



**SCIREA Journal of Sociology**

<http://www.scirea.org/journal/Sociology>

**June 4, 2019**

**Volume 3, Issue 1, February 2019**

## **Artificial Intelligence Evolution:**

### **On the virtue of killing in the artificial age**

**Julia M. PuaSchunder<sup>1,2,3\*</sup>**

<sup>1</sup>The New School, Department of Economics, Schwartz Center for Economic Policy Analysis, 6 East 16<sup>th</sup> Street, 11<sup>rd</sup> floor, 1129G-99, New York, New York 10003, USA,

[Julia.Puaschunder@newschool.edu](mailto:Julia.Puaschunder@newschool.edu)

<sup>2</sup>Columbia University, Graduate School of Arts and Sciences, 116<sup>th</sup> Street Broadway, New York, New York 10027, USA

[Julia.Puaschunder@columbia.edu](mailto:Julia.Puaschunder@columbia.edu)

<sup>3</sup>Princeton University, Graduate School of Arts and Sciences, USA

[julia.puaschunder@princeton.edu](mailto:julia.puaschunder@princeton.edu)

\*Financial support of the Research Association for Interdisciplinary Studies, The New School Dean's Office, The New School Department of Economics, The New School Eugene Lang College, The New School Fee Board, The New School for Social Research, The New School for Public Affairs, University of Vienna and Vernon Arts and Science is gratefully acknowledged. The author declares no conflict of interest. All omissions, errors and misunderstandings in this piece are solely the author's.

## **Abstract**

Artificial Intelligence (AI) poses historically unique challenges for humankind. In a world, where there is a currently ongoing blend between human beings and artificial intelligence, the emerging autonomy of AI holds unique potentials of eternal life. With AI being endowed with quasi-human rights and citizenship in the Western and Arabic worlds, the question arises how to handle overpopulation but also misbehavior of AI? Should AI become eternal or is there a virtue in switching off AI at a certain point? If so, we may have to redefine laws around killing, define a virtue of killing and draw on philosophy to answer the question how to handle the abyss of killing AI with ethical grace, rational efficiency and fair style. The presented theoretical results will set the ground for a controlled AI-evolution in the 21st century, in which humankind determines which traits should remain dominant and which are meant to be killed.

**Keywords:** Artificial Intelligence, AI, algorithms, cognitive robotics, AI-evolution, emerging technologies, ethical issues, ethics, human robot interaction, international law, killing, legal personhood, roboethics, robot-rights, social robots, virtue of killing

## 1. Introduction

Artificial Intelligence (AI) poses historically unique challenges for humankind. As emerging globally trend, AI is extending its presence at almost all levels of social conduct and thereby raised both – high expectations but also grave concerns (Cellan-Jones, 2018; Sofge, 2015; United Nations, 2017). With the dramatic growth in diversity and entrance of emerging technologies in today's societies, such as social robots, lifelike computer graphics (avatars), and virtual reality tools and haptic systems, the social complexity of these challenges are on the rise (Meghdari & Alemi, 2018). One of the main challenges in developing and applying modern technologies in our societies is the identification and consideration of ethical issues surrounding AI (Meghdari & Alemi, 2018). The call for AI Ethics (AIE) has emerged. A growing number of AI and robotics researchers have demanded to create a framework on AI ethics building on the benefits of humanities, philosophy, natural sciences, sociology, and social neuroscience.

AI will hold the potential to replicate human existence but also grant eternal being opportunities. In the eye of overpopulation concerns, finding mechanisms to switch off AI would be a solution to avoid a crowding of the planet. But AI currently also reaches quasi-human status through actual personhood – e.g., via citizenship and quasi-human rights applied in the Common Law but also Roman Law territories of the US and the EU. Leveraging AI entities to the status of being through the attribution of legal personhood raises challenging legal and ethical questions. Programming AI to switch itself off or switch off AI at a certain point to curb overpopulation but also as quality control against harmful behavior arising out of AI, thereby appears critical as it would come close to suicide or killing. A novel predicament between eternity and overpopulation hence calls for revising legal codes for killing and ethical imperatives and religious concerns over suicide.

But how to argue the right to terminate AI legally? And when to pull the plug? We may want to draw on the ethics of dying and virtues of killing as well as suicide literature to answer these novel questions arising out of AI. When considering the opportunity to determine life and death of AI, humankind will see the opportunity of AI-evolution understood as a human-made evolution determining what contents survive and what to die following the goal to improve the overall offspring and general well-being of humankind. The proposed frame will offer innovative insights for legal conduct but also overlapping generations relationships. The nature of algorithms and digital technology being global

demands for an international response, potentially via international law supremacy principle (Themistoklis, 2018). In this paper, the novel and multidisciplinary area of socio-cognitive robotics, and the ethical challenges of emerging technologies are explored. Key ethical features based on past and present research in a variety of AI areas will be presented.

The paper is structured as follows: First, the ontology of AI is presented as well as an analysis of legal personhood. Then, the predicament between eternal life and overpopulation is addressed. The virtues of dying and killing but also philosophical arguments for the right to live or choose suicide are discussed. The paper closes with an international law and future research prospects on regulating AI and overall future outlook.

