a privileged privacy position; there are no guardians and prisoners as such. "Norms and values are just as much fictions as money," (Sorgner 2021, 41) writes Sorgner, so it is left up to us to truly *believe* in the value of radical plurality as our digital privacy fades. On a geopolitical scale too, in order to keep pace technologically with China, it will be necessary for us to develop a democratic method for collecting and using Big Data. Not only does the Chinese government already collect data in massive quantities, but its citizens are more amenable to this system given their relatively collectivist orientation towards subsuming the individual to the common good.

If the most promising means for promoting human flourishing are to be found within the carbon-based realm, then how can we most responsibly engage? Here the Nietzschean flavor of Sorgner's transhumanism fully emerges. Nietzsche, like Darwin, viewed humans

as being gradually (rather than categorically) different from non-human animals. Their naturalist and evolutionary perspective broke with a metaphysical tradition that poised humans as part immaterial soul, part material body. Despite this 19th century rupture, a dualistic metaphysical understanding of personhood has persisted in our social and legal systems, as has a Christian one. In Sorgner's view, this is an *immoral* state of affairs because it impinges on the degree to which we can make autonomous choices about our use of biotechnologies.

In particular he directs his discussion towards the decision making of parents or parents-to-be. For example, if it is metaphysically unclear whether a fertilized egg has a special status or not, then it should be up to the specific parents who are using pre-implantation testing (PGD) and in-vitro fertilization, rather than the state, to determine its moral status. Where there is disagreement on reproductive technologies, Sorgner writes, legalization should occur in favor of more freedom.

Bioethical debates about non-human personhood and what counts as harm remain strongly enmeshed in Kantian metaphysics, adding kindling to what many are likely to view as Sorgner's provocative conclusion that the person-object dichotomy simply cannot be upheld. Contra Jürgen Habermas, one of the most prominent German philosophers of the 20th century, Sorgner does not regard parents who would genetically enhance their children as threatening their children's autonomy or as immorally 'instrumentalizing' them. As in many other places throughout *Cyborgs*, Sorgner supports his perspective by drawing a structural and moral analogy between parental education (*i.e*. child rearing) and genetic enhancement.

> "In both cases," he writes, "decisions are being made by parents concerning the development of their child, at a stage where the child cannot yet decide […] Parents usually love their children and want them to have the best possible starting points in life. Of course, parental decisions do not always produce good results. But, as a rule of thumb, parental influence most often leads to better outcomes than those from chance or without any guidance." (Sorgner 2021, 86)

Sorgner counters Habermas's other concerns about the genetic modification of children -– foremost among them, its seeming irreversibility and embodied expression– by reference to cutting edge research in the fields of epigenetics, siRNA therapy, and gene silencing.

In the wrap up to *Cyborgs*, Sorgner resolves any lingering doubts that readers may have about the details of his transhumanist alignment with Nietzsche. There are certain aspects of Nietzsche's thinking that Sorgner regards as immoral, such as his vision of a two-class, master and slave society. Such a hierarchical structure is deeply illiberal, and Sorgner cherishes liberalism above all else.

Yet Nietzsche's basic impulses, based as they were upon self-discovery, truthfulness, and the will-to-power, are those that Sorgner finds indispensable to a radically pluralistic understanding of perfection, and he offers Nietzsche's theory of the Eternal Recurrence as providing a possible framework for life's meaning. "Someone lives a good life by following their very own idiosyncratic psychophysiological demands, their very own desires, passions and fantasies…. [But] to become aware of one's very own drive is much more difficult than is often believed" (Sorgner 2021, 166)[2]. Plumbing one's depths is a personal responsibility, and it is often painful. Yet it is the avenue by which we can experience "that special moment, which is worth all the suffering [we] have to endure, as [we] know that this one special moment will recur again and again." (Sorgner 2021, 178)

Transhumanism is not a religion, despite the focus that many of its retinue place on ethereal aims like immortality, disembodied consciousness, and perfection. Nor can transhumanism be accurately glossed as an ideology, given its diversity of opinions on possible 'beyond human' outcomes. *We Have Always Been Cyborgs* instead renders transhumanism much more terrestrial, demonstrating that it is better understood as a philosophically informed *positive attitude* towards the use of technologies. It is interested in providing innovative options for our existence rather than introducing dystopian constraints on our individuality. Sorgner's book encourages us to realize that before we can avail ourselves of emerging technologies, we need to do the work of questioning the boundaries between human, nature, and technology. If done well, such an endeavor will inevitably lead to even more provocative questions: What do these categories even mean? How and

[2] See also Sorgner 2016, 141–157.

why did we come to take them for granted? And what makes us feel that we need to cling to them so tightly?

All around us are fissures containing new ways of thinking and radically different options for living and being in the world. Like gymnasts in flight, when we break these fissures open, we create the poetry of unanticipated liberation. And that too, is perfection.

