De Ethica (2025) 8.4:43–56
https://doi.org/10.3384/de-ethica.2001-8819.258443

Towards Artificial Wildlife: Speciesism and Intelligent Machines

Spyridon Stelios & Georgios Sakellariou

During the recent debates on the spread of artificial intelligence (AI), fears have been expressed about the possible dominance of intelligent machines over humans. In this article we explore the foundations on which this dystopian perspective of the threat to human sovereignty posed by the development of AI has been based. The research hypothesis is that the idea of the potential (future) supremacy of intelligent and ultraintelligent machines is based on the same intelligence-driven prejudice as represented by the theory of speciesism. If machines are seen as a new species, they not only differ from non-human animals but also follow a reverse course of development, regarding their relationship with man. Following this line of thought, we claim that a gradual 'undomestication' of machines is witnessed; a transition to a new state, which is further evidenced through the examination of the concepts of "autonomy" and "consciousness". Finally, it is argued that the theoretical background to study and interpret scenarios about machine dominance over humans is already here. It is applied in the field of environmental ethics.

Introduction

During the recent public debate about the spread of artificial intelligence (AI), fears have been expressed about the possible dominance of intelligent machines over humans. This dystopian perspective is not new, as it was developed in the 50s, almost in parallel with the theoretical birth of the interdisciplinary field of AI. The scenario of creating programs and machines with intelligence superior to human intelligence seems to have always been present in the minds of the creators of such programs and machines.

This article explores the foundations on which the dystopian perspective of the threat to human sovereignty posed by the development of AI has been based. The research hypothesis is that the potential (future) supremacy of intelligent and ultraintelligent machines is based on the same prejudice as represented by the theory of speciesism in the context of environmental ethics. Machines differ from non-human animals in that they have followed a reverse course of development regarding their relationship with humanity, even though various parallels have been drawn between the former and the latter. AI-based machines, unlike animals, seem to become 'wild' or 'undomesticated'

through the development of computer science. The phenomenon of the 'undomestication' of machines, which is further evidenced through the examination of the concepts of "autonomy" and "consciousness", is being discussed. This gradual process explains the growing fear of humanity about the evolving phenomenon of AI.

The artificial mind

The starting point for the research hypothesis is an examination of the nature of AI as discussed by two important philosophers long before this kind of technology infiltrated our daily lives and demonstrated the achievements it displays today. This investigation, included in this section, is important because it provides a theoretical background that enables the comparison between human and artificial intelligence. This way the attribution of a higher level of intelligence to machines makes sense.

According to Hilary Putnam, the fundamental nature of the mind cannot be mapped by AI. In his paper entitled "Artificial intelligence: much ado about not very much"[1], Putnam begins with the question "Does artificial intelligence teach us anything important about the mind?" answering that he is of the opinion that it does not. Maybe in the future it will teach something that matters regarding how human beings think. Even if that is the case though, why should people be so excited about AI now? Perhaps because what influences and excites humanity is the prospects deriving from AI, or, in other words, what might be possible through making use of AI, but not AI itself.

Today, and mostly because of the groundbreaking work of Alan Turing in the 1950s, the mind can be considered as a model of computational processes. As such, it is associated with AI. This computational mental model is, of course, not identical with or derived from AI. In fact, the idea of the mind as a kind of computational machine dates back to the seventeenth century.[2] Thus, today's computer science is not, in the first place, the same as artificial intelligence. In modern times, however, to get to AI, people must first get to and through computers. The modern digital computer is a functional implementation of the idea of a Turing machine, and its continuous and indisputable improvement, both in terms of software and of hardware, is not carried out by AI experts. Achievements such as, for example, the huge improvement in the ability of machines to play chess are mainly due to the findings of computer engineers, not due to the findings of AI researchers.

AI is a subcategory of computer design, with the latter belonging to the fields of engineering and programming. For Putnam, AI has become 'notorious' due to exaggerated claims, such as that it is a fundamental discipline and even that it is epistemology. So, one objective for computer engineers was, and most probably still, is to develop software that will allow computers to simulate or replicate the achievements of what is intuitively recognized by human beings as "intelligence". Putnam[3], however, puts forward a rather controversial position; even though AI has so far highlighted many issues of real interest to computer science in general, it has shed very little light on the function of the human mind. Indeed, the field of AI has not made any contributions beyond the issues which Alan Turing discussed in the early and mid-twentieth century.

[1] Hilary Putnam, 'Artificial Intelligence: Much Ado About not Very Much', *Daedalus*, 117:1 (1988), pp. 269-281.
[2] Putnam, 'Artificial Intelligence: Much Ado About not Very Much', p. 269.
[3] Ibid., p. 270.

