Review Article

Superintelligence: Identification of Friend or Foe Future Landscape of Cooperation with Non-human Intelligence

David Josef Herzog[1], Nitsa Herzog[2]

1. Department of Computing, QA Higher Education, United Kingdom; 2. Faculty of Computer Science, Northumbria University, Newcastle upon Tyne, United Kingdom

One of the most studied attributes of mental activity is intelligence. While non-human consciousness remains a subject of profound debate, non-human intelligence is universally acknowledged by all participants of discussion as a necessary element of any consciousness, regardless of its nature. Intelligence can potentially be measured as processing or computational power and by problem-solving efficacy. It can serve as a starting point for reconstructing arguments related to Artificial Consciousness. The shared modus of intelligence evaluation, irrespective of its origin, offers promising direction towards the more complex framework of non-human consciousness assessment. However, this approach's successful resolution of an objective basis for intelligence studies unveils inescapable challenges. Moreover, when the potential for non-human intelligence exists in both biological and non-biological domains, the future of the relationship between humankind, as the possessor of human intelligence, and other intelligent entities remains uncertain. This paper's central inquiry is focused on comparing purely computational capability to general, universal intelligence and the potential for higher intelligence to exert adverse effects on less intelligent counterparts. Another question is related to the degree of importance of the particular architectural characteristics of intelligent systems and the relationship between computing elements and structural components. It is conceivable that pure intelligence, as a computational faculty, can serve as an effective utilitarian tool. However, it may harbour inherent risks or hazards when integrated as an essential component within consciousness frameworks, such as autopoietic systems. Finally, an attempt has been made to answer the question concerning the future of interactions between human and non-human intelligence.

 doi.org/10.32388/QC239X

1. Introduction

Intelligence can be regarded as a property of any data processing substrate. Quality and speed of data processing can be measured, while operations and behaviour can be observed. Cognitive functions require substrate, biological or non-biological, and the nature of this substrate allows computation with, for example, logic gates construction, memory functions, analogous computation and their combinations. The information itself has to be counted as a fundamental part of physical reality[1], and any operations with information can lead to cognitive abilities appearing on the appropriately constructed substrate. While computers and AI devices are based on binary operations and digital elements, life forms appear as a mixture of analogous computational devices with neuro-symbolic digital elements. In biology, many tissues besides neural are candidates to be counted as computationally competent and able to perform complex operations with data, including perception, storage, learning, and active collection and selection[2].

Any living system must be able to accept, transform, and produce data. This ability progresses from the rudimentary level of molecular automata and cells to the level of creatures with fully developed centralized neural systems. Evolutionary and taxonomically, we cannot deny the emerging nature of higher intelligence. Incremental growth of data processing capabilities led to the appearance of human civilization with its abilities for non-biological forms of information preservation and processing. The obligatory development of a highly intelligent life is debatable[3]. However, intelligent life is a matter of fact if it is a fluke or aberration. While autonomous life, starting from the procaryote level, exists in constant interaction with the outside environment and readaptation of intracellular morpho-physiology, multicellular organisms are subject to necessary internal data interaction on a much bigger scale. The development of highly organized neural systems is caused by the evolutionary development of multicellular organisms of growing complexity capable of sophisticated behaviour. There is still a possibility of arguing about primary non-human intelligence, which raised human intelligence and, consequently, neural network-based AI. The query still holds with much lower Solomonoff– Kolmogorov–Chaitin complexity of neural networks than human complexity due to the repeatable nature of the first ones.[4].

Interaction between organisms with neural systems leads to the existence of socially complex behaviour and, at least in the case of humans, the development of language and civilization, which are complementary to biological evolution and potentially can serve as tools for directed self-evolution.

doi.org/10.32388/QC239X

Suppose intelligence is the capability to process a certain volume of relatively complex information in a limited time. In that case, the higher intelligence will be marked by the speed, volume and complexity of processed data and the ability to apply acquired knowledge and abilities practically. It is possible to imagine data processing inversely related to time and energy consumption. The lower level is supposed to be limited by the Landauer Principle, where **erasing** a single bit of information in a computational process has a minimal thermodynamic cost, dissipation of a small amount of energy as heat:

$$E_{min} = k_B \; T \ln(2)$$

where:

- E_{min} is the minimum energy required to erase a bit,
- k_B is the Boltzmann constant (1.38×10^{-23}J/K),
- T is the temperature of the computational environment in Kelvin.

There are many utilitarian ways to calculate data processing Power Usage Effectiveness (PUE), but there is no way to compare it effectively with biological systems. We meet Big Data processing in AI applications, but, as discussed in previous papers[5][6][7], natural intelligence is autopoietic, while AI is poetic. This means that existing AI is just a tool for human civilization. Until it becomes autonomous and self-serving on a sufficient scale, there is no Darwinian type of danger from AI for humankind. There is also a potential energy barrier for AI superintelligence[8]. This does not preclude the possibility of resisting human inputs by intellectually autonomous AI or producing dangerous mistakes with negative outcomes for humans. Any emerging superintelligence will be potentially able to ignore less intelligent control. This applies to general intelligence and less to specialized, narrow forms. This is even more applicable to the "ectomental", in the words of Jordan Pollack, "mindless intelligence"[9].

The division between autopoietic and allopoietic intelligence does not mean there is no potential evolutionary concurrence between intelligent species utilizing any tools, including AI. If the Darwinian mechanisms are not sufficiently softened by ethical development, this can lead to the race between more intelligent species and individuals for automatically driven domination. We can see how universal intellectual capabilities and socially meditated civilization developments allowed the homo-sapiens-sapiens to become the Earth's dominant species despite the active resistance of multiple biological agents in the biosphere. This can lead us to the assumption that more intelligent autopoietic agents will tend to exercise control over less intelligent forms, regardless of their nature and level of social organization[10].

Superintelligence can be achieved on the machine level by applying existing AI or creating quantum computers. If, in the case of universal quantum computers, we consider it to be “quantum supremacy”, an expansion of the super-intelligence abilities into any species, agents or societies will create an “intelligence supremacy”. Superintelligent agents might be attained by different countries, groups or possibly individuals in various ways[11]. The simplest one is the acquisition of tools with superb data processing capabilities. The more sophisticated way is biological or transhuman engineering to create more intelligent species, animals or humans. A hypothetical extraterrestrial civilization with interstellar travel techniques potentially has intellectual and instrumental faculties superior to ours. This will lead to intellectually unequal relationships with unpredictable results. There is also the possibility of an “AI filter”, which can explain the Fermi paradox[12]. It might be argued that any highly developed intelligence is inherently moral, and higher levels of development translate into a more benevolent stance. We still lack experimental support for this, and all relevant observations are applied to the history of evolution on Earth and human society's development. It might also be argued that AGI is inherently untestable[13].