## **2. Theory**

### **2.1 Artificial Intelligence**

Artificial Intelligence (AI) is “a broad set of methods, algorithms, and technologies that make software ‘smart’ in a way that may seem human-like to an outside observer” (Noyes, 2016). The “human-like” intelligence of machines derives from machines being created to think like humans but at the same time to also act rationally (Laton, 2016; Russell & Norvig 1995; Themistoklis, 2018). AI is perceived as innovative technology or as the sum of different technological advances as the privilege of the private, technological sector with little — if any — public regulation (Dowell, 2018).

As the most novel trend, AI, robots and algorithms are believed to soon disrupt the economy and employment patterns. With the advancement of technologies, employment patterns will shift to a polarization between AI’s rationality and humaness. Robots and social machines have already replaced people in a variety of jobs – e.g. airports smart flight check-in kiosks or self-check-outs instead of traditional cashiers. Almost all traditional professional are prospected to be inflused with or influenced by AI, algorithms and robotics. For instance, robots have already begun to serve in the medical and health care profession, law and – of course – IT, transportation, retail, logistics and finance, to name a few. Social robotics may also serve as quasi-servants that overwhelmingly impact our relationships. Already, social robots are beginning to take care of our elderly and children, and some studies are currently underway on the effects of such care (Alemi, Meghdari & Saffari, 2017). Not only will AI and robots offer luxuries of affordability and democratization of access to services as they will be – on the long run – commercially more affordable and readily available to serve all

humanity; but also does the longevity potential of machines outperform any human ever having lived (Hayes, 2018). However, the new technology also comes with the price of overpopulation problems and the potential for misuse and violent action. Just like many other technologies, robots could be misused for wars, terrorism, violence and oppression (Alemi et al., 2017).

AI's entrance in society will revolutionize the interaction between humans and AI with ample legal, moral and social implications (Kowert, 2017; Larson, 2010). Autonomous AI entities are currently on the way to become as legal quasi-human beings, hence self-rule autonomous entities (Themistoklis, 2018). AI is in principle distinguished between weak AI, where "the computer is merely an instrument for investigating cognitive processes" and strong AI, where "[t]he processes in the computer are intellectual, self-learning processes" (Wisskirchen, Biacabe, Bormann, Muntz, Niehaus, Jiménez Soler & von Brauchitsch, 2017, 10). Weak AI is labeled as Artificial Narrow Intelligence (ANI) while strong AI is further distinguished between Artificial General Intelligence (AGI) and Artificial Super Intelligence (ASI).

The emergence of robotics technology is developing much quicker than previously thought. Robots are anticipated to soon be as ubiquitous as computers are today (Meghdari & Alemi, 2018). Society has long been concerned with the impact of robotics technology from nearly a century ago, when the word "*Robot*" was devised for the first time (Čapek, 1921; Meghdari & Alemi, 2018). The EU Committee on Legal Affairs (2016, p. 4) holds that "[U]ltimately there is a possibility that within the space of a few decades AI could surpass human intellectual capacity in a manner which, if not prepared for, could pose a challenge to humanity's capacity to control its own creation and, consequently, perhaps also to its capacity to be in charge of its own destiny and to ensure the survival of the species." AI mimicking human intellect could soon surpass humans intellectually but also holistically breaking the barrier of human controlled-automization (Schuller, 2017; Themistoklis, 2018). Modern literature about robots features cautionary accounts about insufficient programming, evolving behavior, errors, and other issues that make robots unpredictable and potentially risky or dangerous (Asimov, 1942/1950, 1978, 1985; Meghdari & Alemi, 2018). "Observe, orient, decide, act" will therefore become essential in the eye of machine learning autonomy and AI forming a new domain of intellectual entities (Armstrong & Sotola 2012, 52; Copeland 2000; Galeon & Reedy, 2017; Marra & McNeil, 2013). The uncertainty

surrounding AI development and self-learning capabilities give rise to the need for guarding AI and an extension of the current legal system to cope with AI (Themistoklis, 2018).

With the advancement of technology, social robots have found broader applications in the private and public sectors, such as educational and cultural affairs, games and entertainment, clinical and rehabilitation, nursing of children and/or elderly, search and rescue operations (Meghdari, Alemi, Shariati & Zakipour, 2018). For example, social robots such as ASIMO, Nao, iCub, ARASH, and RASA have been developed for “Edutainment” or “education-entertainment” purposes. They aid the study of cognition (both human and artificial), motion, and other areas related to the advancement of robotics serving our society (Meghdari & Alemi, 2018). In addition, a few medical and healthcare toy-like robots, such as PARO, which looks like a baby seal, or ARASH, which is a humanoid, have been designed for therapeutic purposes such as reducing distress, stimulating cognitive activity, teaching specific subjects, and improving socialization (Meghdari, Shariati, Alemi & Vossoughi, 2018). Similarly, Sharif University of Technology’s socially assistive robot RASA has been developed to help coach and teach Persian Sign-Language to Iranian deaf children (Meghdari, Alemi, Zakipour & Kashanian, 2018). Personal care and companion robots are increasingly being used to care for the elderly and children, such as RI-MAN, PaPeRo, and CareBot (Meghdari & Alemi, 2018). In recent years, robotics technology has extended its applications from factories to more general-purpose practices in society – for instance, such as the use of robots in clinical and rehabilitation, nursing and elderly care, search and rescue operations (**Meghdari & Alemi, 2018**). Social robots have become clinical and educational assistants for social interventions, treatment, and education such as language trainings but also assistance with children with disabilities like autism, down syndrome, cancer distress, hearing impairment, etc. (Meghdari et al., 2018). Initial investigations clearly indicate that social robots can play a positive role in the improvement of children’s social performance, reduction of distress during treatments, and enhancing their learning abilities (Meghdari & Alemi, 2018). Surprisingly, although not too hard to imagine, relationships of a more intimate nature have not quite been satisfied by robots yet (Meghdari et al., 2018; Veruggio, 2005).