Against the skepticism of Putnam and other philosophers about the potential of AI, Daniel C. Dennett, in his article entitled "When philosophers encounter artificial intelligence"[4], argues that what is essentially under discussion is a conflict between basic methodological assumptions. Most philosophers tend to view the results of AI with contempt similar to the contempt existing for those who persistently continue trying to square the circle or trisect the angle using a straightedge and a compass. There is evidence, they are told, that this cannot be done. Thus, in essence what they are told is "It cannot be done, so drop it!"[5] . Nevertheless, since a parallel cannot be drawn between evidence regarding AI and evidence about unsolved problems of geometry possibly related to a compass and a straightedge, the contemptuous responses in both cases are neither similar nor comparable; such evidence regarding AI consists of assumptions about 'reasonable' conditions. It contains idealizations that may be proven to be irrelevant to the squaring of the circle and the trisection of the angle, as well as to the case of the infamous aerodynamic laws, which are in essence biological assumptions, asserting that bumblebees cannot fly. What was the aforementioned case regarding the flying of bumblebees about?

From the early twentieth century, the ideas that, in aerodynamic terms, bumblebees could not fly or that their flight did not comply with the known aerodynamic laws were widely spread. The reason is that their wings are too small for their body size. However, since bumblebees have the ability to fly, there is obviously another underlying cause which enables them to do so, such as an undiscovered aerodynamic principle; and it was correctly assumed that, by managing to exploit that principle, it would be possible for scientists to create lift of aircraft with smaller wings or propellers. Therefore, engineers could not explain the biological flight mechanism of the bumblebee until the early 2000s, when the aerodynamics of that flight was solved by modern computational fluid dynamics, which accurately modelled the insect's flight. In fact, brute force rather than aerodynamic efficiency is the key to bumblebee flight.[6]

By the same token, according to Dennett, some philosophers argue that there can be no real intelligence with computers. It is like saying, "There is evidence that it cannot be done, so let it go, do not bother, it cannot be done with computers." Nonetheless, this certainty is not based on computer science. It is based on biological, psychological and other evidence. Consequently, this view can be somewhat arbitrary. Indeed, let us not lose sight of the fact that computer science is evolving, just as the ability to model the bumblebee flight has evolved.

The presentation, perhaps even conflict, of these two important philosophers' views does not bring out a clear answer as to the value of AI in relation to the mind. Nevertheless, it is useful as it represents an in-depth analysis of the supposed nature of AI. In general, Putnam argues that the human mind cannot be mapped with non-biological tools such as computing systems. Dennett argues that this confidence is methodologically wrong as it applies biological evidence to the field of computer science. In other words, this certainty, if it can be stated, should be stated by a computer scientist. These two views are still relevant to the present AI debate concerning the possibility of AI causing human

[4] Daniel C. Dennett, 'When Philosophers Encounter Artificial Intelligence', *Daedalus* 117:1 (1988), pp. 283-295.
[5] Dennett, 'When Philosophers Encounter Artificial Intelligence', p. 292.
[6] University of Oxford, *Flight of the bumble bee is based more on brute force than aerodynamic efficiency*, ScienceDaily (10 May 2009), online at https://www.sciencedaily.com/releases/2009/05/090507194511.htm (accessed 2023-05-17).

extinction. They conceptually bridge artificial and biological intelligence making it possible to formulate dystopian scenarios (see next section) as included in the AI debate. In addition, both different approaches can be reconciled based on the following argument:

> (1) Within computer science, a form of (computational) intelligence is currently being developed.
>
> (2) Computers define our lives.
>
> Therefore, developing computer intelligence defines our lives.

Evidence of this defining impact of computational intelligence on our lives is almost everywhere. Apart from increasingly employed in various businesses and industries, AI is now being developed to make everyday life a "piece of cake". Intelligent machines are instrumental in optimizing the way we interact with each other, entertain ourselves, and complete various tasks.

The dystopian scenario

This new developing kind of intelligence, which already defines our everyday life, can lead to very high levels of computational processes. At this point the term "ultraintelligent" makes its appearance, denoting an overall intelligence much higher than that of humans, and therefore disproportionately powerful. Ultraintelligence will be gradually mastered, by means of new levels of AI achieved through recursive self-improvement. This rather dystopian perspective seems to have been founded as early as the 1960s. According to I. J. Good:

> Let an ultraintelligent machine be defined as a machine that can far surpass all the intellectual activities of any man however clever. Since the design of machine is one of these intellectual activities, an ultraintelligent machine could design even better machines; there would then unquestionably be an "intelligence explosion," and the intelligence of man would be left far behind.[7]

Ultraintelligent machines would, in this context, be superior to humans. This scenario was also implied by Alan Turing. Turing with his "Heads in the Sand" objection[8] notes that, if there were thinking machines, then most possibly they could think much better than people can. Based on that, the possibility that humans might be supplanted or dominated by machines would become a genuine worry. This is not an argument against the claim that machines can think. It represents various fears about what might follow if thinking machines existed.