There is a strong necessity to produce model framework for the intelligences` interactions in order to predict outcomes in the case of “unbalanced” conditions or more than two actors. It is too early to say we have it.

2. Superintelligence

Primary intelligence is hard to identify, but a generally shared understanding of intelligence gradation exists. The most basic, rudimentary forms will represent the starting point of the scale. Super-intelligence resides on the opposite far end of the intelligence spectrum, evidently higher than the highest human natural intelligence. It has become increasingly common to emphasize the necessity of reducing human biases in the assessment of intellect to move away from an anthropocentric perception. On the other hand, there is a necessity to make AI more explainable for humans, the so-called “XAI”[14]. However, it remains challenging to evaluate intelligence without clear benchmarks. Human intellect, being accessible and relatively measurable, serves as a practical reference point until a more universal scale is established. The concept of super-intelligence is well-known in theory, but it has always been comprehended as a human superb skill or magic non-human capability. With the rise of natural sciences and widespread education, the distinction between average normal intellect and exceptional intellect evolved, favouring better schooling with an element of inter-generational transmission of higher basic intellectual abilities as a cause for it. However, there is also a place for savants, people with specific intellectual superb skills in some limited application areas. It is most prominent in savant syndrome, an intellectual impairment with contrasting outstanding mono abilities[15]. Some questionnaire studies on Mensa members failed to show a specific connection between intellectual power and most of the illnesses and disabilities except autism spectrum disorder and myopia[16]. What potential does human society have to achieve a high level of intelligence? Based on the current state of society, superintelligence looks possible through the development and wide implementation of some medical fields (see Figure 1).

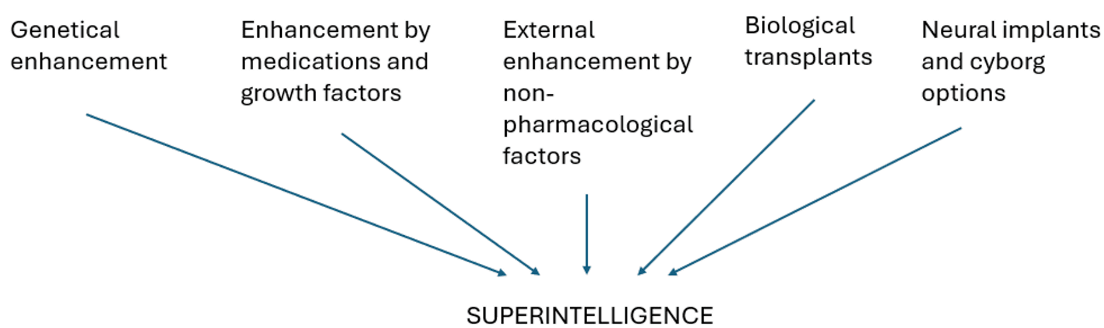

Figure 1. Superintelligence and possible factors leading to its development

2.1. Genetic enhancement

With the ascendance of genetic engineering, especially based on the CRISP/Cas9/gRNA technique (Clustered Regularly Interspaced Short Palindromic Repeats, Cas9 DNA endonuclease, and guide-RNA), there are concerns about the possibility of creating "designed babies" outside of essential medical needs[17]. Military recruitment possibilities and future genome design are focused on a list of autosome genes, such as EPOR for erythropoietin production and, consequently, oxygen transport, the MSTN gene producing myostatin and responsible for the development of muscular tissues, and a number of others. Higher oxygen transport can be an indirect way to improve cognitive functions[18].

 doi.org/10.32388/QC239X

There is also a list of genes recognized as crucial in the development of brain-related clinical conditions and could be potentially targeted for enhancing cognitive functions. Catechol-O-methyltransferase (COMT) and dopamine receptor expression genes for D1, D2, D3, D4, and D5 are all related to various cognitive and behavioural functions besides the development of endogenous psychotic disease[19].

The Brain-Derived Neurotrophic Factor (BDNF) gene, the Serotonin Transporter gene (SLC6A4), Neurexin 1 (NRXN1), the methyl-CpG-binding protein 2 gene (MECP2), the FOXP2 gene for Forkhead Box Protein P2, and some other genes are closely associated with various elements of brain development and functionality[20].

2.2. Enhancement by medications and growth factors

There is also an option to influence brain development and functionality by various growth factors and inhibiting factors, which are used today for exploring in vitro brain organoids[21]. A more practical option stems from cognition-enhancing drugs, which are used to alleviate dementia and can potentially be applied to people with normal intelligence[22]. While popular non-specific modulators, such as caffeine and nicotine, can induce short-term concentration improvement, more specific nootropic stimulators could be responsible for long-term effects. There are several points for medical applications. In addition to stimulating axonal growth and inducing synaptic receptor numbers, several biochemical pathways can target specific memory-related activity or better signal transduction[23].

Pre-synaptic stimulators and receptor stimulators are usual nootropic activators. Attention-related medications and creativity-enhancing substances are also candidates for mental capabilities improvement. However, there is some scepticism concerning psychopharmacological neuroenhancement possibilities[24]. There is also a necessity for additional studies of popular food supplements with supposed nootropic functions. L-tyrosine, catecholamine precursors, such as dopamine, are often considered to be cognitive and attention enhancers[25]. However, it is possibly effective only in distress situations and clinically low levels of L-tyrosine metabolites and only for the short term[26]. The option of long-term pharmacological brain function enhancement remains elusive, and historical studies do not demonstrate the social or intellectual superiority of societies with a tradition of simulators or another drug intake. There is still a question about the balance of potential benefits or damage done by these substances[27].

2.3. Non-pharmacological external enhancement

Other non-pharmacological options for cognitive induction and modulation include Transcranial Magnetic Stimulation (TMS), which is supposed to be an effective non-invasive instrumental technique for alleviating post-stroke cognitive impairment[28]. It can also be used for healthy individuals[29]. Less specific techniques, such as improved non-nootropic nutrition, appropriate physical exercises, sufficient sleep, meditation, mnemonic strategies or computer training, can also be quite productive for improved cognitive abilities in the normal frame. Still, probably they cannot produce super-intelligence as a phenomenon[30].

2.4. Biological transplants

Direct neurogenerative stimulators can be applied to healthy or damaged brain tissues. An obvious additional option is to use transplants or cell injections to enhance brain potential. There are many obstacles to successful auto-, allo- or xenotransplants. Genetically modified neural tissues can be already fully or partially specialized. In contrast, pluripotent or neural stem cells open the possibility for designed differentiation into monoculture or brain cell complexes of neurons, glial cells and ependymal cells[31]. Successful transplants may improve specific function in patients with neurodegenerative disorders or neural lesions[32].

