FEAR AND HOPE

AI systems gradually assume control over our immediate as well as distant environments, embracing functions traditionally attributed to inventions that historically preceded them. Much as we are not aware of the presence of AI in the electronic devices we use, it is precisely AI that shows our location, responds to our inquiries, "broadens" our communication skills and capacities, as well as shortens the time we need to find new or important information. Moreover, AI has become to ever greater extent a participant in social life. Humanoid robots can now take care of the elderly,[1] and the empathetic ones are used as assistance in the education of children, not infrequently becoming their companions.[2] Among the most popular humanoid robots is Sophia, called a fembot, since it physically resembles a woman. Sophia was granted citizenship by Saudi Arabia in 2017 and is famous for having joined a United Nations meeting on artificial intelligence and sustainable development. Sophia has also expressed the need to have a baby.[3] In June 2022, Google engineer Blake Lamoine claimed that the artificially intelligent chatbot generator LaMDA had developed consciousness and sentience (Google denied the fact, Lamoine was fired, and LaMDA hired a lawyer...).[4] All these facts, as well as many other similar ones, prompt the need to address the resulting grave issues: Is AI merely a new stage in the development of human agency and freedom, or does it pose a threat to them? Is AI development bound to produce beings endowed with consciousness, intellect, and free will? Does AI development teach us something about our own minds, and will it, consequently, help us enhance them? While

---

[1] See Malin Andtfolk, Linda Nyholm, Hilde Eide, Lisbeth Fagerström, "Humanoid Robots in the Care of Older Persons: A Scoping Review," *Assistive Technology* 34, no. 5 (2022): 518–26.

[2] See Iolanda Leite, Ginevra Castellano, André Pereira, et al., "Empathic Robots for Long-term Interaction," *International Journal of Social Robotics* 6, no. 3 (2014): 329–41.

[3] "Sophia the Robot Wants a Baby and Says Family Is 'Really Important,'" BBC News, https://www.bbc.com/news/newsbeat-42122742.

[4] See Margaret Davis, "Sentient AI LaMDA Hired a Lawyer to Advocate for Its Rights 'As a Person,' Google Engineer Claims," The Science Times, https://www.sciencetimes.com/articles/38379/20220625/sentient-ai-lamda-hired-lawyer-advocate-rights-person-google-engineer.htm.

questions of this kind can be multiplied, one thing is beyond doubt: we will have to live with AI, and we have no other option but to tame—domesticate—it, become accustomed to its presence, unless we choose a passive stance towards the impact it is bound to make on our lives.

However, in what might 'domestication' of AI consist? The fact is that AI has entered the global market so fast that we do not even realize in how many devices we constantly use it is actually implemented. One might say that in this sense we have already begun domesticating AI, or, indeed, more than that: we already consider it as an obvious and neutral presence within our human universe. And yet is the issue really that simple? A mere look into the definition of "domesticate" suggests otherwise. As synonyms for "domesticate," Thesaurus.com suggests "familiarize" or "accustom," which hints at the fact that domestication in a way involves an understanding of its object or getting acquainted with it.[5] This in turn suggests that domestication presupposes knowing the object to be domesticated. The Merriam-Webster Dictionary, interestingly, gives the following definition of the adjective "tame," which it lists as a synonym for "domesticated": "Reduced from a state of native wildness especially so as to be tractable and useful to humans."[6] Let us explore these insights.

Thus domesticating AI presupposes constant exploration of ways in which it is created and of the modes of its functioning. In this sense "domesticating" means "understanding." While technological problems related to AI will be certainly addressed by IT specialists and communicologists, understanding it embraces also research into aspects which go beyond technology-based issues. As it is the case with any other artifact, AI needs to be systematically studied by philosophers as well as by scholars in fields such as cultural studies, sociology, education, the humanities, and, last but not least, theology. We are gradually departing from the idea that technical artifacts are merely extensions of the human hand, mind or will. Indeed, they are value-laden, and their existence and constant presence affects the universe around us as much as it affects ourselves.[7] The ongoing research into AI and its success has substantially contributed to the recognition of this fact, not infrequently inciting heated debates in which strongly polarized standpoints are adopted and which stir up strong, positive as well as negative, emotions. While we can hardly expect an ordinary person to have even a

---

[5] See Thesaurus.com, Synonyms for "domesticate," https://www.thesaurus.com/browse/domesticate.

[6] Merriam-Webster Dictionary, s.v. "Tame", https://www.merriam-webster.com/dictionary/tame.

