

Iker Nabaskues Martínez de Eulate*

On Transhumanism. A Socio-legal Approach beyond T-800 and The Replicant with Reference to Arendt and Aristotle

Über Transhumanismus. Ein rechtssoziologischer Ansatz jenseits von T-800 und The Replicant unter Bezugnahme auf Arendt und Aristoteles

<https://doi.org/10.1515/zfrs-2025-2010>

Abstract: Hannah Arendt, with her teleological concept of human action, places the individual at the centre of politics, in the sense that politics can only be the result of bringing into play all his or her capacities. For Aristotle, the human condition has three attributes: the animal, the social and the spiritual. The debate on transhumanism cannot be limited exclusively to the possibilities of improvement that such a philosophical movement can bring to society. It has deeper implications that require taking into consideration the very nature of the human condition as developed by Arendt and the imperishable Stagirite. The T-800 in the film “Terminator” (USA 1984) and the Replicant in “Blade Runner” (USA 1982) are both expression of the projection of human anxiety about the effects of the relationship between culture and technology as well as antithetical expressions of a possible future evolution of the android. Only a clear understanding of what constitutes human nature through reflection on its most unique capabilities will enlighten us in judging whether the enhancements of transhumanism can be considered beneficial or dangerous in the future.

Zusammenfassung: Hannah Arendt stellt mit ihrem teleologischen Konzept menschlichen Handelns den Einzelnen in den Mittelpunkt der Politik, in dem Sinne, dass Politik nur das Ergebnis der Einbeziehung aller seiner Fähigkeiten sein kann. Für Aristoteles hat der menschliche Zustand drei Attribute: das Tierische, das Soziale und das Spirituelle. Die Debatte über den Transhumanismus kann nicht ausschließlich auf die Verbesserungsmöglichkeiten beschränkt werden, die eine solche philosophische Bewegung für die Gesellschaft bringen kann. Sie hat tiefere

***Korrespondenzautor:** Iker Nabaskues Martínez de Eulate, University of The Basque Country, Faculty of Law, Paseo Manuel de Lardizábal, 2, 20018, Donostia-San Sebastián, Gipuzkoa, Spain, E-Mail: iker.nabaskues@ehu.eus

Implikationen, die eine Berücksichtigung der Natur des menschlichen Zustands erfordern, wie er von Arendt und Stagirite entwickelt wurde. Der T-800 im Film „Terminator“ (USA 1984) und der Replikant in „Blade Runner“ (USA 1982) sind beide Ausdruck der Projektion menschlicher Angst vor den Auswirkungen des zukünftigen Verhältnisses von Kultur und Technologie, vor der möglichen zukünftigen Entwicklung des Androiden. Nur ein klares Verständnis dessen, was die menschliche Natur ausmacht, durch Reflexion über ihre einzigartigen Fähigkeiten, wird uns bei der Beurteilung darüber helfen, ob die Verheißungen des Transhumanismus als vorteilhaft oder gefährlich angesehen werden müssen.

Keywords: Law and Popular Culture, Hannah Arendt, Aristotle, transhumanism, human nature.

Just as the Internet once was, robotics and artificial intelligence systems are the transformative technology of our time. They are becoming part of everyday life, leading to substantial changes in society: robotic mechanisms are replacing a large number of jobs and this is expected to increase in the future. The war in Ukraine itself is being fought in the air through the use of war drones, which are carrying out missions that were once carried out by soldiers. The same is happening in the domestic sphere where houses are evolving into interconnected and intelligent homes. The profound transformations of artificial intelligence have led to the germination of the first EU legislation to regulate it. It is essential for the law to deal with this disruptive technology to ensure respect for and compliance with fundamental rights.

Artificial intelligence refers to different systems that display intelligent behaviour. AI-based systems can be purely software-based, acting in the virtual world (voice assistants, image analysis software, search engines, speech and face recognition systems) or AI can be embedded in hardware devices (advanced robots, autonomous cars, drones or Internet of Things applications) (European Commission 2018: 2). In this case, intelligence is understood as “the ability to pursue goals, plan, foresee consequences of actions and use tools to achieve goals” (Cortina 2022: 6). That is, the ability to solve problems with the use of tools. To give a more concrete definition of AI, the High-level experts Group on Artificial Intelligence created by the European Commission in June 2018, defines AI systems as: “systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals” (European Commission 2018: 2).

Artificial intelligence presents two major types of problems. Firstly, the problem of identifying the limits between the therapeutic or merely curative application of the new emerging technologies, and their use for the betterment of the individual.

Secondly, the extent to which these innovations may affect the principle of equality and distributive justice, which based in principle on an optimistic vision of liberation marks the path of human self-destruction (Campione 2019: 57).

Transhumanism is an international cultural and intellectual movement whose ultimate goal is to transform the human condition through the development and manufacture of widely available technologies that enhance human capabilities, whether physical, psychological, or intellectual. Transhumanist thinkers have in common the study of the potential benefits and dangers of new technologies that could overcome fundamental human limitations, as well as the appropriate techno-ethics of developing and using such technologies. They speculate that human beings may become capable of transforming themselves into beings with extensive capabilities. As we shall see, some of the transhumanist thinkers believe that the phenomenon will contribute to the shaping of an enhanced human person. Others consider that a new human subject will emerge, for which it would be appropriate to use the term post-human. Still others believe that transhumanism poses dangers to the human condition itself.¹