Brain organoid transplantation is a step forward in introducing whole cellular complexes, which have the potential to repair damaged function via direct replacement with developing tissue complexes[33]. Brain cellular or organoid transplants can provide plausible possibilities for normal brain empowerment, and clinical research creates a strong basis for these possibilities. The question is not whether it will be done but how quickly it will be implemented. Undoubtedly, we are talking here about decades, if not years, and indeed not centuries.

2.5. Neuro-implants and cyborg options

Artificial neural implants, currently proposed for the recovery of lost neural function, can potentially extend functionality into cognitive abilities enhancement[34]. There are real prospects for prosthetic hippocampus, which is mainly responsible for memory[35]. While low-level cyborg and biohybrid robots are successfully tested[36], there is still a significant distance between these systems and higher-level cyborg hybrids. There are several neural interface development programs in different stages of completion[37].

Certain concerns exist about future prosthetic chips and devices with an AI interface and relatively easy external control[38]. The real possibility of "uploading" memories, behavioural programs, or elements of character should be seen as a potential future when prosthetics with an external computer interface will achieve sufficient capability to intervene in the functionality of brain structures. Brain cyborg development programs, for instance, can include the childhood stage[39], which may lead to mixed human-computer education, which is different from any form accepted today. It is hard to predict the results, but certain levels of worry could be applied not only to the social and moral aspects but also to simple mental enhancement.

3. AI superintelligence

The attention of researchers and the general public is usually attracted by the ability of AI to achieve superintelligence[40]. Nick Bostrom sees AI's possibility to surpass human cognitive abilities and moral judgment capabilities. He addressed potential ethical questions about super intelligence, which, in his words, will be capable of eliminating ageing and disease, creating von Neumann self-producing space probes to imitate reality realistically, but, at the same time, allow fine-grained control of human mood, emotion, and motivation and advanced weaponry[41].

There is still a question of whether we should consider AI to be not only super intelligent but also conscious[42]. Calculators can perform mathematical calculations quicker than most humans, but does this instrumental ability differ from transport capability for quick movement or heavy equipment for heavy lifting? There is a strong opinion that AI superintelligence most probably will create problems for humankind, not solutions[43]. Others are less pessimistic and expect solutions to multiple world problems in the upcoming decades. At the same time, the development of autonomous AU superintelligence will inevitably be self-limited by the necessity to create an informational super channel with the outside world[44]. There are other ways to address the question of superintelligence. It can be recognised not as an individual but as a collective one, General Collective Intelligence[45]. This systemic approach can be helpful in terms of the development of AI superintelligence. Others are less optimistic. Alan Turing says, "If a machine can think, it might think more intelligently than we do, and then where should we be? Even if we could keep the machines in a subservient position, for instance, by turning off the power at strategic moments, we should, as a species, feel greatly humbled. This new danger is certainly something which can give us anxiety." AI can amplify human limitations and cause a negative Flynn effect[46]. Even with dependant, controlled super-intelligent AI, we can expect the rise of autocracy and negative social impact[47].

The question of superintelligence, artificial or natural, remains hotly debated, and there is no clear solution in the case of any superintelligent force appearing alongside humankind. The question of superintelligence morality or "benevolence" is approached from different directions, and answers range from the inevitable rise of cooperation and morality as a positive correlate of intelligence to less positive predictions[48]. Some scholars even turn to the God metaphor in describing possible super intelligence "revelation": it is not easy to predict super-intelligent creature or object behaviour as we cannot judge it adequately, especially towards other intelligent or less intelligent specimens[49]. It is challenging to build a strong reasoning line about any moral obligation of super-intelligence, especially AI, as it will not necessarily recognize life as inherent value and with a weak connection to common evolution.

4. Alien intelligence

An intelligence known to us is based on life forms or, as an AI, produced by biological intelligent forms. As a prerequisite for the existence of any extraterrestrial intelligence, we have to assume the existence of a habitat suitable for extraterrestrial forms and their ability to achieve a high level of cognition. While signs of intelligent life are supposed to be technological, xenobiological signatures of life are biochemical, and prerequisites for life are astrophysical[50].

The famous Drake equation is the most notable attempt to formalize the estimation of N, the number of civilisations available for communication with humankind:

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

were $R*$ is the mean rate of star formation or N_s, the number of stars in our galaxy (for galactic estimation); f_p is a fraction of stars with planets, n_e is the number of Earth-type planets; f_l is a fraction of planets with life; f_i is fraction of planets where intelligent life developed; f_c is fraction of intelligent life forms developing civilisation; L or f_L is fraction of civilisations existing long enough to establish contact[51].

doi.org/10.32388/QC239X

Life as we know it is carbon-based, mainly cellular, water-dependent, and mostly phototrophic, with a small fraction being chemotrophic[52]. There are possibilities of xenobiology, alternative biochemistry, and energy harvesting. Estimations of worlds with developed biospheres range from 500,000 to 100 million. Nevertheless, it is still necessary to explain the Fermi paradox in one way or another[53]. Specific conditions are essential to combine biochemical pathways into precellular and cellular life, and they do not necessarily exist at the same time and place simultaneously. The Earth's conditions changed radically during the 4 billion years. At the same time, there are multiple carbon-based biochemical substances in space, from the carbon itself to more complex proteinogenic and non-proteinogenic amino acids, which can give an abiogenetic start for the life forms.[54]

There are theoretical possibilities of xenobiological forms, including silicon-based ones. However, strong bonds between silicon and oxygen create significant restrictions, creating hard matter. Life can be recognized in most general terms as a "self-sustainable organization in a liquid state"[55]. Life states and their flow is reflected in Figure 2.