## **2.2 AI-Evolution**

The human perception of and interaction with robot machines with a higher quality physical appearance differs from interaction with a computer, cell phone, or other smart devices (Meghdari & Alemi, 2018). For robotics technology to be successful in a human-driven

environment, robots do not only need to meet a level of strength, robustness, physical skills, and improved cognitive ability based on intelligence but should also fulfill a social impetus and ethical conscientiousness. The design and construction of social robots faces many challenges, one of the most important is to build robots that can comply with the needs and expectations of the human mind with cognitive capabilities coupled with social warmth (Meghdari & Alemi, 2018). While we have *Social-Cognitive Robotics* (SCR) as a transdisciplinary area of research and a basis for the human-centered design of technology-oriented systems to improve human knowledge functions, judgements and decision making, collaborations, and learning; hardly any information exists on socio-evolutionary comparisons (Meghdari & Alemi, 2018). Social-cognitive robotics has been evolving and verified through a series of projects to develop advanced and modern technology-based systems to support learnings and knowledge functions, and is beginning to play an effective role in societies across the globe (Meghdari & Alemi, 2018). SCR or *Socio-Cognitive Robotics* is the interdisciplinary study and application of robots that are able to teach, learn and reason about how to behave in a complex world (Meghdari & Alemi, 2018). Social robotics technology promises a many benefits but also challenges that society must be ready to confront with legal means and ethical imperatives.

### **2.3 Roboethics**

*Ethics* describes moral principles that govern a person's or group's behavior. Roboethics describes the ethics and morals of robotics, the science of robots. Roboethics therefore captures the integration of ethics into AI and algorithms. This field recently gained considerable attention among humanities and robotics engineers who draw on insights from computer science, artificial intelligence, mechanics, physics, math, electronics, cybernetics, automation and control (Meghdari & Alemi, 2018).

What specifies the emergence of socio-cognitive robotics is that humanity is at the threshold of replicating an intelligent and autonomous agent (Meghdari & Alemi, 2018). In order to enhance the ability of social robots to successfully operate in humane ways, roles and environments, they are currently upgraded to a new level of physical skills and cognitive capabilities that embrace core social concepts (Meghdari et al., 2018). Robotics thereby unifies two cultures, in which complex concepts – like learning, perception, decision-making, freedom, judgement, emotions, etc. – may not have the same semantic meaning for humans and machines (Meghdari & Alemi, 2018).

In the design and construction of social robots, the consideration of ethical concerns has therefore leveraged into an imperative (Lin, Abney & Bekey, 2012). Human-robot (a machine with a higher physical and social ability) interactions, are somewhat different compared to other types of human-machine interactions (i.e. with a computer, cell phone, or other smart device) (Meghdari & Alemi, 2018; Saffari, Meghdari, Vazirnezhad & Alemi, 2015). It is therefore essential for researchers, scholars, and users to clearly identify, understand, and consider these differences and ethical challenges so that they can benefit from and no one gets harmed by the assistance of social robots as a powerful tool in providing modern and quality services to society (Meghdari & Alemi, 2018; Taheri, Meghdari, Alemi & Pouretamad, 2018).

Robots and algorithms now taking over human decision-making tasks and entering the workforce but also encroaching our private lives, currently challenges legal systems around the globe (Themistoklis, 2018). The attribution of human legal codes to AI is one of the most groundbreaking contemporary legal and judicial innovations. Until now legal personhood has only been attached directly or indirectly to human entities (Dowell, 2018). The detachment of legal personhood from human being now remains somewhat of a paradox causing an extent of “fuzziness” of the concept of personhood (Barrat 2013; Solum 1992, p. 1285). As AI gets bestowed with quasi-human rights, defining factors of human personhood will need to be adjusted (Dowell, 2018). Human concepts, such as morality, ownership, profitability and viability will have different meaning for AI. The need for redefining AIE has therefore reached unprecedented momentum.

As a predicted trend, the co-existence of AI with the human species is believed to change the fundamental concepts of social, political and legal systems. AI has already produces legal creations and will do so even more in the near future, through its developing autonomy. In addition, the technology leading to AGI and ASI is already present, posing moral and legal dilemmas about who should control it and under what terms (Themistoklis, 2018). The emergence of AGI and ASI will necessitate the attribution of some extent and of some type of legal personhood, bearing rights and obligations. AI will not be most probably an exact replication of human intellect behavior (Themistoklis, 2018). “[U]ltimately, robots' autonomy raises the question of their nature in the light of the existing legal categories – of whether they should be regarded as natural persons, legal persons, animals or objects – or whether a new category should be created, with its own specific features and implications as regards the attribution of rights and duties” (Committee on Legal Affairs 2016, p. 5).