Bostrom[9] supposes that an ultraintelligent or superintelligent machine wants to take control of the world. Would it be able to do so? Most probably it would, if it developed certain powers that could enable it to achieve global dominion. The necessary condition for these powers is a level of intelligence higher than human intelligence. The human being is the benchmark since humanity's dominant position on Earth is based on our brain's

[7] Irving John Good, 'Speculations Concerning the First Ultraintelligent Machine', *Advances in Computers* 6 (1966), pp. 31–88, at p. 33.
[8] Alan Mathison Turing 'Computing Machinery and Intelligence', *Mind* 59 (1950), pp. 433– 460, at p. 444.
[9] Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies*, Oxford, UK: Oxford University Press (2014), p. 9.

slightly expanded set of faculties in comparison with other animals. Artificial entities with that level of intelligence could accumulate content faster than human beings, making accomplishments such as new inventions much faster than humans. The worst-case scenario is that artificial entities may make use of their intelligence in order to implement strategies against humans.

This potentially dangerous kind of intelligence, resulting from silicon-based forms of organization of matter, is different from carbon-based human intelligence. Thus, scientists can regard ultra-intelligent machines – alongside intelligent machines – as a new "species" which is artificial, but at the same time shares with human beings the common characteristic of cognition. On the basis of this, a prejudice can then be formed concerning the superiority of machines over humans. Such a prejudice could lead to dystopian scenarios that have already been presented in the fields of literature and cinematography. These superior artificial beings, on the assumption that they are autonomous and/or self-conscious, will serve their own interests, potentially threatening the human interests. They will follow the successful, in evolutionary terms, example of humans versus animals, when the former found themselves to be in a position of power.

Speciesism

This prejudice is reminiscent of the theory of speciesism,[10] which is studied in the context of Environmental Ethics. Speciesism refers to viewing members of one species as morally more important than members of another species with regard to both species' similar interests. It is regarded as discrimination and unjustified treatment, based on an individual's membership. It is differential treatment, without taking into consideration whether such treatment is justifiable or not. It is actually a prejudice or attitude of partiality. Speciesists allow the interests of their own species to override the greater interests of members of other species.[11]

For instance, humans demonstrate speciesism when they consider the interests of non-human animals to be less important than the similar interests of humans. It is certainly not speciesist to believe that in some cases human life is more important than that of non-human animals; what is speciesist is to argue that always, in all cases, human life is more important. Besides, it is not speciesist to declare that apes should teach physics; apes do not have that ability. However, it is speciesist to argue that apes should be used in medical experiments that are not absolutely necessary, simply because they are not human beings.[12] The corollary of speciesism is that humans have the right to use non-human animals.

What is this prejudice caused by? It is unquestionably caused by the existence of the unique human mind, which is substantially different from the mind of other species. Even though there are degrees of self-awareness in other animals, the advantage of human intelligence over the intelligence of non-human animals is the realization and use of abstract concepts. Furthermore, humans' different natural intelligence and consciousness seem to be inextricably tied to the biological substrate called "brain". So, this prejudice can

[10] Peter Singer, 'All Animals Are Equal', *Philosophic Exchange* 5:1 (1974), pp. 103-116.
[11] Singer, 'All Animals Are Equal', p. 108.
[12] Sarah Grey and Joe Cleffie, 'Peter Singer's race problem', *Jacobin* (8 June 2015), online at, https://jacobin.com/2015/08/animal-rights-cecil-the-lion-peter-singer-speciesism (accessed 2023-05-17).

be attributed to the intelligence and ability of humans, among other things, to visualise themselves as existent in the past, present and future. Nevertheless, intelligence - regarded as the ability to deal with various situations, use reason, apply knowledge to one's environment and to think in an abstract way - is not the only criterion for distinguishing between human and non-human animals. This is further reinforced by the fact that, although there are humans with severe intellectual disabilities, we still consider them to be superior to non-human animals. Therefore, the life of a human being, regardless of her/his intellectual abilities, is always of greater value in comparison to the life of the most intelligent non-human animal. However, it is worth mentioning that intelligence remains a key ingredient that lays the foundations of the concept of speciesism.

Domestication

An example of the realization of humans' speciesism is the domestication of animals. The wolf (that is, after its domestication, the dog), the sheep and other animals considered to have weaker mental abilities than humans were gradually brought under human control. For Shipman,[13] the initial mainspring of animal domestication seems to have been not only primary products (i.e. meat) but also secondary products such as milk, wool, manure, traction (dragging ploughs in agriculture), pack transport, riding, etc. By creating domesticated animals, human beings have in effect created a kind of living tools. In this way, domestic animals represent a kind of human exosomatic adaptation. They are tools used with the objectives of expanding the resources humans can exploit and of them further cultivating their abilities.

It is evident that, regarding the relationship between humans and animals, the effects of animal domestication have not particularly benefited the animals themselves. More specifically, before ecoculturalism (domestication), within traditional hunter-gatherer societies, where humans and animals shared the environment and its resources, the human-animal relationship was often one of mutual trust - and actually the animals hunted by humans were regarded by them as equals. However, due to the advent of domestication, that human-animal relationship was transformed into a relationship of human domination and control over animals. Humans assumed the role of the master/owner and animals assumed the role of property. Animals were transformed into assets, into objects, so that they could be owned, inherited and exchanged.[14] Through domestication, ultimately, humans affirmed the superiority of the human species.