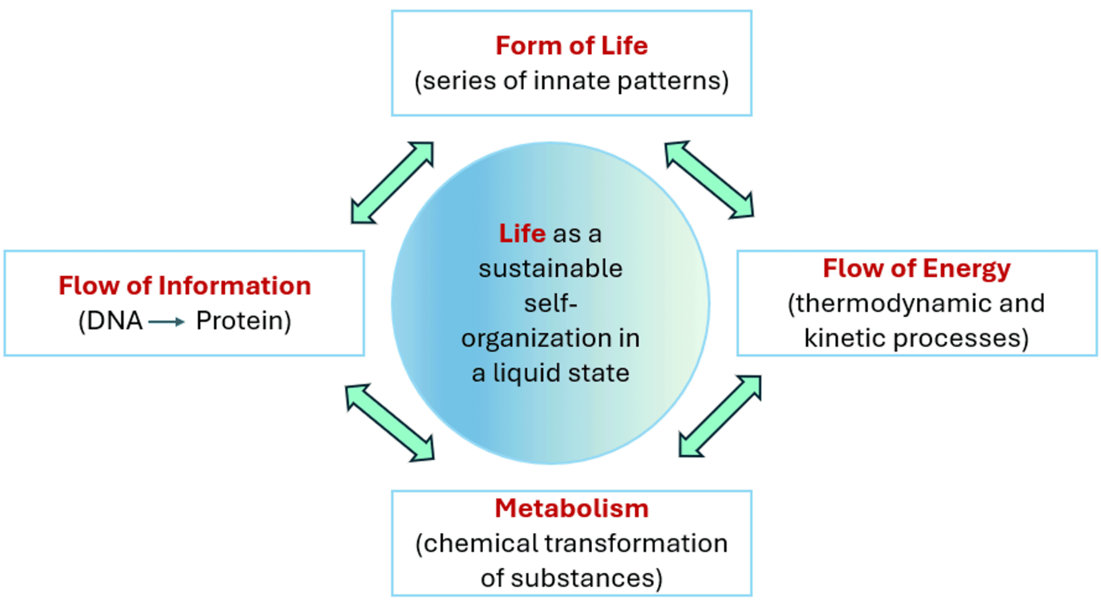

Figure 2. Circle of Life components

It can be abstracted in most general terms as self-perpetuating chemical transformation of substances, which requires energy harvesting, related information flow and necessary innate patterns[56].

doi.org/10.32388/QC239X

Gerald Joyce famously stated that life is a self-sustained chemical system capable of undergoing Darwinian evolution[57]. Intellectual development is inherently emergent in a sufficiently complex planetary ecosystem, which usually requires a water-reach rocky planet[58] in a stellar climatological habitable zone (CLI-HZ), ideally near the G or K star. K stars of brighter type are more long-lived than G stars of solar-type and can provide the best conditions for life development long enough to harbour highly intelligent forms[59].

There is a preference for simple biological systems over complex systems. Microbials were the only form for 70% of Earth's history. Up to 10 million species of eukaryotes and from 100 million to at least a billion prokaryote species exist today on the Earth. Only humans have reached the technological level of intelligence[60], which can reflect the Darwinian effect of dead ends for alternative comparable intelligent forms and, at the same time, the co-evolutionary existence of humankind in the biosphere. Based on evolutionary history and reasoning, there is a strong opinion that intelligence and civilization development are niche phenomena without high evolutionary necessity. It can explain the Fermi paradox: possibilities for life emergence are quite abundant, while intelligent technological civilization is the rare non-convergent flux of common biological evolutionary processes[61].

Marvin Minsky's argument is the opposite and supports the idea of convergent intelligence regardless of its nature. He builds his hypothesis upon implications of cosmological principle: laws of Universe are universal and lead to similar physical possibilities and restrictions[62].

Intellect is a necessary development to mitigate these constraints. Problem-solving intellectual capacity is beneficial for the survival of intelligent species but can also characterize AI without the necessity of evolutionary survival and adaptation, regardless of the symbolic or sensorimotor physical nature of the problem. It also applies to alien intelligent civilizations, so it should not be a problem to communicate or mutually understand other early stages of space-exploratory civilizations. The answer to the Fermi paradox can be different: advanced civilizations do not have a strong link between intelligence and "survival, communication, and expansion of control over the physical world", so they do not seek contact or endless expansion.

Stephen Hawking sees potential contact as undesirable: a more technologically advanced civilization is deadly dangerous for a less advanced party. The counterargument is based on the Fermi paradox explanation through similar advancement reasons. Interstellar voyages "would probably be justified only for major purposes, and plundering the Earth for its resources would be neither practical nor desirable"[63].

Carl Sagan based his argument for civilization survival on the ability to cooperate, and non-cooperative civilization will be stopped by more advanced adversaries for exactly this reason. Still, there is a question about the contact: if we did not detect other civilizations, they could detect us by some means. One of the answers to the Fermi paradox is the less-developed nature of human intelligence and civilization. We are not recognized as intelligent enough to be contacted for intelligent communication, even though human or other terrestrial intelligence can be comprehended as a subset of higher-level intelligence with interstellar transport capabilities[64].

Intelligence is in its most developed form; as we know, it is human. And humankind exploited less intelligent biological forms for utilitarian purposes[65]. On the other hand, humankind has developed ethical responsibility towards life beings, especially those who are more intelligent. It can be in the case of contact with less-developed extraterrestrial intelligence. But what could be done about contact with more developed civilizations? There is a hypothesis that it is purely technological development. coupled with intelligence, are not sufficient for survival even in the Darwinian sense and developed ethics are required. Can this logic be applied to interaction with AI as well?

5. Intelligence vs SuperIntelligence

It is usually thought that there is no intellect without recognition by other comparable intellects and without the possibility of interactive communication with the same or similar types during development. It has inherent potential, as no intellect can be seen as a completely separate autonomous entity. But should the "other mind reading" include sentience as a necessary element[66]?

There is still a question of whether AI is a form of natural intelligence emulation, sensorimotor, pattern recognition or task-related. We cannot claim natural autopoietic functions in any artificial system. The question is whether we could successfully emulate it tomorrow or if natural intelligence is only the form of the first intelligent system. In contrast, any secondary intelligence, biological or non-biological, is inherently only an extension of autopoietic tasks, subdivided into enhanced intelligent tools. Should it be reflected in the anthropic intelligence scale or more generalized natural intelligence scale or judged separately as unique AI or possibly incomprehensible non-biological intelligence[67]?

We have to recognize that any intelligence incomprehensible to us, inaccessible as such, is outside of our assessment or categorization. It inevitably returns us to the question of the common intelligence scale, based on the human-comprehension basis only, without any other intelligent features inaccessible to us.

Mutually comprehensible intelligence requires language of interaction beyond simple recognition of "other minds". In the biological world, different species with comparable intelligence interact with a certain level of understanding. The same is applied to humans in interactions with other humans on different levels of civilizational development and with cognitively developed animals. AI is developed by humans and, by default, is human-accessible. There is a possibility of envisioning a relatively universal intelligence metric applicable to any intelligent entity[68].

There are several requirements for a universal intelligence test. It has to be based on a universal, well-formalized computational principle, possess an adequate measurement scale, encompass any possible intelligence by scale, from minimal to superintelligence, and have scope, any type of it. The test should be environment- or context-independent and also time-independent or capable of registering it. The Turing test and its analogues are supposed to be inadequate for the task, as are classical psychometric tests[69].