AI – including automatic profiling, decision-making and other machine learning technologies – affects the right to privacy and other rights, including those relating to health, education, freedom of movement, freedom of peaceful assembly and association, and freedom of expression. AI tools are widely used to seek insights into patterns of human behaviour. With access to the right data sets, it is possible to draw conclusions about how many people in a particular neighbourhood are likely to attend a certain place of worship, what television shows they may prefer and even roughly what time they tend to wake up and go to sleep. AI tools can make far-reaching inferences about individuals, including about their mental and physical condition, and can enable the identification of groups, such as people with particular political or personal leanings. AI is also used to assess the likelihood of future behaviour or events. AI-made inferences and predictions, despite their probabilistic nature, can be the basis for decisions affecting people's rights, at times in a fully automated way. Many inferences and predictions deeply affect the enjoyment of the right to privacy, including people's autonomy and their right to establish details

¹ The contemporary meaning of the term transhumanism was forged by one of the earliest professors of futurology, Fereidoun M. Esfandiary, known as FM-2030, who thought of “new concepts of the human” in The New School around 1960, when he began to identify people who adopt transitional to “posthuman” technologies, lifestyles, and worldviews as “transhumans”. This hypothesis would be supported by the work of the American philosopher Max More, who would begin to articulate the principles of transhumanism as a futurist philosophy in 1990, and to organize an intellectual group in California that has since grown into what is now called the international transhumanist movement.

of their identity. They also raise many questions concerning other rights, such as the rights to freedom of thought and of opinion, the right to freedom of expression, and the right to a fair trial and related rights (UN 2021: 5). The privacy, autonomy and responsibility impact of AI, three basic legal concepts for law, is therefore a fundamental issue.

There are two main models of transhumanism: techno-scientific or cybernetic transhumanism and biological transhumanism. The former takes its inspiration from artificial intelligence, software engineering and robotics, seeking the hybridisation of person and machine. The technoscientific transhumanism seeks to techno-manufacture a post-humanity, a new hybridised species with machines endowed with physical capabilities and an AI superior to the human one. The biological transhumanism sets less ambitious goals, since it advocates human bio-improvement based on new advances and discoveries in the fields of biology, medicine, pharmacology and genetics (Cortina 2022: 5).

Authors such as Nick Bostrom, one of the representatives of the technologist transhumanism, goes back to classical authors to show that since ancient times, there was already a conception of human nature, from which the human being would be understood as an infinitely “plastic being” (Ortiz de Zárate 2020: 100). Modern man would be characterised by his capacity for self-moulding, for being his own artisan and creator. This notion of man as a self-moulding being would be the fundamental assumption on which technological transhumanism would be based. From this conception of the human condition, transhumanism derives what has become one of its most famous slogans, that “there is nothing wrong with manipulating nature” and “there is nothing to be ashamed of” in this respect (Bostrom 2005: 99).

For this conception of transhumanism, there would be no essence in human nature. Thus, it is an anti-essentialist movement according to which there is nothing that we could call “human nature” worth preserving. Therefore, the human being is understood as a whole susceptible to any manipulation or improvement as long as it is accepted and consented to by oneself by virtue of the right of free choice that each individual possesses (Ortiz de Zárate 2020: 101). For technological transhumanism, there is nothing worthy of the name “essence” or “human nature” that prevents us from modelling and improving the human being at will, given that the most important attribute of man is freedom of choice. We are free to choose our destiny, our own bodily form and to give ourselves our own law.

On the contrary, on the philosophical-political level, the group of thinkers who oppose transhumanist doctrines, called bio-conservatives, include, among many others, Francis Fukuyama. Fukuyama identifies as the most significant threat of contemporary biotechnology the possibility that it may alter human nature and thus lead to a “posthuman stage of history” (2002: 32). He defends the relevance of

humanity as a valid term that has “provided a stable continuity to our experience as a species” and “along with religion, what defines our most basic values” (Fukuyama 2002: 60). Fukuyama warns of the danger that emerging technologies will eventually make us lose our humanity, our essence.

Other bio-conservatives defend the morality of gratitude and humility, which are based on recognising the character of the “gift of human achievements and capabilities” (Campione 2019: 56) against the transhumanist ethics of perfectionism that “forgets that freedom consists in a certain sense in a permanent negotiation of what has been received” (Sandel 2007: 127). Habermas, for his part, presents objections to transhumanism, because of “the imprecision of the boundaries between the nature we are and the organic endowment we give ourselves” (2002: 35), exposing the unbridgeable distances between the debate on transhumanist proposals and the practical issues they entail. The various representations of human enhancement throughout history allow us to philosophically discern the field we address. Popular culture is full of expressions and symbols of this idea of human enhancement.

Human Enhancement, Androids and Cyborgs in Popular Culture

Prometheus is one of the Titans, the supreme trickster, and God of fire in the Greek religion. His intellectual side is emphasised by his name, “Forethinker” (Britannica 2024). The myth says that he developed into a master craftsman, and in this connection, he was associated with fire and the creation of mortals. The myth of Prometheus, who steals fire from the gods, inspired artists and writers throughout history to refer to the audacity of men to do or possess divine things; romantics saw in him a prototype, the one who tries to push the limits.

Prometheus had a diverse treatment in Western culture (Sasani & Pilevar 2016). It was seen firstly, as a beneficent figure, who makes possible the progress of humanity and tries to equalise man with gods; secondly, as the romantic prototype of the rebel who defies gods and nature; finally, as a disastrous figure, since knowledge, science and technology break humans’ innocence and cause disasters and suffering.

Mary Shelley’s lovely character, Victor Frankenstein, who made his first appearance in her 1818 novel, is a scientist whose creature rises up against his creator. This is the representation of the punishment that derives from the irresponsible use of technology (Keese 2011: 16). “The Creature,” as he is called in the work, would later become the monster called just “Frankenstein”, giving The Creature a more recognisable name as well as the surname of its creator. Frankenstein – both the doctor

and his creation – have appeared so many times in films, television shows, and video games that they have long since become part of the universe of our popular culture.