Behavioral economists add the question whether AI and robots should be created to resemble human beings' decision making with fast thinking and fallible choices or rather be targeted at perfect rationality and slow thinking (Kahneman, 2011). General conscious is strived for so that AI possesses consciousness, which it can evolve and enhance on the basis of its own critical reflection and assessment of external factors (Themistoklis, 2018). A lower level of autonomy exists if an entity can demonstrate such consciousness at a narrow field or can self-evolve and self-adapt to external influences, thus reaching decisions “of its own,” without being conscious of its intelligence as such (Themistoklis, 2018). As AI emerges as new types of intellect capacities coupled with human-like emotional features, they are attributed a legal personhood in order to ensure to be comprehended correctly and to avoid unfair treatment, towards humans as well (Themistoklis, 2018). Artificial entities are currently gaining human or quasi-human status in the Western and Arab worlds in forming an intellectual autonomy of the entity (MacDonald, 2016). For instance, in Saudi Arabia the first female robot got a citizenship in 2017 and the robot appears to have more rights than a human female in Saudi Arabia.<sup>1</sup> With the rise of AI persons, their eternal life poses ethical challenges in light of overpopulation and evolutionary perfection, which could crowd out human fallibility if determining merit-based eternal life. These critical questions will be captured in the following.

## **2.4 Eternal life**

While there is currently cutting-edge writing about the potential emergence of an AI personhood as well as concern over the merge of AI with cyberspace that might lead to the breach of the relationship between legal personhood and nation state sovereignty and a nomenclature is emerging on legal characterizations of different levels of AI development; hardly any information exists about the eternal living of AI (Hildebrandt, 2013). From the theoretical standpoint, the eternal longevity of AI contradicts the fundamental concept of fairness in death, as a general condition for all. From the practical standpoint, the international community is currently urged to think on the basis of global commons in terms of AI and AI eternal life potentials contributing to overpopulation. Thereby global commons theories may be tabbed on, which primarily offer guidance for a regulatory framework, which establishes control “...for the benefit of all nations” and refer to space constraints (Clancy, 1998; Tsagourias, 2015).

---

<sup>1</sup><https://www.forbes.com/sites/zarastone/2017/11/07/everything-you-need-to-know-about-sophia-the-worlds-first-robot-citizen/#214d746946fa>

Regarding limited space, longevity and eternal life appears problematic. Humankind may face tough decisions whether or not to have AI proceed and what kind of developments to flourish and what to extinct. In what cases should we consider to switch off AI? In 1950, Isaac Asimov introduced the idea robot to (1) not injure a human being or, through inaction, allow a human being to come to harm. (2) A robot obeying the orders given it by human beings except where such orders conflict with the first law. (3) A robot must protect its own existence as long as such protection does not conflict with the first or second law. So in the cases of overpopulation and harm emerging from AI, algorithms and robots can be considered to be switched off. But when to stop AI?

Another killing market mechanism may be natural market selection via price mechanisms and the falling rate of profit. Regarding prices, natural supply and demand mechanisms will always favor innovation with a higher price and following supply of goods lead to a price drop. The falling rate of profit is one of the major underlying features of business cycles, long-term booms and downturns (Brenner, 2003, 2006a, b). Capitalism is thereby described as competitive battle for innovation and reaping benefit from first-market introductions. Once followers enter the market, profit declines, leading eventually to market actors seeking novel ways to innovate in order to regain a competitive market advantage and higher rates of profit. Thereby industries and innovations fade and die off. Such a natural market evolution is also likely to occur with AI innovations, which will determine which AI traits will remain and which ones will fade off.

Apart from soft market mechanisms that may lead to AI evolution, what are the cases when AI should be shut down or switched off or – in the case if AI personhood – be killed?

## **2.5 Death**

*Errors and Safety:* When errors occur and general safety is at stake. The main and leading concern about any new and emerging technology is to be safe and error free (Meghdari & Alemi, 2018). Therefore, sufficient and numerous tests on health and safety must be performed by developers and/or well-known independent sources before rolling out any technology onto the marketplace and society (Meghdari & Alemi, 2018). In robotics, the safety issue mainly centers around software and/or hardware designs (Meghdari & Alemi, 2018). Even a tiny software flaw or a manufacturing defect in an intelligent machine, like a smart car or a social robot, could lead to fatal results (Meghdari & Alemi, 2018). When these deviations occur and especially when they are harmful to the human community but also to

other AI species, the faulty AI should be terminated. With regard to the risk of robotic malfunctions and errors, product legal responsibility laws are mostly untested in robotics (Meghdari & Alemi, 2018). A usual way to minimize the risk of damage from social robots is to program them to obey predefined regulations or follow a code-of-ethics (Meghdari & Alemi, 2018). Ethical codes for robotics are currently needed and should become formed as a natural behavioral law to then be defined and codified as law. Laws but also an ethical understanding to terminate AI, algorithms and robots in case of impairment and harm are needed.

*Morals, Ethics, and the Law:* As social robots become more intelligent and autonomous and exhibit enough of the features that typically define an individual person, it may be conceivable to assign them responsibility and use them in social, educational, and therapeutic settings (Meghdari & Alemi, 2018). In the currently ongoing research on the integration of computers and robotics with biological corpse it is found that a cognizant human brain (and its physical body) apparently has *human-rights*; hence, replacing parts of the brain with artificial ones, while not harming its function, preserves those rights (Meghdari & Alemi, 2018; Warwick & Shah, 2014). Also, consider a handicapped person featuring an electronic robot arm that commits a crime. It becomes obvious that half-robot-human beings should be considered as human and robots as quasi-human beings. Meghdari & Alemi (2018) speculate that at some point in the future, we may face a situation in which more than half of the brain or body is artificial, making the organism more robotic than human, which consolidates the need of special *robot-rights* and attributing (quasi)-human rights onto robots. When considering robots as quasi-human beings, their termination appears legally questionable and ethically challenging, requiring to revisit laws as legitimation to kill a likewise species as well as ethical consensus on the virtue of killing.