REFERENCES

Bostrom, Nick (2003). "Are You Living in a Computer Simulation?" In *Philosophical Quarterly* 53 (211): 243-255. Available at <https://www.simulation-argument.com/simulation.html>, last time accessed in

Olivera, Lieve J. (2021). "It Feels as If the Uneven Bars Speak to Me, and In Return, I Listen." In *International Gymnast Online,* March 10, 2021. Available at <https://www.intlgymnast.com/interviews/nina-derwael-it-feels-as-if-the-uneven-bars-speak-to-me-and-in-return-i-listen/>, last time accessed in

Rees, Tobias (2018). *After Ethnos.* Durham: Duke University Press.

Reeve, Elle (2021). "Simone Biles and the 'Twisties': How Fear Affects the Mental Health and Physical Safety of Gymnasts," *CNN*, July 29, 2021. Available at <https://www.cnn.com/2021/07/28/us/simone-biles-olympics-gymnastics-physical-mental-health/index.html.>, last time accessed in

Sorgner, Stefan Lorenz (2016). "Three Transhumanist Concepts of Human Perfection." In *Perfecting Human Futures: Technology, Secularization and Eschatology*, B. Hurlbut, H. Tirosh-Samuelson (eds.). Wiesbaden: Springer, 141–157

Sorgner, Stefan Lorenz (2021). *We Have Always Been Cyborgs: Digital Data, Gene Technologies, and an Ethics of Transhumanism.* Bristol: Bristol University Press.

NOTES FOR CONTRIBUTORS

THE ANNALS OF THE UNIVERSITY OF BUCHAREST
PHILOSOPHY SERIES

The authors of the articles and book reviews are requested to observe the following publication guidelines:

- ✓ The articles should be edited in English.
- ✓ The articles should be submitted electronically (by e-mail or CD) in a WORD format (formats .doc or .rtf).
- ✓ The articles should contain the author's full name and affiliation, along with the author's e-mail address. The authors are requested to supply an auto-bio-bibliography (approximately 10 lines), in a footnote.
- ✓ The articles should contain an abstract (10-15 lines), followed by 5-7 Keywords (Palatino Linotype, 9, single spaced).
- ✓ All the articles and book reviews must be edited using diacritical marks; if there are special Fonts, these should also be sent.
- ✓ The page format: paper A4 (no Letter, Executive, A5 etc.).
- ✓ The page margins: top – 5,75 cm; bottom – 5 cm; left and right – 4,25 cm; header – 4,75 cm; footer – 1,25 cm.
- ✓ The articles submitted for publication must be typed single spaced, in Palatino Linotype, 11.
- ✓ ***The title*** of the article should be centered, bold, all capitals (Palatino Linotype, 11).
- ✓ ***The author's name*** (bold capitals) should be centered, under the title (Palatino Linotype, 9).
- ✓ ***The abstract*** (with the translated title, if the article is written in other language than English; Palatino Linotype 9, single spaced) precedes the text of the article; the Keywords (Palatino Linotype, 9,

bold) follow the abstract and they are preceded by the word ***Keywords*** (in italics, bold).

- ✓ ***The notes*** should be indicated by superscript numbers in the text and typed at the bottom of the page (single spaced, Palatino Linotype 9).
- ✓ ***The references*** or ***the quotations sources*** should be indicated in the text, following the format: (Author year:(space)page) – (Pop 2001: 32); (Pop/Ionescu 2001: 32).
- ✓ ***The abbreviations*** or ***abbreviated titles*** (RRL, tome L, nos 3-4, p. 216) can be used in the papers; they will be included completely in the listed references at the end of the article, as it follows:

RRL – *Revue Roumaine de Linguistique*, tome L, nos 3-4, 2005.

- ✓ ***The references*** should observe the following styles:

1. Books *Basic Format:* Author, A.A. (, B.B. Author, C.C. Author), Year of publication, *Title of Work,* Location, Publisher.

Kleiber, Georges, 2001, *L'anaphore associative*, Paris, Presses Universitaires de France.

2. Edited Books *Basic Format:* Author, A.A. (, B.B. Author, C.C. Author) (ed./eds.), Year of publication, *Title of Work,* Location, Publisher (only the name of the first editor inverted).

Zafiu, R., C. Stan, Al. Nicolae (eds.), 2007, *Studii lingvistice. Omagiu profesoarei Gabriela Pană Dindelegan, la aniversare*, Bucureşti, Editura Universităţii din Bucureşti.

3. Articles or Chapters in Edited Book *Basic Format:*

Rand Hoare, Michael, 2009, "Scientific and Technical Dictionnaries", in A.P. Cowie (ed.), *The Oxford History of English Lexicography*, Oxford, Oxford University Press, pp. 47-94.

4. ***Articles in Journals*** *Basic Format:* Author, A.A. (, B.B. Author), Year of publication, "Title of the article", in *Title of Periodical*, volume number (issue number), pages.

Fischer, I., 1968, « Remarques sur le traitement de la diphtongue *au* en latin vulgaire », in *Revue Roumaine de Linguistique*, XIII, nr. 5, pp. 417-420.

All the bibliographical references should appear in the final bibliography.

All the papers will be peer-reviewed by a committee of specialists in different philological fields: linguistics, literature, cultural studies, translation studies.

The first version of the articles should be submitted to the e-mail address: annals.philosophy@ub-filosofie.ro

Tiparul s-a executat la
Tipografia Editurii Universităţii din Bucureşti