While many later adaptations of Frankenstein portrayed Victor Frankenstein as the prototypical “mad scientist,” Shelley’s original novel depicted him as a man tragically driven by ambition and scientific curiosity, unable to deal with the consequences of his actions in pretending to be God. Later, some decades after, mostly since the beginning of the seventies of the past century, the character was portrayed as a ruthless sociopath. And, over the last few decades, the characterisation has shifted away, linking him to a more complex character in relation to technology.²

The Romantics believed in the capacity of the mind to recreate; thus, the individual’s mind acts as the Almighty God. Prometheus’ quality of rebellion and liberty was appreciated and welcomed by many Romantic authors. This attribute was embedded everywhere in 18th century Romantic literature owing to a great change happening in Europe and America, which led to the breaking down of the dominant rules and tyrannies (Sasani & Pilevar 2016). But, despite living among the Romantics, Mary Shelley did not show any special political affiliation with freedom and the limitless ability of human beings. On the contrary, she used the character of the monster to decry the romantic-related ideas of human beings’ limitless capabilities (Sasani & Pilevar 2016).

The change in the way Dr. Victor Frankenstein has been depicted in popular culture over the last few decades may have to do with the way today’s society looks at human enhancement projects. This change may be understood as an expression of today’s anxiety linked to the technology-humanism duality (Serra 2022: 171). We can frame the portrait of science as a danger for the human condition of the past decades, in the bio-conservative approach. Today, our culture seems more interested in the questions and challenges that arise in the minds of these prototypes. Why are many conservatives convinced that social engineering almost invariably leads to backfiring? One answer they might give is an inductive one. In many past instances, it has been observed that attempts to socially engineer society have led to suboptimal outcomes (Kayali & Clarck 2020: 247)

A key tenet of political conservatism is the view that human nature imposes severe constraints on the very possibility of deliberately improving human society (Kekes 1998: 41). Humans suffer from severe cognitive and affective limitations and so are not capable of deliberately improving the complicated societies they inhabit

2 It would be impossible to list all the films in which Frankenstein’s characters, both the doctor and the beast, appear in this work. I will only refer to some films that can be considered referential in the history of cinema, such as: “Frankenstein” (USA 1931); “The Ghost of Frankenstein” (USA 1942); “The Curse Of Frankenstein” (UK 1957); “Frankenstein Must Be Destroyed” (UK 1969); “The Bride” (UK 1985); “Frankenstein” (UK 1992); “Frankenstein” (USA 2004); “Frankenstein” (USA 2015).

(Kayali & Clarck 2020: 248). The conservatives who rely on this objection place themselves in an “uncomfortable position” (Buchanan 2011: 9). Humans have sufficient cognitive power to know that they have severe and permanent cognitive limitations and to appreciate that new inventions embody superior knowledge, but they lack cognitive power to overcome these limitations (Buchanan 2011: 150).

Certainly, one of the potentialities of human beings has been to anticipate through imagination the technological changes in the future. There is an unquestionable nexus between science fiction and reality. Both the idea of looking to the future and the possibility of using fiction to do that, are relatively new in history. A fundamental change in human thinking about the future began in the 18th century, as technological change accelerated to a point where its effects were easily visible in the course of a single lifetime, and terms such as progress and development entered human discourse. Speculation about the future became more common as human beings increasingly reshaped the world during the 19th and early 20th centuries, though it was seen largely as entertainment, a diversion from the often stark realities of everyday life. Yet, some of that speculation proved surprisingly close to the mark (Rejeski & Olson 2006: 17).

This was the period that gave birth to the earliest examples of what contemporary readers might yet recognise as science fiction. Edgar Allan Poe may well be called the father of “scientifiction”. It was he who really originated the romance, cleverly weaving a scientific thread into and around the story. One of the most representative writers of science fiction literature as an anticipator of a future to come was Jules Verne. Many of the inventions novelised by Verne in his narrations became reality some decades later. A little later came H. G. Wells, whose scientifiction stories, like those of his forerunners, have become famous and immortal (Westfahl 2007: 50).

It was Wells who advanced what is now a vibrant literary tradition of predicting the onward march of technology. Wells’s vision of future technology is rich. Wells imagined technological developments that altered the physical landscape. By 2100, people were concentrated in huge cities (the projected population of London is thirty-three million) that are walled, not against any external threat but rather as a convenient means of controlling the weather ... Wells also anticipated television, the videocassette, and powered commercial and combat aircraft (De Canio 1994: 77).

In Verne’s novels, the stories are packed with world exploration and mystical monsters that Verne explains and hypothesises abundantly in typical Victorian 19th century, science-loving fashion. Verne is not the only science fiction writer to envision inventions and cultural changes long before they became a part of everyday life. Aldous Huxley predicted antidepressants in his novel *Brave New World* (1931), George Orwell predicted widespread governmental surveillance in his novel *1984* (1949), and John Brunner, school shootings, electric cars and hookup culture in his work *Stand on Zanzibar* (1969) (Kerr 2020).

The combination of technological innovation in industrial society can be transferred to another author and another technological period. The American writer Philip K. Dick can be considered a reference author who portrays in his works the political, socio-legal and even metaphysical consequences of a society involved in the evolution of artificial intelligence in relation to the transhumanist debate.