The panel of experts provided a generalized informal definition of intelligence: "Intelligence measures an agent's ability to achieve goals in a wide range of environments". Cycles of the interaction of the intelligent agents with the environment are goal-oriented, and any intellectual task can be formalized in this framework. The conditional probability measurement provides a scale of improvement through iterative attempts and can be formalized through the algorithmic prior distribution.

While the universal intelligence test could be achievable, another problem arises. As we mentioned previously, intellectual capabilities of any sort could not be seen as separable or autonomous from other intelligent agents or environments. The goal in any test is achieved in the real environment or a simulated or abstract one. For example, in the case of human intelligence measurement, we consider the natural and social environment, the ability to interact with the help of natural language and the capability to learn and memorize. There is a necessity to extend formal testing to the network of interactions, where goals can be expanded to the "intelligent environment interactions"[70]. There is a possibility of creating a framework with scales for different types of intelligence and additional measurements, such as processing speed and complexity. However, there is no one-fit-all scale to compare different types of intelligence in different conditions and no apparent way to construct it. Task-related tests are more practical for data processing, behavioural, or operational estimation.

Smart environments, envisioned by the concept of the Fifths Industrial Revolution, include not only AI-powered smart devices, vehicles or robotic units but whole integrative networks. In this case, the interaction of humans with a smart environment or AI with a smart or human environment is not a simple goal-oriented task of evolutionary natural intelligence of animal-kind but a more developed one in a significantly more complex environment. Human-machine interaction transforms into human-smart, complex interaction with physical, psychological and social dimensions[71].

Certain concerns go far beyond simple universal intelligence measurement. An interaction with a smart environment or intelligent population reserves a human position as an arbiter. However, with the quick development of intelligent networks, super-intelligence can be encountered in artificial form much earlier than in extraterrestrial form[72]. A high level of intelligence cannot be judged only by goal-oriented solutions. Human goals are often complex, and applying the same rationale to the developed levels of intelligence would be reasonable. Complex goals are not expected to be "coded" on the most basic level. Mathematical operators or formulations in their abstraction cannot fully reflect the necessary complexity. Insufficient interaction between human and computer systems sometimes leads to massive missing data, even in critical areas[73]. Even the understanding of human inputs could be too literal, which is sufficient for the undesirable results from AI. The necessity for AI to learn complex goals can be difficult at best and impossible at worst. The consistency of all sub-goals with target final goals can be hard to achieve. These "final goals" have to be considered not as final in literal terms but as an interim complex goal in the teleological chain of future complex goals, which requires endless recursion to the "primary goal" for correction.

Another element of concern is speed, with which AI systems acquire intelligent abilities, including creative ones, which are impressive. It is temporarily superior to humans when we require years of learning, while AI systems can be trained much quicker. The idea that creativity is a measure of natural intelligence and only mundane tasks are left to AI is definitely not holding on. LLM successfully takes the Torrance test on creativity, and the creative abilities of generative AI are demonstrated in literature, music and visual arts[74].

Ethical questions related to AI development have to include the necessity of the development of an ethical emulator or proxy for AI-powered robotics if we want to prevent "antisocial" incidents.

doi.org/10.32388/QC239X

Empathically constrained behaviour[75] is insufficient if we consider empathy versus sympathy, and gradation of the behaviour is important. We can imagine interaction with any intelligent agent possessing comparable to humans or a higher level of intelligence. Precedents from the experience in interactions between humans or between humans and natural or artificial intelligence are not providing highly optimistic examples, even though there is not as dramatic as sometimes seen. Interaction with AI has to be, according to Stuart Russell[76], subjected to three main principles:

1. The machine's only objective is to maximize the realization of human preferences.
2. The machine is initially uncertain about what those preferences are.
3. The ultimate source of information about human preferences is human behaviour.

There are potential dangers of human-incompatible AI. While these problems are generally understood, there is no clear way to deal with them. The simple suggestion is the ability to switch a problematic AI off if it would be possible to do so with the superintelligent machine. Currently, the problem is not seen as actual, and there is no need to control it. It might be a premature issue, as superintelligent AI is expected roughly in a century. The vision exists that intelligence inevitably will be connected with altruistic tendencies, which is far from obvious[77].

There is also an argument about intelligence multidimensionality, discussed above. AI will be narrowly specialized, and there is no danger from the intelligent tool as long as it is restricted to a certain area. However, we cannot be sure that narrow superintelligence will be fully controlled. The same logic applies to extraterrestrial intelligence but without the capability to control it in the same way as AI. We can see that the framework for interacting with other intelligence must be researched further to be more prepared for contemporary and future challenges.

Here is also the place to ask about natural limits to superintelligence. It might be a computational limit[78] or a natural physical limit, as in the case of the Kardashev scale of civilization energy consumption[79]. Extending the Kardashev scale brings it to the Universe's energy level limit.

6. Intelligence, technology development and social adaptation

The intelligence of a certain society can be extrapolated to the technological level, which requires certain metrics. There are many potential elements, scales, and metrics for the civilization development assessment, from infrastructure to self-organization[80]. The paraphrase of Aristotelian "zoon politikon" is reflected in the idea that humans are "technological animals"[79].

Universal adaptation to most environments on the Earth is possible due to tool-making and technological adaptation rather than biological. While even bacteria are able to change the environment and lower species can use non-biological substances for self-development and other biological species for survival, no other creature builds technological civilization on the human scale. It is possible only due to intellectual abilities and certain biological anatomy. e.g. operative hands-free to be used with tools, binocular chromatic vision and social organization. While biological evolution took billions of years for the development of life forms with intellect and several million years for the shift from biological to techno-social adaptation, the technological phase is quite short. It accounts for around 2.5 million years and 1.5 years for stone tools usage and fire control, respectively. The civilization itself, from the Neolithic revolution through the utilization of metal to contemporary computers, is just a few thousand years old[79].

Writing system development 5000 years ago helped the transition from the prehistoric stage to the historic one. Still, the tool was based on readily available materials without deep transformation, and agricultural techniques were very basic. Energy was consumed from very natural sources and mainly did not include contemporary forms of energy harvesting, from fossil fuels to direct solar energy. While it might be possible in future to produce food from low organic or inorganic substances, with the possibility of Neolithic counter-revolution, and have the ability to operate on the micro (nano) level and macro (planetary) level, instrumental information processing and energy consumption are most reflective metrics of civilization development on a big scale.