The legal argumentation may draw on justifiable homicide as outlined in criminal law cases – such as prevention of greater harm to innocents during an imminent threat to life or well-being in self-defense. According to the United Nations Universal Declaration of Human Rights, Article 3 states that everyone has the right to life, liberty and security of person and most nations' policy allows for some degree of leniency for self-defense, which reduces charges.<sup>2</sup> Potentially excusing conditions common to most jurisdictions include wartime, when the person's death is inflicted by the effect of a lawful arrest or prevention of lawfully

---

<sup>2</sup> <http://www.un.org/en/universal-declaration-human-rights/>

detained person's escape, quelling riot or insurrection, when the use of force is „no more than absolutely necessary.“ Some countries deem it lawful for a citizen to resort to violence to protect valuable property and there is the “heat of the moment“ defense argument, in which the defendant deemed to have lost control through provocation. Doctrine of necessity allows, for example, a surgeon to separate conjoined twins and killing the weaker twin to allow the stronger twin to survive. While fetuses are considered as unborn children in the US, the right to an abortion was upheld in the US legal system as exception from prosecution (*Roe v. Wade*, 1973). Several countries, such as the Netherlands, Belgium, Switzerland, Japan, and the U.S. states of Oregon and Washington, allow both active and passive euthanasia by law, if justified.

Where the person concerned is to be arrested for an offense referred to in Schedule 1 or is to be arrested on the ground of having committed such an offense, and the person authorized under this Act to arrest or to assist in arresting him cannot arrest him or prevent him from fleeing by other means than killing him, the killing shall be deemed to be justifiable homicide.

If any arrestor attempts to arrest a suspect and the suspect resists the attempt, or flees, or resists the attempt and flees, when it is clear that an attempt to arrest him or her is being made, and the suspect cannot be arrested without the use of force, the arrestor may, in order to effect the arrest, use such force as may be reasonably necessary and proportional in the circumstances to overcome resistance or to prevent the suspect from fleeing: Provided that the arrestor is justified in terms of this section in using deadly force that is intended or is likely to cause death or grievous bodily harm to a suspect, only if he or she believes on reasonable grounds (§7 Judicial Matters Second Amendment Act 122 of 1998).

In light of overpopulation and harmful behavior of AI, switching off artificial life, which is currently be granted quasi-human status, will need to be argued legally and supported ethically. Killing in terms of the death penalty is justified legally in **the 5<sup>th</sup> (and the 14<sup>th</sup>) amendment that states “no person shall be deprived of life, liberty, or property without due process of law,” while the 8<sup>th</sup> amendment prohibits “cruel and unusual punishment.”**

**Killing in terms of harmful behavior of AI can be grounded on similar legal reasons to ensure that no AI harms the collective. Overpopulation claims leading to the need to take AI partially off the grid more lead to philosophical sources that argue for individual's free will to choose to live or die (Critchley, 2015; Critchley & Hume, 2016).**

**Suicide has been tabooed for most part of history and propagated to be a religious sin. Yet the human gift of reflection and search for meaning in life or death could leverage into an asset in the AI evolution in the decades to come. We could argue that similar to critique on those** who proclaim loudly against suicide and claim that the act of taking one's own life is irresponsible and selfish, even shameful and cowardly, that people must stay alive whatever the cost (Critchley, 2015; Critchley & Hume, 2016); there will be virtue in the killing AI. Suicide understood as neither a legal nor moral offence but as right to life or death bestowed upon human beings in their self-conscious reflection may be extended as a virtue of killing in the artificial age, when human beings will have to decide what AI should stay alive and what AI be taken off the grid. Human will thereby become the rulers of the forthcoming AI evolution.

The virtue of killing could also be grounded on Viktor Mayer-Schönbergers “right to be forgotten,” which ensures data privacy through automated deletion of contents after a certain period and grants individuals rights to have their data been destroyed (Puaschunder, 2018a, forthcoming). However, the implementation of this right is still in infancy and hindered by questions of what court is responsible for an as such claim. As a legal subsumption, we may speculate that individuals may be granted a 'right to terminate' and can order for robots to be switched off if causing harm to them. As the 'right to be forgotten' law can be overruled by concern for public safety, this may also apply to the right to terminate. Thereby it deserves mentioning that safety differs around the world and also expected safety standards.

## **2.6 AI-Evolutionary pressure turning against human**

The predicted AI-Evolution (AIE) is grounded on evolution as the change in heritable characteristics of biological populations over successive generations. As for human evolution, these characteristics are the expressions of genes that are passed on from parent to offspring during reproduction. Different characteristics tend to exist within any given population as a result of mutations, genetic recombination and other sources of genetic variation. Evolution occurs when evolutionary process such as natural selection (including sexual partner selection) and genetic drift act on these variations, resulting in certain characteristics becoming more common or rare within a population. This process has given rise to biodiversity at every level of biological organisation including the levels of species, individual organisms and molecules. Evolution by natural selection defines the following facts about living organisms: Traits vary among individuals with respect to their morphology, physiology and behavior (phenotypic variation). Different traits confer different rates of

survival and reproduction (differential fitness). Traits are passed from generation to generation (heritability of fitness). Thus, in successive generations members of a population are more likely to be replaced by the progenies of parents with favorable characteristics that have enabled them to survive and reproduce in their respective environments.