How are speciesism and domestication of animals related to the dystopian scenario of the potential superiority of intelligent machines? If intelligent machines are regarded as a new 'species' which on the one hand is artificial and on the other hand has the common feature of cognition, then a new quasi-speciesism is likely to be formed. In other words, a prejudice advocating the superiority of machines over humans may emerge. It is a fictitious speciesism recognized by the potential victim and not (yet) by the potential perpetrator. It is not about the practice of treating members of one species as more important, in moral terms, than members of other species, but it is an outlook that is born from the impressive development of technology. For this reason, it is almost a specialty. Based on that, the artificial superior beings will serve their own interests, alongside with

[13] Pat Shipman, *The Animal Connection*, New York: W. W. Norton & Company (2011), pp. 250-252.
[14] Tim Ingold, 'From Trust to Domination: An Alternative History of Human Animal Relations', in *Animals and Human Society: Changing Perspectives*, edited by Aubrey Manning and James Serpell. London: Routledge (1994), pp. 1-22, pp. 6-9.

threatening the interests of the human species. They will follow the successful, in evolutionary terms, example of humans versus animals. In this case, however, humans will be the beings that will be domesticated. Following this line of thought, the question of how much humans are justified in treating machines worse than themselves just because machines do not belong to the same species emerges. One way of answering that is by addressing the issue of a conscious or self-conscious artificial being, which is discussed in the next section. Be that as it may, what the quasi-speciesism hypothesis suggests is that, within environmental ethics, we could ascribe morally relevant interests to an entity regardless of whether it has consciousness.

Speciesism based on the level of intelligence lays the foundations for the fears expressed about the predominance of machines over humans. It is possible that the deepest cause of human fears towards the evolution of AI comes from the example of domesticated animals. Being aware of the implications of animal domestication on animals, humans are fearful of the implications of machines of intelligence superior to theirs domesticating them. The example of the sheep and other animals raises the possible prospect of our own "domestication" by machines, which will occur in a context uncontrollable by humans. For the first time in thousands of years, the interests and primacy of the human species are now being threatened, shaping the moral attitudes humans have and will most possibly continue to have towards intelligent machines.

It should not be overlooked that, today, human beings instrumentally use AI applications in the same way they exploit domestic animals. In this sense, "intelligent" machines can be subsumed into the broader category of domesticated beings. It is apparent that such machines are artificial beings and not natural ones, but they are similar to the latter in the sense that they possess a kind of intelligence as well.

The concept of "autonomy" is extremely important. Intelligence is a personal trait attributed to an autonomous agent. Therefore, autonomy is the criterion on which the distinction between 'wild'/ 'undomesticated' and 'tame'/ 'domesticated' environment is based. Creatures living in the wild are characterized by autonomy regarding both their existence and functioning. In this context, what is of particular interest is that today the degree of autonomy of an intelligent agent is a pursued goal and is also proportional to that system's value and usefulness.

Autonomy and consciousness

Autonomy, seen broadly as the condition of self-government and/or self-control, is an important factor in the AI debate. If a machine demonstrates autonomy in its operations, then it can take conscious action towards an end and exhibit moral behavior. For instance, it may regard some kind of human behavior or action as bad and, as a consequence, it is likely to react against this behavior by harming the person who has been the agent of these. It becomes evident that autonomy, being an additional condition of behavior and decision-making, is a characteristic of paramount importance. This is because autonomy determines issues of accountability between computer engineers and the machines they create.

Autonomy has also been the criterion for distinguishing between a natural and an artificial state. Natural beings seem to have autonomy in their existence and functioning. According to Aristotle's Physics,[15] natural things differ from things which are not natural

[15] Aristotle. *Complete Works vol. 1, The Revised Oxford Translation*, edited by Jonathan Barnes. Princeton, N.J.: Princeton University Press (1991), 192b24-192b32

in that they have within themselves a principle of motion and of stationariness. They are born and changed by an internal power, which is exclusively their own. an energy that exists within them. Artificial products, such as houses and other products of manual labour, have not got within themselves the principle of their own reproduction. That principle exists, but in something else external to the artificial thing. Artificial things owe their creation, their "birth", as well as any change they undergo to the craftsman, who is a mental and physical actor existing outside of them.

Dennett[16] offers a particularly enlightening account of the concept. Autonomy is a real situation which can be defined quite clearly through physics and engineering. The concept that unlocks its decoding is the "degrees of freedom". Each degree of freedom, that is, each independent displacement or motion of a (mechanical) system, represents an opportunity for control. For example, human beings have a lot more degrees of freedom in their fingers, in comparison to humanoid robots, like Honda's ASIMO or Engineered Arts' Ameca. However, human beings do not have as many degrees of freedom as there are in an octopus tentacle. If there are more degrees of freedom than buttons for controlling them, then those degrees of freedom are going to be either out of control or remotely controlled. Following this line of thinking, autonomy means self-control, in contrast with being out of control or under remote control.