Philip K. Dick's work was replete with whimsical and absurdist presentations of the greatest challenges to reason and to humanity – paradox, futility, paranoia, and failure – they have been the inspiration for several well-known and thought-provoking films such as “Blade Runner” (USA 1982), based on the novel *Do Androids Dream of Electric Sheep?* (1968), “Total Recall” (USA 1990), “Minority Report” (USA 2002), “The Adjustment Bureau” (USA 2011) and “The Man in The High Castle” (USA 2015). Other films that deal with the implications of technology on the human condition and where a certain influence of K. Dick's works can be seen are “A Scanner Darkly” (USA 2006), “Imposter” (USA 2001), “Next” (USA 2007), “Screamers” (Canada 1995), “Paycheck” (USA 2003), “Matrix” (USA 1999) and “Inception” (USA 2010) (Sullins 2011: 10).

The T-800 and The Replicant: Two Models of Technological Development

The Replicant and the T-800 – The Terminator – represent two antithetical points about the limits, implications and development of artificial intelligence. The Terminator, also known as a Cyberdyne Systems Model 101 or the T-800, is the name of several film characters portrayed by Arnold Schwarzenberger. The Terminator is a formidable robotic assassin and soldier, designed by the military supercomputer Skynet for infiltration and combat duty, towards the ultimate goal of exterminating the Human Resistance (Khouw 2020: 11).³

The Terminator can be seen as a metaphor for humans trying to regain control from technology. This does not imply a prediction for the future but definitely does set a premise for how human beings feel about technology: anxiety. The term “Technological pessimism” refers to the sense of disappointment, anxiety, even menace, that the idea of “technology” arouses in many people these days. There is some-

³ When we refer to the Terminator in this article, we are referring to the first sequel's Terminator. The second shows the same T-800 model, which has evolved towards political commitment to good and the fight against machine totalitarianism. It is a model that feels and empathizes with human beings. Therefore, it is important to note that we are referring to Cameron's T-800 of “Terminator” (USA 1984) and not to the one of the second sequel (1991).

thing paradoxical about the implication that technology is somehow responsible for today's widespread social pessimism. The modern era, after all, has been marked by a series of "spectacular scientific and technological breakthroughs" (Marx 1994: 15). We are reminded of the astonishing technical innovations of the last century in, say, medicine, chemistry, aviation, electronics, atomic energy, space exploration, or genetic engineering.

But, above all, there is one area where this anxiety can be exponentially increased: the armaments and warfare sphere. If World War II meant motorisation, which surpassed the logic of the trenches of the Great War, the warfare of the 21st century is taking place in the digital realm. We are witnessing this in the Ukrainian War, where remote-controlled drones are proving to be a decisive factor in carrying out surprise attacks. The drones used in Ukraine, coupled with an effective data processing system, pose new military paradigms that force strategists to adapt their knowledge to the use of artificial intelligence. Long-range sensor systems and precision weapons are being employed to attack the enemy. Drones are an integral element of both Russian and Ukrainian reconnaissance and strike complexes, providing an enormous amount of data and allowing commanders to identify and prioritize targets more efficiently. The massive use of drones has signalled an evolution in the nature of combat (Cropsey 2024).

The T-800 possesses all human-enhanced capabilities. It can be more effective at eliminating enemies than an entire division of soldiers, and in this sense, the Terminator's message is both powerful and disturbing. The T-800 may be the embodiment of science at the service of the war machine as an indispensable element of power. There is no doubt that in this sense, the message of the Terminator saga is philosophically bleak. The project of human improvement has at its point of origin the question of how to eliminate human life faster and in a more effective way. This would be the negative version of technology's potential to bring about a more just, liveable and healthy world.

The Replicant is a fictional humanoid featured in the 1982 film "Blade Runner" and the 2017 sequel "Blade Runner 2049" (USA) which is physically indistinguishable from an adult human and possesses superhuman strength and intelligence. "Blade Runner" became an influential popular and cultural icon of the eighties. What is striking about the film, apart from elements that encourages deep philosophical questions, is the way in which it confronts and develops the question of the content and essence of humanity.

Blade Runner's Replicant represents the inverse of James Cameron's Terminator. The Replicant is a synthetic robot, which nevertheless aspires to a human life (Shanahan 2020). Life and death of Replicants is scheduled by humans so they know when they will die. The story shows the existential crisis that this consciousness produces in the Replicant. The philosophical approach is then paradoxical. A syn-

thetic android questions his existence as a human. He does not want to die. Replicants do feel emotions, and they have memories. Does that make them human? The replicants share enough qualities with humans that they deserve protection. “It’s a very strong case for treating – a non-human – with the same legal rights we give a human. We wouldn’t call – the Replicant Rachel – a human, but maybe a person,” (Boissoneault 2017:12).

On Transhumanism, Arendt and Aristotle

Both “Terminator” and “Blade Runner” are expressions of the hybridisation of the human body and the machine. In this sense, they are a representation that anticipated the transhumanism that was to come in the 21st century. According to Bostrom, transhumanism is a tool to end unwanted and unnecessary aspects of the human condition, such as suffering, disease, the effects of ageing and even mortality (Bostrom 2005). The transhuman individual is already here (Rejeski & Olson 2006). Technology that looks for human capabilities’ enhancement has reached the point where it has been brought about by advances in artificial intelligence. Today, transhumanism represents a challenge for bioethics, with profound anthropological and ethical implications. In the face of transhumanist proposals, it would be good to avoid both technophobia and uncritical technophilia (Martínez-Córcoles et al. 2017: 23). Modern technologies generate, in the same measure, comfort and disasters. At the psycho-dynamic level, this ambivalence is expressed by technophilia (attraction to technology) and technophobia (rejection of technology). Technophilia and technophobia are the two extremes of the relationship between technology and the human being, but especially, between technology and society (Osiceanu 2015: 1139). We should neither accept nor reject anthropotechnical projects per se but rather value them one by one. To do this, of course, we need criterions.