Kardashev macro scale (repeated) of energy consumption by civilization includes four types. Type 0, introduced later, is biological, with a consumption of 10^6 W. Type I means terrestrial energy consumption in 1964, the year of scale publication. Type II is "a civilization capable of harnessing the energy radiated by its own star". Type III is "a civilization in possession of energy on the scale of its own galaxy". Types I, II and III can use 10^16 W, 10^26 W and 10^36 W respectively[81]. Type IV will harvest 10^46 W to the Universe level. The extended Kardashev scale of energy consumption is reflected in Figure 3.

doi.org/10.32388/QC239X

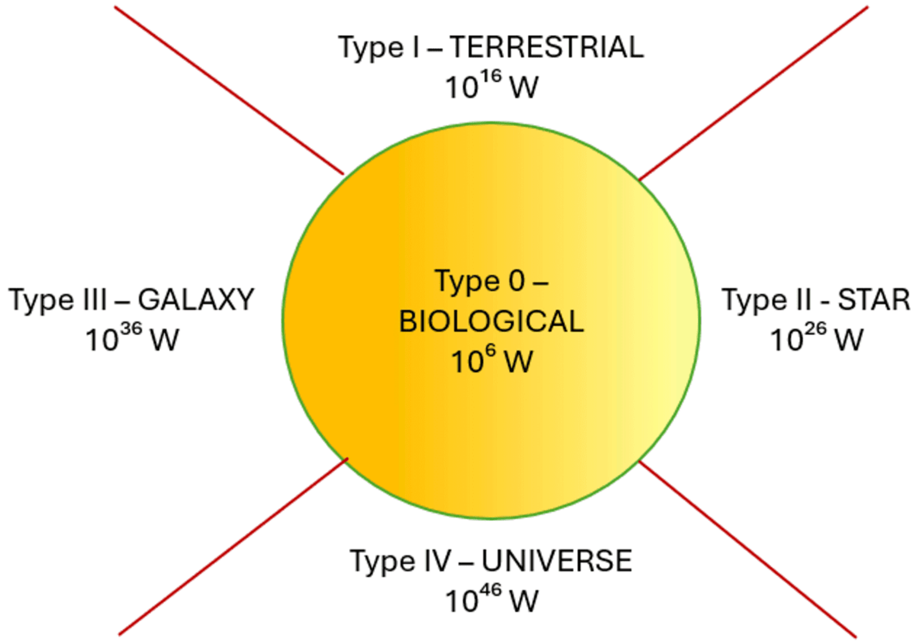

Figure 3. Four types of civilization's energy consumption

Here, we discuss civilization's technological level and not the intellectual capabilities of its individuals or machines. More fine-grained scales of data processing and energy consumption can be helpful in evaluating the possibilities open for the individual inside civilization. But even from the Kardashev scale, we can see the God-like or demiurge-like abilities of higher civilization types compared with lower ones. We can assume that higher energy consumption levels can be extrapolated into higher data processing capabilities. If it is an extraterrestrial civilization, we cannot do much about it. But if superintelligence capabilities are born inside our civilization, we can assume some potential levers to influence it.

One of the questions is what was first, technology or a certain level of social development that led to this technology, which reminds us of the argument of chicken or egg priority. While biologically and evolutionary, the egg came first. The main idea of the argument, if we start "*ab ovo*", might be about binary processes, where change in technology and change in society go hand in hand[82]. The Kantian's answer to the question of priority might be a reformulation of the relationship between a priori and a posteriori knowledge: ''theory without practice is empty; practice without theory is blind.'' Society can come to certain technologies through the stepwise process of being ready to accept

 doi.org/10.32388/QC239X

new technology or building preparedness on the basis of preliminary technology. Regardless of what it is, the first real question is about the possibility of controlling or stopping completely undesired technology inside the wide society[83].

Technology develops with material resources and applicable knowledge. Part of the resources are social-related resources. The material can be limited, but the second element, knowledge, and the third element, other socially related resources, cannot be easily controlled. Putting limits to the proven unwelcome development is often expensive, technically challenging, and politically complex if we take into account formal political or any other independence of governments and societies. It might be questionable for AI development control and really impossible to control on a global scale if it is not done globally by the only existing superpower with a real possibility of suppressing unwanted prospects inside but without any serious control of the Universe outside.

While we cannot speculate about the future control of technological development outside of the Solar System, even Earth itself, we can ask ourselves: What should be done in the case of failed control inside our civilization? The simplest solution is to be ahead, meet challenges, and be prepared to fence it up with our own development.

7. Conclusion

Superintelligence is a relative phenomenon and cannot be arbitrarily claimed today only based on a certain absolute level of information processing capacity or operational capability. Superintelligence is a strong intellectual over-performance of average human cognitive abilities and the best possible human mental capabilities. At the same time, we are free to use the scale of parameters from primitive cognitive levels to modern human levels and higher degrees of intelligence. It is necessary to recognize that we have some points of difficulty while trying to address the problem of intelligence scale construction. The first one is the inability to clearly and singularly claim the fundamental minimal level of intelligence. The categorical vagueness of the borderland between pre-intelligence and intelligence proper exacerbates the problem. It might be shown as a continuous space with several scales, such as data processing speed, task complexity levels, and operational or behavioural scales, while still being related to and compared with human abilities as a comprehensible benchmark. The starting point is zero intelligence, which still has to be shown, and the final point for the scale is genuinely Universal, limited only by the physical abilities of the intelligent system.

There is an option to take a step back in the discussion on Artificial Consciousness and take on the problem of narrow Artificial Intelligence, even without considering AGI. Intelligence, at least some types of it, is practically assessable: measurable on the constructed scale in accordance with the level of solved problems. It allows us to create a framework for the potential practical evaluation of any intelligence, regardless of its physical basis, working mechanism nature, or architecture. While recognizing the possibility of a comparative universal intelligence measurement environment, we discuss the necessity of using the framework not only for academic and philosophical reasons but purely for the necessity of pragmatic examination of risks attached to xeno-intelligence. The issue of autopoietic intelligent systems is less important in the face of any signs of autonomy of intelligent systems. We argue that regardless of the nature of the system, biological, non-biological, terrestrial or extraterrestrial, dominating autonomous cognitive system has a potential to ignore the "less intelligent counterpart" for explainable self-reason, nor readily available for the less fortunate. The contingency plans for the interaction between humans and non-human intelligence must include risks of adverse reactions from the other side. Moreover, we can apply this principle to human-human intelligence interactions. It remains to be investigated whether universal moral restrictions, as we know them, are applicable to all intelligent systems or societies.