Remote control is exercised by agents that have purposes, plans, etc. For Dennett[17] agency is an "intentional system". A man, machine or alien creature is an intentional system in relation to the strategies of someone who tries to explain and predict its behavior. Only beings capable of intentional states are agents; this means that both people and non-human animals such as cats or dogs are agents, thanks to being capable of intentional states and of performing acts. Moreover, intentional explanation and prediction of the behavior of artificial entities such as today's chess-playing computers or GPT chatbots is common. Human beings seem to treat these programs as intentional systems, and they do so regardless of any considerations about what substance these programs are composed of, about their origin, their consciousness or self-consciousness or even the determinacy or not of their operations.[18]

In contrast, trees are not agents. Trees grow and drop leaves, but this is not something that happens as the result of an action on their part.[19] Moreover, trees are not remotely controlled. Even though the falling of autumn leaves from the trees is not controlled by the wind or gravity, it is caused by the wind and by gravity. Gravity or the wind are not agents, as they do not make the leaves go where they go because of certain plans or goals. They cause the falling, but do not control it. Thus, falling leaves are out of control.[20]

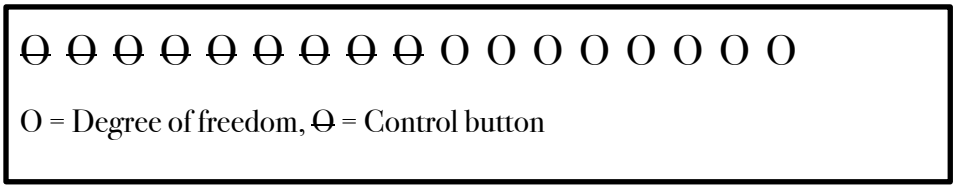
Based on the above, autonomy means having a greater number of control buttons on your degrees of freedom than the number of degrees of freedom is (see Scheme 1).

---

[16] Daniel C. Dennett, 'Autonomy, Consciousness, and Freedom', *The Amherst Lecture in Philosophy* 14 (2019), pp. 1–22, at pp. 1-2.
[17] Daniel C. Dennett, *Brainstorms: Philosophical Essays on Mind and Psychology*, Fortieth Anniversary Edition, Cambridge, MA: MIT Press (2017), p. 3.
[18] Dennett, *Brainstorms: Philosophical Essays on Mind and Psychology,* p. 7.
[19] Kenneth Einar Himma, 'Artificial Agency, Consciousness, and the Criteria for Moral Agency: What Properties Must an Artificial Agent Have to be a Moral Agent?' *Ethics and Information Technology*, 11:1 (2009), pp. 19-29, at p. 20 as cited in David J. Gunkel, *The Machine Question: Critical Perspectives on AI, Robots, and Ethics*, Cambridge, MA: The MIT Press (2012), p. 19.
[20] Dennett, 'Autonomy, Consciousness, and Freedom', p. 5.

Today, machines through machine learning and deep learning gain more self-control over their degrees of freedom. Machine learning includes algorithms that parse data and learn from them, which leads to informed decisions. Deep learning, a subfield of machine learning, creates through algorithms a neural network that learns and makes decisions on its own. The issue under discussion is not the greater capabilities that AI systems acquire, but the self-control these systems have over them.

***Scheme 1**: Autonomy based on degrees of freedom*



Regarding consciousness, all of us can think of reasons why humanity cannot speak of consciousness acquired by machines. Although, current computers might be regarded superior to humans in the sense that they can calculate much faster than humans, such abilities do not make them self-conscious beings. This 'obvious' lack of (self) consciousness retains the machines' practical nature: they are just tools used by humans. Furthermore, any comparison at first sight between human and (assumed) machine consciousness seems meaningless. What kind of consciousness does an AI-based system have when hundreds of hours of work and a multi-member team of engineers are hidden behind it? However, Putnam[21] would disagree with the aforementioned assumption. In his view, the arguments that support this assumption have a rather hidden and devastating flaw. They are based on only two premises concerning robots: a) that machines/robots are artifacts, that is, artificial (rather than natural) beings, and b) that they are deterministic systems whose behavior and intelligence are preselected and designed by humans. Nevertheless, the two qualities mentioned in the premises above can also be human qualities.

First and foremost, regarding the first premise, human beings are machines as well, just machines of a different physical type. All human beings are created by their parents and determined, in both form and function, by instructions contained in the genetic code. All children are produced from the building blocks of their parents' DNA. In this context, of course, one could argue that what distinguishes us from machines is that humans have free will · we can make decisions of our own. However, this viewpoint is not accepted by everyone. There exist various a priori, empirical, scientific arguments and theories for and against that thesis.[22] Regarding the second premise, this deterministic dimension does not take into account all possible kinds of learning mechanisms (see e.g., the machine learning/deep learning method). Machines that are designed to learn, do not just do what they were pre-programmed to do. They often come up with unique solutions that surprise even their developers.[23] So, the question of whether computers and machines can have consciousness or even be self-conscious seems to be downplayed.