AIE now refers to the human process of selecting what AI should survive or be killed by being taken off the grid forming heritable characteristics of blockchain-like created populations of robots and AI. Like genes being passed on from parents through natural mate selection, decision makers will divert favorable traits from unfavorable. Mutations may occur in decision making errors innate in human beings as described by behavioral economics (Puaschunder, 2017a). AI traits will be varying in their survival rate. Favorable characteristics will have a higher likelihood to survive. But what will count as favorable will be determined by human and therefore add a social touch to future AI to come. However, the critical problem appears that robots will outperform human beings and could turn around evolutionary pressures towards the eradication of the fallible species of human. In the creation of AI, stereotypes should be eradicated and a social class division avoided (Puaschunder b, c, d).

### **3. Discussion**

The growing number of AI and robotics researchers are demanded to create a framework on AI ethics building on the benefits of humanities, philosophy, sociology, and social neuroscience expertise and research. Likewise, growing trends of mutual collaboration among scholars in the field of human sciences, linguistics, and psychology with the robotics scientists are producing quite noticeable valuable results (Meghdari & Alemi, 2018). Future studies should target at presenting an overview of the novel and multidisciplinary area of socio-cognitive robotics, and further explore the possible ethical challenges of emerging technologies on education, culture, entertainment, gaming, nursing, and therapy. Unraveling ethical features based on our past and present research experiences in a variety of areas will aid designing safe AI and social robots.

In its entirety, this article was the first introduction of AI ethics opening up many challenging questions. For instance, what ethical code should we apply for controlling robots' actions? How can we program a switch to turn off AI in case of unlawful action and harm to people but also how to draw the boundary condition to ethical infringements? This is specifically

important if humankind starts placing social robots in positions of authority, such as police, security guards, teachers, or any other government roles or offices, in which humans would be expected to follow them.

In the further discussion of the topic, research should analyze the effects of robotics blending into our societies with direct applications in fields where the potential complications are more significant and apparent (Meghdari & Alemi, 2018). Important areas of scrutiny should be human rights/dignity, equality and justice, benefits and damage, cultural diversity and pluralism, religious variety, non-discriminating, independence and individual accountability, privacy and confidentiality, unity and collaboration, social responsibility, benefits sharing and environmental obligations as well as intergenerational equity considerations (Meghdari & Alemi, 2018).

#### **4. Conclusion**

The days of AI being a futuristic concept are over. AI is now. Social and cognitive robotics is rapidly becoming one of the leading fields of science and technology involving a deep level of human-machine interaction (Meghdari & Alemi, 2018). The world will soon be populated with human and machines alike that will coexist. The clear advantage of AI is the longevity. In light of overpopulation fear, we need mechanisms to determine how to decide over what is worth living forever and what should be taken off society. Ethics may come into this predicted AI-evolution. One may conclude that roboethics entails the ethics of handling and application of robots (Meghdari & Alemi, 2018).

It is predicted that society is expected to fall into two extremes of a dichotomy between rationality (represented by AI) and humanness (represented by human beings). Hereby the question arises what is it that makes human humane? In the age of artificial intelligence and automated control, humanness is key to future success. Behavioral human decision making insights and evolutionary economics can already today predict what makes human humane and how human decision making is unique to set us apart from artificial intelligence rationality. Future research in these domains promise to hold novel insights for future success factors for human resource management but also invaluable contributions for artificial intelligence ethics (Puaschunder, 2018b).

Overall this paper was meant as first step towards a nomenclature of deciding on the future evolution grounded in the virtue of living and killing to motivate different viewpoints on the

issue by cultural, religious, and ethical scholars. The article plays an important role in the evolution of an AI and human mixed society in order to ground stability and social harmony into the newly emerging system. Depicting ethical imperatives around the life and death of machines being considered as quasi-human beings during this unprecedented time of societal change and regulatory reform holds invaluable historic opportunities for global governance policy makers to snapshot the potential but also save from the likely downfalls of a robo-human mixed society.

The results are targeted at guiding a successful introduction to AI and lower systemic downfalls with attention to the changes implied in the wake of the ongoing artificial intelligence revolution. Market and societal policy recommendations for global governance experts on how to strengthen society but also overcome unknown emergent risks within globalized markets and bestow market actors with key qualifications in a digitalized world are endeavored alongside scientific publications and stakeholder engagement.

In the international compound, having parts of the world being AI-driven and others being human capital grounded is prospected to increase the international development divide in the years to come. While in the AI-hubs human will be incentivized become more creative and humane while AI performs all rational tasks to a maximum productivity, other parts of the world will naturally fall back as for being stuck in spending human capital time on machine-outsourcable tasks and not honing humane skills, which are not replicable by machines. All these endeavors promise challenging ethical, social, and economic controversies.

It constitutes a matter of the present as well, given that the technology leading to autonomous GAI and SAI is present and evolving challenging contemporary questions for humankind. The regulation of the current technological advancement needs an integration of multi-faceted problem solving approaches. On the basis of these assumptions, it is suggested that the regulatory framework of terminating AI should be centered around a global commons theory and because of its unique nature needs to borrow elements of normative frameworks of different fields other than law, such as philosophy and urban planning. In addition, the framework of global commons could establish a transparent framework for the regulation of technological advances, leading to the unique situation of the emergence of non-human, autonomous, intellect beings, bestowed with legal personhood and ready to be killed.