If we want to discern the scope and possibilities of transhumanism to promote a better society, we must take into consideration some aspects of human nature that seem elementary to us. Two giants of philosophy like Arendt and Aristotle help us in this. In relation to Arendt, I will use the concept of “human action” in her political theory and in relation to Aristotle, I will use his concept of “human nature” in his philosophical thought.

Hannah Arendt used “human action” as the key concept of her theoretical approach (Patierno & Crisorio: 2016). According to Arendt, action is the greatest of our faculties. It is above work, because of its capacity to define us and give meaning to existence. Action means ethical and political action, the old praxis of the Greeks at the polis, in a few words, the virtuality of changing our life through our freedom

to decide by ourselves. Arendt's insights here again converge more closely with sociologists than with philosophers or political theorists. In particular, theorists of the "risk society" and "knowledge society" have arrived at very similar conclusions to Arendt concerning the transformed and transforming role of scientific knowledge within human societies. Thus, the originator of the concept of risk society, Ulrich Beck, has pointed out that "the category of risk society reflects the response to uncertainty, which nowadays often cannot be overcome by more knowledge but is instead a result of more knowledge" (Walsh 2011: 132). In this sense, technological advance will not necessarily bring with it something more positive by expanding knowledge. Therefore, we should be sensible when qualitatively and politically assessing the object of what is considered human enhancement in the sphere of transhumanism.

The Arendtian world of action is not free of worries or suffering, quite the contrary: there, suffering and worries are wanted, in the name of freedom and autonomy and not of pleasure or comfort. Hannah Arendt wrote that a miracle happens every day, when a human being is born. She was fonder, by the way, of the description of men as "the born" than as Heidegger's "the mortal" (Rivara 2010). Among the works where Arendt deals with the content of the human condition are *The Origins of Totalitarianism* (1951) and *The Human Condition* (1958) as fundamental works when interpreting the concept of a body subordinated to the political sphere. Arendt denies the existence of an innate, prescribed and unquestionable human nature. What is born – the human – is not simply a living organism, but a being endowed with the capacity to decide its environment and relations. Its proposal fundamentally focuses on investigating what we do. The activities that are within the reach of every human being, grouped into the concepts of "labour," "work" and "action," are of particular interest to Arendt, to the extent that these activities largely define human existence. When Arendt points out the "human condition", she alludes to the reciprocal relationship between what is "produced by human activities" (Arendt 2014: 23) and an individual's own existence. Activity and existence are categories that merge in the individual's relationship with the world. No human can inhabit the world without doing anything in it, while what they do – the activities framed under the notion of "vita activa" – define their humanity (Paterno & Crisorio 2016: 11).

Naturally, action brings consequences, and it is impossible, as Arendt herself pointed out, to fully calculate the consequences of any action. The best of intentions can bring the worst of evils, and, on the contrary, the worst intentions can bring positive results. It is also almost better this way: if everything were predictable, if every problem had a specific solution, we would live a zombie life. Therefore, it would be good if we did not analyse transhumanism from the point of view of the comforts that it can bring to human beings, but from the methodological caution of believing that human action can come from something that is not human.

Aristotle's concept of human nature can be a good point of reference to assess the normative dimension of transhumanism (Ogunyomi & Ogundele 2021). Aristotle made the distinction between the plane of "logos" and that of "physis" (Prevosti 2011: 40). Concepts and definitions are one side of human thought, but physical reality of things is something different. If we would build our life around concepts only, we would reduce empirical reality to something separated from our bodies.

Anyone who has read the first few chapters of the *Nicomachean Ethics* will know that Aristotle considers his political approach in relation to intuitions and sensations that shape virtue. We can see why Aristotle thinks that politicians should study his theory of the human good, since the aim of a ruler should be to provide the best life for the lucky few who are capable of living it and have the right to participate in government. This also explains why Aristotle insists on the necessity of legal rules for education. Since virtue of character is defined as a disposition to make the right choices in accordance with reason, we also get a discussion of choice and deliberation. The best virtue, namely excellence in pure theoretical thought, is what makes for the best life. Individual morality, according to Aristotle, cannot be studied in isolation from the political framework in which we live as agents. Only in the best city will the virtue of good men in general coincide with the virtue of the best citizens, and hence the best life can only be realized in the best social order, and the best social order must be linked to the human condition, since virtue emerges from the human condition (Striker 2007: 118–141).

Concepts are outside of time, while physical reality is in time, in space, immersed in movement. Like the human body: our cells are in continuous change. Concepts are universal, whereas physical reality is composed of the concrete (Marcos 2018: 120). Remember Prince Hamlet' when he claims: "there are more things in heaven and earth, than are dreamt of, in your philosophy Horatio" (Shakespeare 2011: 89).

Feyerabend also said something similar when pointing out that "we try through the concept to conquer the abundance of the real" (Feyerabend 1999: 26). There are, according to this, no universal principles of scientific rationality; the growth of knowledge is always peculiar and different and does not follow a prefixed or determined path. Feyerabend strongly defends the value of inconsistency and anarchy in science, from which, he claims, science has derived all its positive characteristics and argues that a combination of criticism and tolerance of inconsistencies and anomalies, as well as absolute freedom, are the best ingredients of a productive and creative science (Balandier 2003: 50). Feyerabend points to the problem of the scientific method, and the conclusion that follows is that it makes no sense to formulate in a general way questions such as what criteria one would follow in preferring one theory to another. To put it more clearly, successful research does not obey general standards: it relies on one rule or another, and the movements that make it advance are not always explicitly known (Vásquez 2006: 88).

The concept of “human nature” in the Aristotelian tradition is neither primarily a concept nor can it be compressed into a standard definition. Human nature is an individual form, present in each concrete human being. It is a physical entity, embedded in time and space and affected by change, but with the stability of substance. The content of this concept of human nature includes three different features: the animal, the social and the spiritual. We cannot separate each from the other (Marcos 2018: 121).