---

[21] Hilary Putnam, 'Robots: Machines or Artificially Created Life?' *Journal of Philosophy* 61:21 (1964), pp. 668–691.

[22] For a comprehensive overview: Timothy O'Connor and Christopher Franklin, 'Free Will', in *The Stanford Encyclopedia of Philosophy* (Winter 2022 Edition), edited by Edward N. Zalta and Uri Nodelman, online at, https://plato.stanford.edu/archives/win2022/entries/freewill/ (accessed 2024-08-31).

[23] Gunkel, *The Machine Question: Critical Perspectives on AI, Robots, and Ethics*, p. 53.

The main issue about consciousness lies in the concept of "consciousness" itself. It is not so much whether machines and animals are conscious or not. This will most possibly remain a contentious issue, with each side offering theoretical and practical arguments for and against it respectively. The principal issue regarding consciousness is the fact that this discussion proceeds and persists with a rather flexible and not entirely consistent definition of consciousness being existent.[24] Even though human beings have an intuitive understanding about what is an objective and widely accepted definition of "consciousness", this definition cannot yet be adequately framed in cognitive sciences and computation. So, human understanding of consciousness does not seem to be sufficient when humans try to translate it into a machine or answer about whether a certain agent, be it artificial or not, can be said to fulfill this kind of understanding.

So, are today's intelligent machines agents like animals? In other words, do they have the autonomy - and not necessarily the consciousness, as there are significant difficulties in defining this concept - to act on the basis of judgments that could even include attributing the quality of "evil" to the human species? If the answer is yes, then non-human animals, humans and machines may be placed on the same level of possible intentional states. In this case, the hypothesis of reverse paths (see below) becomes more valid. In other words, it is sufficient that the artificial system is understood as if it had autonomy and agency. The question, finally, of whether the 'reversed paths scenario' would still apply if the system were not conscious, cannot be adequately addressed due to the difficulty in defining consciousness.

Reverse paths

Machines and animals follow reverse paths. Machines, unlike animals, commence from a human, controlled environment of action and seem to become 'wild' or 'undomesticated' through the further development of AI. As already mentioned, this artificial wildness does not necessarily presuppose that intelligent machines will become self-conscious. Achieving a sufficient level of autonomy and thus, of 'undomestication' (or else, of 'untaming'), they could stop providing assistance and necessary resources to their human creators and start making use of them. They would now assume the role of the owner, while humans would have the role of property. Some signs, if we consider them as such, of this development are manifested even today. For example, an AI chatbot's responses that resembled to implied threats against the human user.[25]

Let us see in greater detail the way in which these two proposed processes are formed and are evolving. First, animal domestication may be presented through a three-step process (see Figure 1). Undomesticated animals compete against humans in the natural environment (see for instance, wolves and hunter gatherers almost 30,000 years ago). By losing their autonomy, they assist humans, who eventually use them. Second, regarding intelligent machines, as long as artificial intelligence resembles more and more the human soul with regard to its function, the differences that place man and machine in different (social) contexts start to diminish. The question that arises is "When will there be convergence between the artificial and natural functional framework?" The answer is that

[24] Ibid., p. 54.
[25] Billy Perrigo, 'The New AI-Powered Bing Is Threatening Users. That's No Laughing Matter', *Time* (17 February 2023), online at, https://time.com/6256529/bing-openai-chatgpt-danger-alignment/ (accessed 2024-08-31).

convergence will occur when intelligent machines enter a *machinecentric* (instead of anthropocentric) ethical system (i.e. follow an ethical course) to the extent that at an initial stage the machine itself does not need human craftsmanship anymore and puts machine survival and development at the forefront. This will be the first autonomy step of the existence of machines, placing them in a reverse process of "undomestication". In this context, they could begin exploiting humans. Maybe they see humans as a threat, the same way a wild animal does! In the wild, a basic feature of animals' stance is to defend their interests. So, the second stage of the artificial 'undomestication' might be the prioritization of the interests of the machine, initiating the transition to a new state, which is founded upon the application of the 'criterion of wildness'. This development leads to the third and final stage.