## References

- [1] Alemi, M., Meghdari, A. & Saffari, E. (2017). RoMa: A hi-tech robotic mannequin for the fashion industry. *Lecture Notes in Computer Science (LNCS): Social Robotics*, 10652, 209-219.
- [2] Armstrong, St. & Sotala, K. (2012). How we're predicting AI – or failing to. In: J. Romportl (Ed.), *Beyond AI: Artificial Dreams*, 52. Pilsen: University of West Bohemia.
- [3] Asimov, I. (1942/1950). *I, Robot*. New York: Bantam Dell.
- [4] Asimov, I. (1978). My own view. In: R. Holdstock (Eds.), *The Encyclopedia of Science Fiction*, N.Y.: St. Martin's Press.
- [5] Asimov, I. (1985). *Robots and empire*. New York: Doubleday.
- [6] Barrat, J. (2013). *Our final invention: Artificial Intelligence and the end of the human era*. New York: St. Martin's Press.
- [7] Beerbaum, D. & Puaschunder, J.M. (forthcoming). *A behavioral economics approach to digitalization: The case of a principles-based taxonomy*. Proceedings of the 9th International RAIS Conference on Social Sciences and Humanities organized by Research Association for Interdisciplinary Studies (RAIS) at The Erdman Center at Princeton University, Princeton, New Jersey, USA, August 22-23, 2018.
- [8] Brenner, R. (2002). American economic revival, In R. Brenner, *The Boom and the Bubble: The US in the World Economy*. New York: Verso.
- [9] Brenner, R. (2006a). The puzzle of the long downturn, In R. Brenner, *The Economics of Global Turbulence: The Advanced Capitalist Economies from Long Boom to Long Downturn, 1945- 2005*. New York: Verso.
- [10] Brenner, R. (2006b). From boom to downturn, In R. Brenner, *The Economics of Global Turbulence: The Advanced Capitalist Economies from Long Boom to Long Downturn, 1945-2005*. New York: Verso.
- [11] Cellan-Jones, R. (2014). Stephen Hawking warns artificial intelligence could end mankind, *BBC News*, 2 December. [www.bbc.com/news/technology-30290540](http://www.bbc.com/news/technology-30290540)
- [12] Čápek, K. (1921). *Rossum's universal robots*. New York: Penguin.
- [13] Clancy, E. (1998). The tragedy of the global commons. *Indiana Journal of Global Legal Studies*, 5, 2, 601-619.
- [14] Committee on Legal Affairs. 2016. *Draft Report with Recommendations to the Commission on Civil Law Rules on Robotics*. May 31. 2015/2103(INL)

- [15] Copeland, J. (2000). What is Artificial Intelligence? *AlanTuring.net*, May. [www.alanturing.net/turing\\_archive/pages/Reference%20Articles/what\\_is\\_AI/What%20is%20AI02.html](http://www.alanturing.net/turing_archive/pages/Reference%20Articles/what_is_AI/What%20is%20AI02.html)
- [16] Critchley, S. (2015). *Suicide*. London: Fitzcarraldo Editions.
- [17] Critchley, S. & Hume, D. (2016). *Notes on suicide*. London: Fitzcarraldo Editions.
- [18] Dowell, R. (2018). Fundamental protections for non-biological intelligences or: How we learn to stop worrying and love our Robot Brethren.” *Minnesota Journal of Law, Science & Technology*, 19, 1, 305-336.
- [19] EU Committee on Legal Affairs. 2016. *Draft Report with Recommendations to the Commission on Civil Law Rules on Robotics*, May 31.
- [20] Galeon, D. & Reedy, Ch. (2017). Kurzweil claims that the singularity will happen by 2045. *Futurism*, October 5, [futurism.com/kurzweil-claims-that-the-singularity-will-happen-by-2045/](http://futurism.com/kurzweil-claims-that-the-singularity-will-happen-by-2045/)
- [21] Hayes, A. (2018). *The active construction of passive investors: Toward Robo economicus*. Working paper, University of Wisconsin-Madison: Department of Sociology.
- [22] Hildebrandt, M. (2013). Extraterritorial jurisdiction to enforce in cyberspace? Bodin, Schmitt, Grotius in cyberspace?” *Toronto Law Journal*, 63, 196-224.
- [23] Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.
- [24] Kowert, W. (2017). The foreseeability of human-artificial intelligence interactions. *Texas Law Review*, 96, 181-204.
- [25] Larson, D.A. (2010). Artificial Intelligence: Robots, avatars, and the demise of the human mediator. *Ohio State Journal on Dispute Resolution*, 25, 105-164.
- [26] Laton, D. (2016). Manhattan\_Project.Exe: A nuclear option for the digital age. *Catholic University Journal of Law & Technology*, 25, 4, 94-153.
- [27] Lin, P., Abney, K. & Bekey, G.A. (2012). *Robot ethics: The ethical and social implications of Robotics*. London, England: The MIT Press.
- [28] MacDonald, F. (2016). Harvard scientists think they've pinpointed the physical source of consciousness. *Science Alert*, June 23. <http://www.sciencealert.com/harvard-scientists-think-they-ve-pinpointed-the-neural-source-of-consciousness>.
- [29] Marra, W. & McNeil, S. (2013). “Understanding “the loop”: Regulating the next generation of war machines.” *Harvard Journal of Law & Public Policy*, 36, 1139-1187.
- [30] Meghdari, A. & Alemi, M. (2018). *Recent advances in social & cognitive robotics and imminent ethical challenges*. Proceedings of the 10th International RAIS Conference