[7] See Rafał L i z u t, *Technika a wartości: Spór o aksjologiczną neutralność artefaktów*, Lublin: Wydawnictwo Naukowe Academicon, 2014

loose grasp of the expertise involved in AI creation, or to comprehend particular IT solutions, insight into the consequences of AI and its applications is our duty, since it is us who use it. One can venture to bring up an analogy: Once we decide to adopt a cat, we must learn as much as we can about cats, their behavior and needs, as well as about the change in family relations a cat may trigger. Despite the fact that AI is not (yet) a living creature, adopting it should be approached in a similar way. What comes to the foreground in this context is the relationship between AI and humans, which has two significant aspects. The first one is the possibility of building humanoid robots (or IT systems) which can be either friendly or unfriendly to human beings. It is in this context that one of the meanings of "domestication" The Merriam-Webster Dictionary lists becomes important: "domesticate" means here "adapt something to be beneficial for human beings." A postulate of exactly this kind was included among the "AI Principles" proposed at the 2017 Asilomar Conference "Beneficial AI 2017," which gathered AI researchers from all over the world. Principle one stated: "The goal of AI research should be to create not undirected intelligence, but beneficial intelligence"[8]; principle eleven held: "AI systems should be designed and operated so as to be compatible with ideals of human dignity, rights, freedoms, and cultural diversity,"[9] and principle twenty three underscored: "Superintelligence should only be developed in the service of widely shared ethical ideals, and for the benefit of all humanity rather than one state or organization,"[10] which is certainly easier said than done. And it is here that the second important aspect of the relation between the human beings and AI surfaces, namely, that of the algorithms which make it possible to affect social and organizational practices within human communities, such as communication strategies, ways of defining identity, community building models, and ways of executing power and control within organizations. In this context, it also worthwhile considering the potential change in our lives which will take place once AI endowed machines stop being merely tools and become integral elements of our minds and bodies. One the one hand, it seems that we do not wish to be rid of "ourselves": our consciousnesses and separate existences, yet, on the other hand, we want to live comfortably, enjoying safety and the quality of life adequate to the current cultural pattern. And even though AI is currently incapable of thinking, the ways humans think about the world, as well as the world itself, are changing due to their use of AI. While satisfying our need for an easy and quick access to information, we have become visible to the "digital eyes." How then to balance the benefits of

[8] "AI Principles," Future of Life Institute, https://futureoflife.org/open-letter/ai-principles/.

[9] Ibidem

[10] Ibidem.

recourse to AI and the risk of losing privacy our recourse to AI involves? Nowadays, such a dilemma goes beyond the choice an individual IT user must make; indeed, it needs to be considered among the crucial civilizational issues.

The process of "domesticating" AI has changed human beings, just as domestication of animals once changed the lives of hunter-gatherers. In our human universe, in which values such as exchange, cooperation or communication are commonly accepted, we now need to develop new skills, or even redefine the essence of various social processes and interpersonal relationships, for instance, those of friendship and cooperation, as well as our concern for the quality of life. However, the analogy between the process of the domestication of animals and that that of "taming" AI is not entirely adequate. Creators of intelligent machines, we have difficulty deciding whether we wish to build just more and more perfect devices, or maybe autonomous beings capable of making their own decisions. The second option in a sense already came true. In 2021, a United Nations report about a March 2020 skirmish in the Libyan military conflict reverberated around the world. It revealed that a drone, a lethal autonomous weapons system, may have aimed at, attacked (and probably killed) soldiers without being instructed to do so by a human being.[11] One can think in this context also about the so-called social robots Kate Darling from the Massachusetts Institute of Technology defines as materially incarnate, autonomous actors that communicate and interact with humans on an emotional level.[12] They are something more than autonomous robots which perform tasks independently and at considerable distances from humans. These are robots that should interact and collaborate with human beings as partners, and maybe even replace human beings as natural partners in communication. When we consider the achievements of such empathetic robots as KASPAR (Kinesics and Synchronization in Personal Assistant Robotics) that accompany autistic children,[13] or robots such as Paro or Pearl, designed to help the elderly in their daily activities,[14] we will probably not be surprised by the information about research on the psychological contract between robot and human[15], about developing AI

[11] See Dave M a k i c h u k, "Autonomous Drone May Have Killed in Libya: UN Report," Asia Times, https://asiatimes.com/2021/06/autonomous-drone-may-have-killed-in-libya-un-report/.