The normative analysis blossoms from this approach to human nature. This way, in connection with our animality, notions of “health and well-being” appear. Any transhumanist intervention contrary to them would have to be avoided. On the contrary, any therapeutic intervention should be considered positive. In connection with our sociability, normative terms blossom such as “peace, justice and freedom”. In the same way, they help us to evaluate anthropotechnical proposals and to reject those that undermine peace, justice or freedom. Finally, and connected to our spirituality we have the normative notions of “truth, beauty and goodness”. Only transhumanist contribution which acts in favour of them should be supported. And finally, given that these aspects are not isolated from another, that they are always integrated in each human body, normative notions such as identity will appear. So, interventions that do not break an individual’s identity will be acceptable and the rest will be unacceptable (Marcos 2018: 122).

Conclusion

Artificial intelligence has already had a significant impact on society, both positive and negative. On the one hand, it has led to advances in medicine, industry, security and many other fields, and has improved efficiency and accuracy in tasks that previously required human intervention. On the other hand, its inappropriate or malicious use could have negative consequences, unexpected or unintended, for society and individuals. The development and use of artificial intelligence raises significant ethical challenges that must be addressed to ensure its responsible and fair use. These challenges stem from the very nature of AI, which can make decisions and perform actions without direct human intervention. Throughout this paper we have distilled the possibility of a future where the creation of an artificial intelligence superior to the human one could come to dominate everything and even cause the extinction of the human species.

Nor should general artificial intelligence be confused with conscious minded machines or the artificial superintelligence of science fiction movies. To one day have general artificial superintelligence, an AI far superior in all respects to human

intelligence and its creativity, it should be able to create in turn machines more intelligent than itself. But this is something we cannot take for granted, at least in the sense that these improvements are unlimited.

Certainly, there are those who defend with conviction the possibility that in the future there will be machines capable of recursive self-improvement, which would be the type of self-improvement necessary to reach superintelligence, for instance, machines capable not only of improving their intelligence, but also of improving their capacity to make better machines. This is what many transhumanists think. But, for the moment, it is only a theoretical possibility under discussion. And there are experts who have repeatedly expressed scepticism about it.

But, even if it were ever possible to create a super-intelligent machine, it would not necessarily have to have consciousness. Consciousness, among other things, consists in having subjective experiences and feeling them as belonging to a coherent, temporally continuous and unified stream of subjective experiences. It also implies cognitive access to the mental processes that our relationship with the world elicits, that is, not only to experience the world, but to know that one is experiencing it at the very moment of doing so. It would be, then, a second-degree knowledge. And above all, consciousness is knowing oneself to exist in a reality that is different from oneself, and in which one is situated with a spatio-temporal perspective. These last two aspects could be considered as characteristic of self-consciousness.

Probably in no other field of scientific research as in artificial intelligence is there so much disagreement, not only about what future advances will be able to achieve, but about the correct interpretation of what has been achieved at present. And perhaps nowhere else does industry hype and media hype play such a prominent role. This peculiarity nurtures thought, and it cannot be ignored that one of its main causes lies in the need to keep social interest high in order to justify a growing investment. This should be enough to make us take a cautious approach to statements that are closer to propaganda than to scientific discourse.

However, since the possibility that the predictions of the most convinced will come true, the topic of artificial intelligence in relation to transhumanism deserves attention. If the dangers in relation to the application of artificial intelligence on individuals were to be confirmed and the real possibilities of creating a machine capable of recursive self-improvement were to increase, it would not be unreasonable to consider that any research aimed at achieving super-intelligent machines that would greatly improve the qualities of human beings would be ethically questionable, at least as long as there were no greater guarantees that human beings could always exercise their control over them. It is disturbing that not just a few scientists claim that there are probabilities that in the future an artificial superintelligence will lead to the extinction of mankind or to some severe damage to our species.

There have always been possible dangerous applications of scientific advances, but until today it has seemed that common sense and basic ethics acquired in any socialization process were sufficient for scientists to be able to develop guidelines to orient themselves in their work. Today, many teaching centres that offer degrees in biomedicine, genetics or related disciplines consider it necessary to complement the scientific training of their graduates with a solid bioethics' education. It is high time that the same thing happened in the field of artificial intelligence and computer science. But there is a problem here: "technoethics" or "ethics for artificial intelligence" or "ethics for machines" are still in their infancy. We have no clear criteria for deciding whether the construction of some kind of machines should be prohibited and even whether research in certain fields related to AI should be avoided, even temporarily. There are only a few proposals that have not yet reached a sufficient degree of agreement and about which the feasibility of their implementation is under discussion.

What underlies the transhumanist movement is both an enormous potential market that is emerging and the idea that human beings are biologically too limited to be able to effectively face the challenges of the world's growing complexity. Transhumanism also acquires an almost religious gnostic dimension, given that many authors believe in the possibility of making human beings immortal in the long term, or even in the technological resurrection of the dead, reproducing the scheme of Mary Shelley's *Frankenstein* as we mentioned before. The "post-human" individual that would result from this process would have basic capabilities that radically exceed those of humans today, to the point that they could no longer be qualified as merely human according to our current ways of understanding life. The risks associated with transhumanism's ethical conflicts refer to the possibility that human enhancement technologies may dehumanize people and undermine something as fundamental as human dignity. The degraded human being can become a harm to himself and become a threat to other individuals, since these improved and perfected human beings, by way of technological implementations, could want to establish a dominance and supremacy, which would engender spaces of injustice.