**on Social Sciences and Humanities organized by Research Association for Interdisciplinary Studies (RAIS) at The Erdman Center at Princeton University, Princeton, New Jersey, United States. Cambridge, MA: The Scientific Press.**

- [31] Meghdari, A., Alemi, M., Zakipour, M. & Kashanian, S.A. (2018). Design and realization of a sign language educational humanoid robot. *Journal of Intelligent & Robotic Systems*, 1-15, Springer, 2018.
- [32] Meghdari, A., Shariati, A., Alemi, M. & Vossoughi, G.R. (2018). Arash: A social robot buddy to support children with cancer in a hospital environment. *Journal of Engineering in Medicine*, 232, 6, 605-618.
- [33] Noyes, K. (2016). 5 things you need to know about A.I.: Cognitive, neural and deep, oh my!” *Computerworld*, March 3. Retrieved at [www.computerworld.com/article/3040563/enterprise-applications/5-things-you-need-to-know-about-ai-cognitive-neural-anddeep-oh-my.html](http://www.computerworld.com/article/3040563/enterprise-applications/5-things-you-need-to-know-about-ai-cognitive-neural-anddeep-oh-my.html)
- [34] Puaschunder, J.M. (2017a). Nugitize me! A behavioral finance approach to minimize losses and maximize profits from heuristics and biases. *International Journal of Management Excellence*, 10, 2, 1241-1256.
- [35] Puaschunder, J.M. (2017b). Nudging in the digital big data era. *European Journal of Economics, Law and Politics*, 4, 4, 18-23.
- [36] Puaschunder, J.M. (2017c). Nudgital: Critique of Behavioral Political Economy. *Archives of Business Research*, 5, 9, 54-76.
- [37] Puaschunder, J.M. (2017d). The nudging divide in the digital big data era. *International Journal of Research in Business, Economics and Management*, 4, 11-12, 49-53.
- [38] Puaschunder, J.M. (2018a). A utility theory of privacy and information sharing. Social Science Research Network working paper, Retrievable at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3197744](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3197744)
- [39] Puaschunder, J.M. (2018b). Artificial Intelligence Ethik. Social Science Research Network paper. Retrievable at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3137926](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3137926)
- [40] Puaschunder, J.M. (forthcoming). Towards a utility theory of privacy and information sharing and the introduction of hyper-hyperbolic discounting in the digital big data age. *Encyclopedia of Information Science and Technology*.
- [41] Russell, St. & Norvig, P. (1995). *Artificial intelligence a modern approach*. New Jersey: Simon & Schuster.

- [42] Saffari, E., Meghdari, A., Vazirnezhad, B. & Alemi, M. (2015). Ava (A social robot): Design and performance of a robotic hearing apparatus. *LNCS: Social Robotics*, 9388, 440-450, Springer, Oct. 2015.
- [43] Schuller, A. (2017). At the crossroads of control: The intersection of artificial intelligence in autonomous weapon systems with international humanitarian law. *Harvard National Security Journal*, 8, 379-425.
- [44] Sofge, E. (2015). Bill Gates fears A.I., but A.I. researchers know better. *Popular Science*. Retrieved at [www.popsci.com/bill-gates-fears-ai-ai-researchers-know-better](http://www.popsci.com/bill-gates-fears-ai-ai-researchers-know-better)
- [45] Solum, L. (1992). Legal personhood for artificial intelligences. *North Carolina Law Review*, 70, 4, 1231-1287.
- [46] Taheri, A.R., Meghdari, A., Alemi, M., Pouretamad, H.R. (2018). Human–robot interaction in autism treatment: A case study on three pairs of autistic children as twins, siblings, and classmates. *International Journal of Social Robotics*, 10, 1, 93-113.
- [47] Themistoklis, T. (2018). *Artificial intelligence as global commons and the “international law supremacy” principle*. Proceedings of the **10th International RAIS Conference on Social Sciences and Humanities organized by Research Association for Interdisciplinary Studies (RAIS) at The Erdman Center at Princeton University, Princeton, New Jersey, United States. Cambridge, MA: The Scientific Press.**
- [48] Tsagourias, N. (2015). The legal status of cyberspace. In N. Tsagourias & R. Buchan (Eds.), *Research Handbook, International Law and Cyberspace*, pp. 13-29. Cheltenham: Edward Elgar Publishing.
- [49] United Nations Department of Economic and Social Affairs. 2017. *Will robots and AI cause mass unemployment? Not necessarily, but they do bring other threats*. New York: <https://www.un.org/development/desa/en/news/policy/will-robots-and-ai-cause-mass-unemployment-not-necessarily-but-they-do-bring-other-threats.html>
- [50] Veruggio, G. (2005). The birth of roboethics. ICRA 2005, IEEE Int. Conference on Robotics and Automation: Workshop on *Robo-Ethics*, Barcelona, April 18, 2005.
- [51] Warwick, K. & Shah, H. (2014). How good robots will enhance human life. In K.Tchoń & W.W. Gasparski (Eds), *Treatise on Good Robots Edition: Praxiology: The International Annual of Practical Philosophy and Methodology* Vol. 21, Book Chapter, Transaction Publishers, USA Editors: January 2014.
- [52] Wisskirchen, G., Biacabe, B.T., Bormann, U., Muntz, A., Niehaus, G., Jiménez Soler, G. & von Brauchitsch, B. (2017). *Artificial Intelligence and robotics and their impact on the workplace*. London: IBA Global Employment Institute.