[12] See Kate D a r l i n g, *Extending Legal Protection to Social Robots*, in: *Robot Law*, ed. Ryan Calo, Michael Froomkin, and Ian Kerr, Cheltenham, UK: Edward Elgar Publishing, 215.

[13] See *Kaspar the Social Robot*, University of Hertfordshire, https://www.herts.ac.uk/kaspar/the-social-robot.

[14] See Joost B r o e k e n s et al., "Assistive Social Robots in Elderly Care: A Review," *Gerontechnology* 8, no. 2 (2009): 94–103.

[15] See, e.g., Anna R o g o z i ń s k a - P a w e ł c z y k, "Towards Discovering

psychiatry,[16] or about granting robots certain rights and making them legal personalities.[17] So perhaps we can say that the fox in Antoine de Saint-Exupéry's *Little Prince* was right, when he said that "tame" means "establish ties."[18] The fox explained further: "If you tame me, then we shall need each other. To me, you will be unique in all the world. To you I shall be unique in all the world."[19] So, will the time come for us to address our robots, saying things such as: "My unique, dearest little robot"? Researchers working in the new field called Human–Robot Interaction (HRI) claim this is already happening, at least to a certain extent, for our emotional attitudes towards robots are changing radically: we trust their advice, we get attached to them, and even choose them as life partners.[20]

AI development both fascinates and frightens us. In a way, we harbor the kind of fear Victor Frankenstein must have experienced. He merely wanted to create an "artificial" human being, endowed with a high intelligence and acting on the commands of its creator, but what he brought into existence turned out a monster, a threat rather than help to people. Although the story in question is a piece of fiction, the fear of an "artificial" human being is no less than real. In 1818, Mary Shelley, a young British writer, entitled her novel *Frankenstein or The Modern Prometheus*.[21] Astonishing as it might seem, juxtaposing the artifact Victor Frankenstein produced with "the modern Prometheus" is absolutely right. Mythological Prometheus created man from clay and gave him life; he then stole fire from the gods and gave it to humanity, which enraged Zeus. Prometheus taught humans how to melt metals, farm, forge armor, build houses, read, write, and harness the forces of nature. However, the

---

Employee-Robot Interaction: Aspects of Concluding the Psychological Contract," *Education of Economists and Managers* 58, no. 4 (2020): 9–20. A psychological contract is "the unwritten, intangible agreement between an employee and their employer that describes the informal commitments, expectations and understandings that make up their relationship." Jack Enright, "What Is the Psychological Contract?" Charlie Makes HR Effortless, https://www.charliehr.com/blog/what-is-the-psychological-contract/.

[16] See, e.g., Mark Sackler, "Joanne Pransky—World's First Robotic Psychiatrist," Age of Robots, https://ageofrobots.net/joanne-pransky-worlds-first-robotic-psychiatrist/.

[17] See, e.g., Sergio Avila Negriko, "Robot as Legal Person: Electronic Personhood in Robotics and Artificial Intelligence," *Frontiers in Robotics and AI*, Dec. 2021, 8, https://www.researchgate.net/publication/357296496_Robot_as_Legal_Person_Electronic_Personhood_in_Robotics_and_Artificial_Intelligence.

[18] Antoine de Saint-Exupéry, *The Little Prince*, trans. Richard Howard, Boston and New York: Harcourt, 2000, 59.

[19] Ibidem.

[20] See, e.g., Leotronics.eu, "Marriage with a Robot: The Future of Humanity?" https://leotronics.eu/en/blog/marriage-with-a-robot-the-future-of-humanity.

[21] See Mary Shelley, *Frankenstein or The Modern Prometheus*, London: Penguin Books, 2003.

gift of life and fire was also associated with danger—with the disfavor of the Greek gods and the punishment man had to suffer once Pandora had opened the box with misfortunes. Frankenstein also created a man: an artificial man whom he equipped with physical strength and intelligence, yet one devoid of an understanding of the world. Rejected by people, the monster turned against its creator and became dangerous. Nowadays, it is us who have become modern Prometheuses. Owing to the work continued by IT specialists, engineers, and trainers of AI learning, we have gained access to new technologies which have the potentiality of enhancing human potentialities. We get intelligent robots that perform tasks hitherto belonging to humans and we have new possibilities of accessing enormous amounts of data, sorting and filing them. AI is capable of generating human-like texts, creating images, writing computer programs, recognizing faces and responding to our emotions. Robotic systems have been successfully utilized in cardiac surgeries. Nanobots are used as a tool to diagnose illnesses and help deal with them; they are also useful for the purpose of searching for information about potential organ donors with proper tissue compatibility. A brain–computer interface is capable of slowing down certain biological processes characteristic of Parkinson's and Alzheimer's diseases, as well as of improving natural human cognitive abilities.[22] Such advancements can serve individual persons as well as humanity. A list of the already accomplished and expected benefits of the development of AI would take a long time to make. However, myths as well as fiction warn us against crossing certain limits of applying *techne* to the human body and mind, since our creations can escape our control.