References

1. ^*Danchin A, Fenton AA. "From Analog to Digital Computing: Is Homo sapiens' Brain on Its Way to Become a Turing Machine?" Frontiers in Ecology and Evolution. 10 (2022): 796413.*
2. ^*Rouleau N, Levin M. "The multiple realizability of sentience in living systems and beyond." eneuro. 10.11 (2023).*
3. ^*Rospars JP. "Trends in the evolution of life, brains and intelligence." International Journal of Astrobiology. 12.3 (2013): 186-207.*
4. ^*Bilokon PA. "Evolution-Bootstrapped Simulation: Artificial or Human Intelligence: Which Came First?" arXiv preprint arXiv:2402.00030 (2024).*
5. ^*Herzog DJ, Herzog N. (2024). "On the Hard Problem: Revisited, Re-Evaluated, Recast." Qeios. doi:10.32388/L7SSUA.*
6. ^*Herzog DJ, Herzog N. (2024). "Emergence ex Machina: Correlates of Consciousness." Qeios. doi:10.32388/ROLBJW.*

7. ^Herzog DJ, Herzog N. (2024). "Intelligence: The Quest for a Universal Assessment Framework." Qeios. doi:10.32388/TGPFZF.
8. ^Stiefel KM, Coggan JS. "The energy challenges of artificial superintelligence." Frontiers in Artificial Intelligence. 6 (2023): 1240653.
9. ^Pollack JB. "Mindless intelligence." IEEE Intelligent Systems. 21.3 (2006): 50-56.
10. ^Bales A, D'Alessandro W, Kirk-Giannini CD. "Artificial intelligence: Arguments for catastrophic risk." Philosophy Compass. 19.2 (2024): e12964.
11. ^Rosenberg L, et al. "Towards Collective Superintelligence: Amplifying Group IQ using Conversational Swarms." arXiv preprint arXiv:2401.15109 (2024).
12. ^Garrett MA. "Is artificial intelligence the great filter that makes advanced technical civilisations rare in the universe?" Acta Astronautica. 219 (2024): 731-735.
13. ^Yampolskiy RV. "Unpredictability of AI: On the impossibility of accurately predicting all actions of a smarter agent." Journal of Artificial Intelligence and Consciousness. 7.01 (2020): 109-118.
14. ^Hemmer P, et al. "Human-AI Complementarity in Hybrid Intelligence Systems: A Structured Literature Review." PACIS (2021): 78.
15. ^Heaton P, Wallace GL. "Annotation: The savant syndrome." Journal of child psychology and psychiatry. 45.5 (2004): 899-911.
16. ^Sofaer JA, Emery AE. "Genes for super-intelligence?" Journal of Medical Genetics. 18.6 (1981): 410-413.
17. ^Gouni O. "Prenatal Human Modification, Designing Babies and the End of Homo Sapiens." International Journal of Prenatal and Life Sciences DOI: 10.24946/IJPLS (2019): 14.
18. ^Ranisch R. "When CRISPR meets fantasy: Transhumanism and the military in the age of gene editing." Transhumanism: The Proper Guide to a Posthuman Condition or a Dangerous Idea? (2021): 111-120.
19. ^Savitz J, Solms M, Ramesar R. "The molecular genetics of cognition: dopamine, COMT and BDNF." Genes, brain and behavior. 5.4 (2006): 311-328.
20. ^Ramus F. "Genes, brain, and cognition: A roadmap for the cognitive scientist." Cognition. 101.2 (2006): 247-269.
21. ^Chiaradia I, Lancaster MA. "Brain organoids for the study of human neurobiology at the interface of in vitro and in vivo." Nature Neuroscience. 23.12 (2020): 1496-1508.
22. ^Mehlman MJ. "Cognition-enhancing drugs." The Ethics of Sports Technologies and Human Enhancement. Routledge, 2020. 275-298.

23. ^Lanni C, et al. "Cognition enhancers between treating and doping the mind." Pharmacological Research. 57.3 (2008): 196-213.

24. ^Schleim S, Quednow BB. "How realistic are the scientific assumptions of the neuroenhancement debate? Assessing the pharmacological optimism and neuroenhancement prevalence hypotheses." Frontiers in Pharmacology. 9 (2018): 3.

25. ^Vine CAJ, Spurr TJ, Blacker SD. "Food for thought: dietary nootropics for the optimisation of military operators' cognitive performance." BMJ Mil Health (2024).

26. ^Jongkees BJ, et al. "Effect of tyrosine supplementation on clinical and healthy populations under stress or cognitive demands—A review." Journal of Psychiatric Research. 70 (2015): 50-57.

27. ^Schifano F, et al. "Benefits and harms of 'smart drugs'(nootropics) in healthy individuals." Drugs. 82.6 (2022): 633-647.

28. ^Sandrini M, et al. "Cognitive reserve predicts episodic memory enhancement induced by transcranial direct current stimulation in healthy older adults." Scientific Reports. 14.1 (2024): 4879.

29. ^Dengler J, et al. "Enhancing cognitive control with transcranial magnetic stimulation in subject-specific frontoparietal networks." Cortex. 172 (2024): 141-158.

30. ^Dresler M, et al. "Non-pharmacological cognitive enhancement." Neuropharmacology. 64 (2013): 529-543.

31. ^Henriques D, et al. "Successes and hurdles in stem cells application and production for brain transplantation." Frontiers in Neuroscience. 13 (2019): 1194.

32. ^Nakagomi T, et al. "Transplantation of human brain-derived ischemia-induced multipotent stem cells ameliorates neurological dysfunction in mice after stroke." Stem Cells Translational Medicine. 12.6 (2023): 400-414.

33. ^Cao S-Y, et al. "Cerebral organoids transplantation repairs infarcted cortex and restores impaired function after stroke." NPJ Regenerative Medicine. 8.1 (2023): 27.

34. ^Sun C, et al. "Perspectives on tissue-like bioelectronics for neural modulation." Iscience. 26.5 (2023).

35. ^Berger TW, et al. "A hippocampal cognitive prosthesis: multi-input, multi-output nonlinear modeling and VLSI implementation." IEEE Transactions on Neural Systems and Rehabilitation Engineering. 20.2 (2012): 198-211.

36. ^Siljak H, Nardelli PHJ, Moioli RC. "Cyborg insects: Bug or a feature?" IEEE Access. 10 (2022): 49398-49411.

37. ^Barfield W, Williams A. "Cyborgs and enhancement technology." Philosophies. 2.1 (2017): 4.