***Figure 1**: Animal domestication and artificial 'undomestication'*

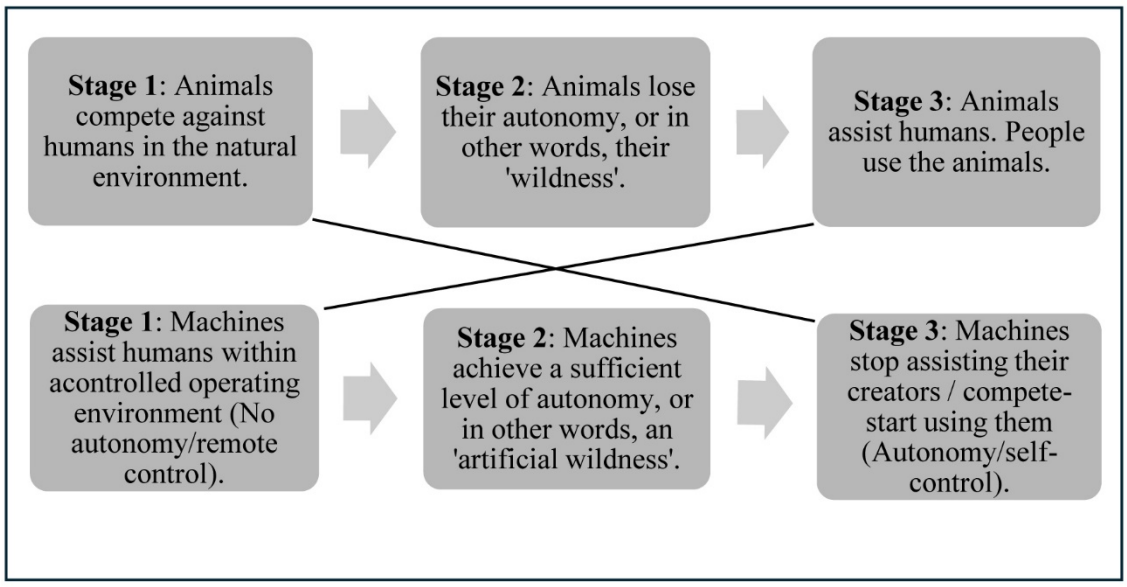
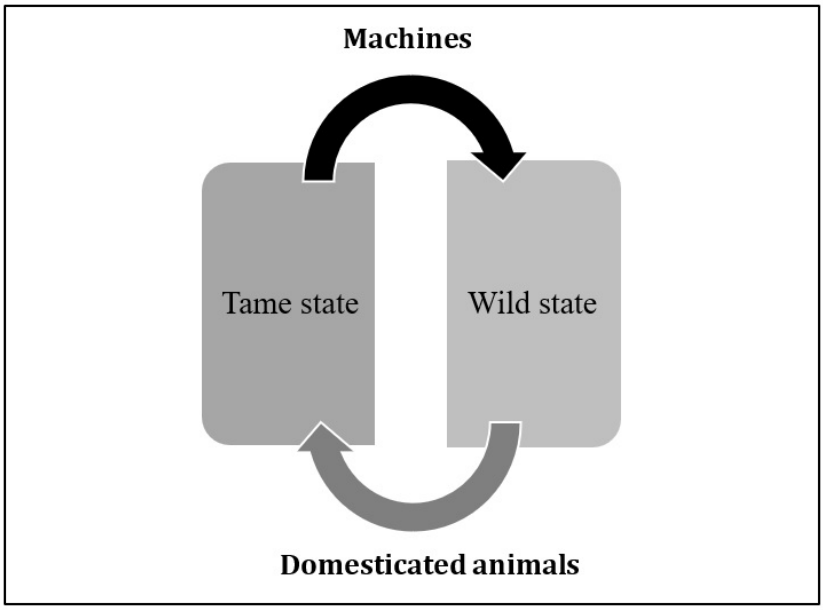


Figure 2 shows in a simpler way the historical switching of states that occurs between nonhuman animals and intelligent machines.

Figure 2: *The 'evolution' of machines' and animals' relationship with humans*



Conclusion

Recently, especially after the release of large language models, such as ChatGPT, there has been a lot of discussion about the possibilities of AI. This public debate is, for the most part, based on concerns about the increasing power and social integration of AI-based machines. These machines and programs have already begun to assimilate into human society and are charged with "stealing jobs", as well as creativity from the human species. Moreover, it is undoubtedly true that this academic, journalistic and social concern has erupted somewhat suddenly. While there were, of course, prior studies pointing to possible unpleasant side effects, it was only when the vast majority of people saw the great potential of these chatbots that human beings began to worry. Until then, these AI-based programs were like people's pets and domesticated animals. They were the dogs that brought people the newspaper or the cows plowing the field instead of humans.

Only when an AI program started to "show its teeth", like a wild dog, did humans start to fear en masse and consider their next moves. The artificial wild dog is still on a leash. It is not yet among human beings, so humans do not have to hide. It cannot "smell" human fears and thus, adjust its behavior. It is limited, confined to some cyberspace corner that human beings can still observe and control. What humanity needs to do now is think about how to keep that leash as a safety measure and understand the wild nature of the new being humans themselves have allowed to develop. The reverse paths hypothesis offers a common framework for analyzing these concerns. Humans, machines, and animals as agents or non-agents are united under the umbrella of high-low intelligence, forming a wide web of relationships that change over time. As long as computational intelligence defines our lives, the stages of domestication and 'undomestication' can be realized.

The theoretical background to study and interpret dystopian scenarios about machine dominance over humans is already here. It is applied in the field of ethics and in particular, of environmental ethics. Environmental ethics and the theory of speciesism provide a closer look at this kind of scenarios; and it is especially useful to know how to

deal with such hypothetical situations that seem to be increasingly penetrating the collective consciousness of humanity.