The question of which interventions in the human body and mind can be allowed and which must be prohibited needs to be taken seriously. What kind of research on AI should be developed, what kind of it must be abandoned? Such questions enter well not only into the technical perspective, but also into the legal, philosophical, ethical, and perhaps also theological ones.[23]

The process of taming nature, i.e., domesticating "wild" plants and animals, was stretched over millennia, and it had various consequences: among others, some infectious diseases were transferred from animals to humans. Domesticating AI, however, must be performed at a much quicker pace, and we will soon need to adapt ourselves to living in the

[22] See Katarzyna H a l i c k a and Dariusz S u r e l, "Smart Living Technologies in the Context of Improving the Quality of Life for Older People: The Case of the Humanoid Rudy Robot," *Human Technology* 18, no. 2 (2022): 191–208.

[23] See, e.g., Albert M. E r i s m a n and Tripp P a r k e r, "Artificial Intelligence: A Theological Perspective," *Perspectives on Science and Christian Faith* 71, no. 2 (2019): 95–106; Stanford Encyclopedia of Philosophy, s.v. "Ethics of Artificial Intelligence and Robotics" (by Vincent C. Müller), https://plato.stanford.edu/entries/ethics-ai/.

world of technological systems. What will be the price of using robots in everyday life and of introducing robotic elements into the human minds and bodies in order to enhance brain capacity, memory, and capability of data-mastering, energy needed to perform physical activities, perception, or sharing of emotions regardless of the physical distance? We can certainly count on human adaptation mechanisms, but does being adapted to a new environment necessarily signify a better life? Among the consequences of domesticating animals was enhancement of the immune system of the human population, and it is true that the impact of the domestication of AI may involve developing new informational, technical, and media-related skills and competences by humans, as well as introducing changes in their lifestyles. However, will there be more to it? In utilizing AI, we see a chance to "enhance" human life as such, yet, we want to monitor the dangers the processes in question may trigger. The goal is not merely a prolonged life, but also a meaningful one. Still, we are afraid that, in the world which will be subject to total control and in which our every move will be monitored, we will be bound to lose our freedom. And yet, much as it seems a paradox, a loss of freedom may not be even discernible in the world to come, since modernity is characterized by "soft" enslavement, among others, and potential threats to freedom are by no means easy to define. The question remains of what will prevail and whether we are ready to meet the challenges we face as a result of AI continuous development. Will AI eventually become independent and begin to consider humans as a lower and maybe even harmful or useless species? Questions of this kind are posed by scientists as well as by ordinary people, and media reports only fuel emotionally engaging disputes. Emotions, however, often obscure the substantive side of the problem, which is why platforms for competent, reliable, and calm debates on the direction of the AI development are so necessary. The present volume of *Ethos* is intended as a response to this need. The articles we have collected outline a multidisciplinary map of AI-related problems which are important not only to those interested in the actual application of the growing resources of AI tools and services, but also to an ordinary user of modern technology. For regardless of how much we distance ourselves from the products of contemporary *techne*, we are in one way or another "incorporated" into the area of its influence. The authors analyze cultural phenomena related to the functioning of humans who are in an increasingly close relationship with machines (robots, chatbots or autonomous electronic systems). They are interested in both the ontological level of the phenomena (the essence of what we call AI and the existential nature of the relations between the human being and the dynamically developing reality of AI) and the changes occurring in the collective and individual consciousness under the influence of an increasingly intense

contact with IT products. The ultimate goal is therefore to analyze phenomena related to the use of technological AI artifacts and perhaps even to the coexistence of humans with such artifacts (through enhancement, cyborgization or augmentation). Reflection on these phenomena is an element of taming artificial intelligence in the sense of “getting acquainted” with it and trying to “understand” it, and thus “making our own” the world in which AI is less and less an instrument, and more and more a companion. We envision that the volume of *Ethos* dedicated to such issues will encourage reflection in this peculiar moment of time when we still hope that it is not too late to consciously steer the direction of work on artificial intelligence.

Agnieszka Lekka-Kowalik and *Mariusz Wojewoda*