The central point of the transhumanist controversy is whether we should accept technology in order to transform ourselves into a being that would have nothing to do with human nature. Behind this dilemma lays the necessity to involve all available forces and factors in order to legally and ethically regulate the uncontrolled exponential growth and comprehensive expansion of technology, where the essential need to protect the fundamental human free will, personal identity and right to choose will be at the centre, bearing in mind all the consequences that an immoderate, unbalanced and unregulated calling and promoting of transhumanist ethics and values can have.

Ethical critiques may call into question the moral conception or worldview that underlies transhumanism. Some transhumanists proclaim themselves as libertarians, so they place freedom on top of human priorities and virtues. They also strongly believe in perpetual progress. Their uncritical, ecstatic acceptance and promotions of new technologies and striving to free themselves from all (including biological) limitations, leads to an extremely optimistic view on technology, or perhaps a kind of techno-utopian view.

But we could say that this framework is not libertarian. On the contrary, transhumanism rests on a utilitarian, teleological or consequentialist view of human progress and freedom, believing that the only morally correct values that really matter are the increase of group happiness and well-being along with the reduction of human suffering. According to the utilitarian doctrine elaborated by Jeremy Bentham, James Mill and John Stuart Mill, it would seek to achieve ends for tomorrow, not to compensate for yesterday's injustices. It is a rational and reasoned vision of laws so that they pursue a moral good and a benefit for all.

When Bostrom in 2014 suggested that intelligent machines should have values with human meaning from which to motivate their decisions and actions, he was already thinking of values linked to the utilitarian understanding based on maximizing happiness and minimizing suffering. However, these approaches highlight the limitations of transhumanist ethical approaches, based mainly on reason and utilitarianism, which leave aside many different visions and analyses present in the different beliefs, cultures and religions that make up humanity.

Transhumanist goals or ends can become threats to humanism or moral values despite the attempt to disguise this doctrine as a defence of moral rights, trying to take the use of technology for human improvement or progress to its ultimate consequences. However, arguing in favour of science does not guarantee that the practice carried out will always be respectful with the moral criteria relevant to human coexistence. That is why it is essential to remain vigilant against possible abuses or overreach of science that may jeopardize the basic moral demands or values of society.

The changing representation of characters such as Victor Frankenstein in popular culture show that society perceives the creator of life differently than it did during the 20th century. There are many questions about transhumanism that have not yet been answered, but the artificial creation of life by human beings is something that today is not seen as something coming from a madman or a sociopath, something that was representative of the 20th century's popular culture. The evolution of the character of Victor Frankenstein in popular culture would be an example of this evolution.

If we had to choose between the Terminator and the Replicant models, it is clear that we would opt for the second. In some way, it represents an expression

of technophilia, to the extent that it shows a representation of popular culture that projects a hopeful future of scientific advances that are associated with transhumanism. Cameron's Terminator model is precisely the opposite; it involves the projection of the fears and risks that technophobes detect on scientific and technological development. In any case, these models show that cinema has projected this debate with profound political, social and philosophical implications into popular culture for decades. In "Blade Runner" we see humanity proclaimed by a synthetic being created by humans. It is the representation of the evolution of technology that puts human beings in front of their own human condition, fallible and limited. In any case, human beings should always feel responsible for the results of humanity's technological advances. There is probably no greater vindication of our humanity than this. If we someday create robots with human-like cognitive and emotional capacities, we owe them more moral consideration than we would normally owe to otherwise similar human beings. We will have been their creators and designers. We are thus directly responsible both for their existence and for their happy or unhappy state.

It seems that the ethical solution to the dilemmas of AI should be provided by science itself. Yet this is not the case. It is ethics that should enlighten science, not the other way around. Every time that the radical positivist approach has been imposed in history, every time there has been an attempt to disregard the goals and values that correspond to human nature itself, the result has been catastrophic. For this reason, the advances in artificial intelligence linked to transhumanism must always be guided by those goals and values. That is precisely the moral philosophy and the philosophy of law approach.

Artificial intelligence technology brings great benefits in many areas, but without ethical barriers it risks reproducing real-world prejudice and discrimination, fuelling divisions and threatening human rights and fundamental freedoms. In no other field do we need an ethical compass more than in this.

We do not yet have enough information to assess all the implications linked to transhumanism, but at least, Arendt's concept of human action and Aristotle's concept of human nature can serve, if not to give us the answers to the challenges of transhumanism, to help us build the questions to face up its risks and contributions. Marcus Aurelius wrote that the great problems of existence demand always a simplification operation (Marcus Aurelius 1977: 33). Looking at what happens at the beginning of human action and what happens at the end, at the telos, can provide us with the right approach to at least formulate the questions we can ask when we face the challenges of transhumanism.

For all that has been expressed in this work, we must start from a basic premise in relation to artificial intelligence. We have stated that there are authors who see it as an advance for society, while others, on the contrary, see it as a great danger. If

we take as a reference the human being as conceptualized by Aristotle and Hannah Arendt, we realize which are the fundamental elements of human nature: virtue, sociability and human action. Any improvement project that contributes to this dimension of human nature is to be welcomed.