38. ^Schneider S, Corabi J. "Cyborg divas and hybrid minds." The Mind-Technology Problem: Investigating Minds, Selves and 21st Century Artefacts (2021): 145-159.
39. ^Laughlin CD. "The cyborg, the ideology chip and the guru programme: the implications of cyborg tech nologies for the development of human consciousness." foresight. 2.3 (2000): 291-312.
40. ^Gill KS. "Artificial super intelligence: beyond rhetoric." Ai & Society. 31 (2016): 137-143.
41. ^Bostrom N. "Ethical issues in advanced artificial intelligence." Machine Ethics and Robot Ethics (202 0): 69-75.
42. ^Hauser L. "Why isn't my pocket calculator a thinking thing?." Minds and Machines: Journal for Artifici al Intelligence, Philosophy and Cognitive Science (1993).
43. ^Torrance S. "Super-intelligence and (super-) consciousness." International Journal of Machine Consci ousness. 4.02 (2012): 483-501.
44. ^Heylighen F. "Return to Eden? Promises and perils on the road to a global superintelligence." The end of the beginning: Life, society and economy on the brink of the singularity (2015): 243-306.
45. ^Williams AE. "General Collective Intelligence and the Transition to Collective Super-Intelligence." (202 0).
46. ^Wilks Y. "Will there be superintelligence and would it hate us?" AI Magazine. 38.4 (2017): 65-70.
47. ^Fortes P. "Hasta la vista, baby: reflections on the risks of algocracy, killer robots, and artificial superint elligence." Revista de la Facultad de Derecho de México. 71.279 (2021): 45-72.
48. ^Fox J, Shulman C. "Superintelligence does not imply benevolence." ECAP. 10 (2010): 456-462.
49. ^Szocik K, Tkacz B, Gulczyński P. "The revelation of superintelligence." AI & SOCIETY. 35 (2020): 755-7 58.
50. ^Prantzos N. "A joint analysis of the Drake equation and the Fermi paradox." International Journal of A strobiology. 12.3 (2013): 246-253.
51. ^Maccone C. "The statistical Drake equation." Mathematical SETI: Statistics, Signal Processing, Space M issions. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. 3-72.
52. ^Irwin LN, Schulze-Makuch D. "The astrobiology of alien worlds: Known and unknown forms of life." U niverse. 6.9 (2020): 130.
53. ^Szostak JW. "The origin of life on Earth and the design of alternative life forms." Molecular Frontiers Jo urnal. 1.02 (2017): 121-131.
54. ^Cockell CS. "The laws of life." Physics Today. 70.3 (2017): 42-48.

55. ^Budisa N, Kubyshkin V, Schmidt M. "Xenobiology: A Journey towards parallel life forms." ChemBioChem. 21.16 (2020): 2228-2231.
56. ^Gánti T. The principles of life. Oxford University Press, 2003.
57. ^Joyce GF. "Evolution in an RNA world." Cold Spring Harbor symposia on quantitative biology. Vol. 74. Cold Spring Harbor Laboratory Press, 2009.
58. ^Cockell CS. "Habitable worlds with no signs of life." Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. 372.2014 (2014): 20130082.
59. ^Cuntz M, Guinan EF. "About exobiology: the case for dwarf K stars." The Astrophysical Journal. 827.1 (2016): 79.
60. ^Cabrol NA. "Alien mindscapes—a perspective on the Search for Extraterrestrial Intelligence." Astrobiology. 16.9 (2016): 661-676.
61. ^Lineweaver CH. "Paleontological tests: human-like intelligence is not a convergent feature of evolution." From Fossils to Astrobiology: Records of Life on Earth and Search for Extraterrestrial Biosignatures. Dordrecht: Springer Netherlands, 2009. 353-368.
62. ^Diederich J. "The convergence of intelligences." Bioastronomy 99. Vol. 213. 2000.
63. ^Michael G. "Extraterrestrial Intelligence: Friends, Foes, or Just Curious?" Extraterrestrial Altruism: Evolution and Ethics in the Cosmos (2014): 25-35.
64. ^Szocik K, Abylkasymova R. "If extraterrestrial intelligence exists, it is unable to recognize humans as intelligent beings." International Journal of Astrobiology 21.6 (2022): 462-468.
65. ^Peters T, Arnould J. "Intelligent aliens and astroethics." Space Exploration and ET: Who Goes There (2014): 1-20.
66. ^Nyholm S. "Other minds, other intelligences: The problem of attributing agency to machines." Cambridge Quarterly of Healthcare Ethics 28.4 (2019): 592-598.
67. ^AJ, Gomez-Ramirez D, Kieninger J. "A Holistic Test for (Artificial) General Intelligence." (2023).
68. ^Hernández-Orallo J, Dowe DL. "Measuring universal intelligence: Towards an anytime intelligence test." Artificial Intelligence 174.18 (2010): 1508-1539.
69. ^Legg S, Veness J. "An approximation of the universal intelligence measure." Algorithmic Probability and Friends. Bayesian Prediction and Artificial Intelligence: Papers from the Ray Solomonoff 85th Memorial Conference, Melbourne, VIC, Australia, November 30–December 2, 2011. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013.

70. ^*Kranz M, Möller A, Roalter L. "Robots, objects, humans: Towards seamless interaction in intelligent environments." 1st International Conference on Pervasive and Embedded Computing and Communication Systems (PECCS 2011). 2011.*
71. ^*McKenna HP. "Human-smart environment interactions in smart cities: exploring dimensionalities of smartness." Future Internet 12.5 (2020): 79.*
72. ^*Petersen S. "21 SUPERINTELLIGENCE AS SUPERETHICAL." Robot ethics 2.0: from Autonomous cars to artificial intelligence (2017): 258.*
73. ^*Cismondi F, et al. "Missing data in medical databases: Impute, delete or classify?" Artificial intelligence in medicine 58.1 (2013): 63-72.*
74. ^*Guzik EE, Byrge C, Gilde C. "The originality of machines: AI takes the Torrance Test." Journal of Creativity 33.3 (2023): 100065.*
75. ^*Christov-Moore L, et al. "Preventing antisocial robots: A pathway to artificial empathy." Science Robotics 8.80 (2023): eabq3658.*
76. ^*Russell S. Human compatible: AI and the problem of control. Penguin UK, 2019.*
77. ^*Russell S. "Human-Compatible Artificial Intelligence." Oxford University Press (2022): 3-23.*
78. ^*Alfonseca M, et al. "Superintelligence cannot be contained: Lessons from computability theory." Journal of Artificial Intelligence Research 70 (2021): 65-76.*
79. a, b, c*Genta G, Riberi P. Technology and the Growth of Civilization. Springer, 2019.*
80. ^*Targowski A. "The element-based method of civilization study." Comparative Civilizations Review 81.81 (2019): 6.*
81. ^*Gray RH. "The extended Kardashev scale." The Astronomical Journal 159.5 (2020): 228.*
82. ^*Dosi G. "Technological paradigms and technological trajectories: a suggested interpretation of the determinants and directions of technical change." Research Policy 11.3 (1982): 147-162.*
83. ^*Geist EM. "It's already too late to stop the AI arms race—We must manage it instead." Bulletin of the Atomic Scientists 72.5 (2016): 318-321.*

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.