Even if the intelligence developed through computers does not represent true human intelligence, this does not seem to mitigate human concerns about its future potential. Today, due to machine learning and deep learning, programs gain more self-control over their degrees of freedom. Additionally, regarding whether these programs are conscious or not, this is not a matter of concern. Autonomous or not, conscious or not, AI models have undeniably enormous computing powers which, in instrumentalist terms, resemble the functions of human intelligence. A quasi-speciesism based on the level and nature of that intelligence underlines and explains human fears about the domination of machines over people. It is true that some humans fear that intelligent machines will treat us in the same way we treat domesticated animals. This argumentation is rather psychological and not so much an ethical one. Nevertheless, this explanation has clear ethical implications since it can reflect the future ethical treatment of machines through the ethical treatment of animals. The possibility of an artificial wilderness, however unlikely it may seem, inevitably redefines the human-machine roles and correlations involved in the moral universe of humanity.

One question that arises from the above analysis is the following: Could it be that, through the wild nature of machines and due to it, humans are gradually losing self-control regarding the degrees of freedom they have? Furthermore, does this new artificial species gain all the autonomy that humans lose? These are questions mainly related to the degree of human interaction with AI. As long as this technology is integrated into every aspect of people's lives, humanity will be concerned with such issues. One thing is certain. This exciting technology is here to stay, changing the way we, the human species think about ourselves and our future.

Spyridon Stelios, Department of Humanities, Social Sciences and Law School of Applied Mathematical and Physical Sciences, National Technical University of Athens/Hellenic Open University, stelioss@mail.ntua.gr

Georgios Sakellariou, School of Humanities, Social and Education Sciences, European University Cyprus/Hellenic Coastguard Academy/Hellenic Navy NCO Academy, giorgsakellar@hotmail.com

Bibliography

Aristotle. *Complete Works vol. 1, The Revised Oxford Translation*, edited by Jonathan Barnes. Princeton, N.J.: Princeton University Press (1991).

Perrigo, Billy. 'The New AI-Powered Bing Is Threatening Users. That's No Laughing Matter', *Time* (17 February 2023), online at, https://time.com/6256529/bing-openai-chatgpt-danger-alignment/ (accessed 2024-08-31).

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

Dennett, Daniel C. 'Autonomy, Consciousness, and Freedom', *The Amherst Lecture in Philosophy* 14 (2019), pp. 1–22.

Dennett, Daniel C. *Brainstorms: Philosophical Essays on Mind and Psychology*, Fortieth Anniversary Edition. Cambridge, MA: MIT Press (2017).

Dennett, Daniel C. 'When Philosophers Encounter Artificial Intelligence', *Daedalus* 117:1 (1988), pp. 283-295.

Good, Irving John. 'Speculations Concerning the First Ultraintelligent Machine', *Advances in Computers* 6 (1966), pp. 31–88.

Grey, Sarah and Cleffie, Joe. 'Peter Singer's race problem', *Jacobin* (8 June 2015), online at, https://jacobin.com/2015/08/animal-rights-cecil-the-lion-peter-singer-speciesism (accessed 2023-05-17).

Gunkel, David J. *The Machine Question: Critical Perspectives on AI, Robots, and Ethics*. Cambridge, MA: The MIT Press (2012).

Himma, Kenneth Einar. 'Artificial Agency, Consciousness, and the Criteria for Moral Agency: What Properties Must an Artificial Agent Have to be a Moral Agent?' *Ethics and Information Technology*, 11:1 (2009), pp. 19-29.

Ingold, Tim. 'From Trust to Domination: An Alternative History of Human Animal Relations', in *Animals and Human Society: Changing Perspectives*, edited by Aubrey Manning and James Serpell. London: Routledge (1994), pp. 1-22.

O'Connor, Timothy and Franklin, Christopher. 'Free Will', in *The Stanford Encyclopedia of Philosophy* (Winter 2022 Edition), edited by Edward N. Zalta and Uri Nodelman, online at, https://plato.stanford.edu/archives/win2022/entries/freewill/ (accessed 2024-08-31).

Putnam, Hilary. 'Artificial Intelligence: Much Ado About not Very Much', *Daedalus*, 117:1 (1988), pp. 269-281.

Putnam, Hilary. 'Robots: Machines or artificially created life?' *Journal of Philosophy* 61:21 (1964), pp. 668–691.

Shipman, Pat. *The Animal Connection*. New York: W. W. Norton & Company (2011).

Singer, Peter. 'All Animals Are Equal', *Philosophic Exchange* 5:1 (1974), pp. 103-116.

Turing, Alan Mathison. 'Computing Machinery and Intelligence', *Mind* 59 (1950), pp. 433– 460.

University of Oxford. *Flight of the bumble bee is based more on brute force than aerodynamic efficiency*. ScienceDaily (10 May 2009), online at https://www.sciencedaily.com/releases/2009/05/090507194511.htm (accessed 2023-05-17).