References

- Arendt, Hannah (2014) *La Condición Humana*. Buenos Aires: Paidós.
- Arendt, Hannah (2006) *Los Orígenes del Totalitarismo*, Madrid: Alianza.
- alandier, Georges (2003) *La Teoría del Caos y las Ciencias sociales*. Barcelona: Gedisa.
- Boissoneault, Lorraine (2017) Are “Blade Runner’s” Replicants “Human”? Descartes and Locke Have Some Thoughts Enlightenment philosophers asked the same questions about what makes humans, humans as we see in the cult classic. *Arts & Culture*, October.
- Bostrom, Nick (2005) Transhumanist values, *Review of contemporary philosophy* 4(1–2) 87–101.
- Britannica Online Encyclopedia. Prometheus. Published 27 March, 2024.
- Buchanan, Allen (2011) *Beyond humanity? The ethics of biomedical enhancement*. Oxford: Oxford University Press.
- Campione, Roger (2019) A vueltas con el Transhumanismo: cuestiones de futuro imperfecto. *Cuadernos Electrónicos De Filosofía Del Derecho* 40: 45–67.
- Cortina, Adela (2022) Los Desafíos Éticos del Transhumanismo. *Pensamient*, 78(298): 471–483.
- Cropsey, Seth (2024) Drone Warfare in Ukraine. Historical Context and Implications for The Future. *Strategika*.
- De Canio, Stephen (1994) The Future Through Yesterday: Long-term Forecasting in the Novels of H. G. Wells and Jules Verne. *The Centennial Review* 75–93.
- European Commission (2018) Ethics Guidelines for Trustworthy AI. Available at: <https://www.aepd.es/sites/default/files/2019-12/ai-ethics-guidelines.pdf>.
- Feyerabend, Paul (1999) *Contra el Método*. Barcelona: Altaya.
- Fukuyama, Francis (2002) *Our Posthuman Future: Consequences of the Biotechnology Revolution*. Profile Books.
- Habermas, Jürgen (2002) *El futuro de la Naturaleza Humana: ¿Hacia una eugenesia liberal?* Barcelona: Paidós.
- Keese, Andrew (2011) The Myth of the Monster in Mary’s Shelley’s Murder Mystery, Frankenstein, *Jostes: The Journal of South Texas English Studies* 2(2).
- Kekes, John (1998) *A Case for Conservatism*. Ithaca: Cornell University Press.
- Kerr, Breana (2020) Jules Verne: The Sci-Fi Author Who Predicted the Future. *The Hustle*. June 30.
- Khouw, Sebastiaan (2020) Terminators and the Philosophy of Empathy. Can a Terminator Learn to Feel? *Medium*, August 16.
- Marco Aurelio (1977) *Meditaciones*. Madrid: Gredos.
- Marcos, Alfredo (2018) Bases filosóficas para una crítica al Transhumanismo. *ArtefaCToS. Revista de Estudios de la Ciencia y Tecnología* 7(2): 107–125.
- Martínez-Córcoles, Mario, Teichmann, Mare & Murdvee, Mart (2017) Assessing Technophobia and Technophilia. *Technology in Society* 51: 183–188.
- Marx, Leo (1994) “The idea of technology and postmodern pessimism” in *Technology, Pessimism, and Postmodernism*. *Sociology of the Sciences Yearbook, SOSCS*, Vol 17, pp. 11–28.

- Ogunyomi, Abidemi & Ogundele, Emmanuel (2021) Aristotle, Confucius and Rousseau on Human Nature and the Golden Mean: a comparative analysis. *Prajña Vihara: Journal of Philosophy and Religion* 22(1) January-June.
- Ortiz de Zárate, Lucía. (2020) El Transhumanismo o el Fin de las Esencias: el (bio)conservadurismo y su Reminiscencia Aristotélica. *Logos* (53): 99–118.
- Osiceanu, Maria-Elena (2015) Psychological Implications of Modern Technologies: Technophobia versus Technophilia”, *Procedia. Social and Behavioural Sciences* (180): 1137–1144.
- Patierno, Nicolas & Crisorio, Ricardo Luís (2016) Cuerpo y Naturaleza Humana en la obra de Hannah Arendt. *INTERthesis Revista Internacional Interdisciplinar* 13(2) Maio-Agosto: 1–18.
- Prevosti, Antoni (2011) La Naturaleza Humana en Aristóteles. *Espíritu LX* 141: 35–50.
- Rejeski, David & Olson, Robert L (2006) Has Futurism Failed? *The Wilson Quarterly* 30(1): 14–21.
- Rivara, Greta (2010) Apropiación de la Finitud. Heidegger y el Ser para la Muerte. *En-Claves del Pensamiento* 4(8).
- Sandel, Michael (2007) *Contra la Perfección: La Ética en la Era de la Ingeniería Genética*. Barcelona: Marbot.
- Sasani, Samina & Pilevar Hamid Reza (2016) No Romantic Prometheus: Mary Shelley’s Frankenstein and Rejection of Romanticism, *International Journal of English Language & Translation Studies* 4(3): 50–59.
- Serra, Miquel Àngel (2022) Human Enhancement and Functional Diversity. Ethical Concerns of Emerging Technologies and Transhumanism. *Metode. Science Studies Journal* 12: 169–175.
- Shakespeare, William (2011) *Hamlet*, Madrid: Alianza.
- Striker, Gisela (2007). Aristotle’s ethics as political science in Aristotle in *The Virtuous Life in Greek Ethics*. Universität Hamburg: Cambridge University Press.
- Sullins, John (2011) “Replicating Morality” in *Philip K. Dick and Philosophy: Do Androids Have Kindred Spirits?* Illinois, Open Court, 197–206.
- UNESCO. Recommendation On Ethics of Artificial Intelligence. Unesco, 2022.
- United Nations. The Right to Privacy in The Digital Age. Report of the United Nations High Commissioner For Human Rights. 13 September, 2021.
- Vásquez, Adolfo (2006) La Epistemología de Feyerabend; Esquema de una Teoría Anarquista del Conocimiento. *Revista Observaciones Filosóficas*, Abril.
- Walsh, Philip (2011) The Human Condition as Social Ontology: Hannah Arendt on Society, Action and Knowledge. *History of the Human Sciences* 24 (2) 120–137.
- Westfahl, Gary (2007) *Gernsback, Hugo and the Century of Science Fiction. Critical Explorations in Science Fiction and Fantasy*, North Carolina: McFarland.