Artificial Intelligence: The Rise of the Machines

Will Artificial Intelligence, continually advancing at an exponential rate, be detrimental to the human race – financially, socially, and physically – within the next century?

Morgan S. Navarro
Columbia University
February 2016
Artificial Intelligence is defined as the theory and development of computer systems able to perform tasks that normally require human intelligence, including visual perception, speech recognition, decision-making, and translation between languages. In addition, it is an academic field of study, which involves the creation of computers, and computer software that is capable of intelligent behavior. This intelligence works to study and design intelligent agents – systems that perceive their environments and take action to maximize the chances of overall success. Artificial intelligence, if it continues to advance rapidly in the near future, has the potential to make lives easier through the understanding of human desires along with more efficient product manufacturing, but if left uncontrolled, the technology could pose a serious security threat to our society as a whole.

It is likely that any technology that we can currently foresee will be rapidly developed by the first superintelligence, a theoretical being which possesses intelligence far greater than that of the most advanced and developed human brains, along with many other technologies of which we are, as of yet, unaware. The foreseeable technologies of artificial superintelligence will involve advanced weaponry and security devices capable of safely disarming nuclear armaments, space travel, and von Neumann probes (self-reproducing interstellar probes), fine-grained control of human mood, emotion, and motivation, and the reanimation of cryonics patients. In addition, uploading (neural or sub-neural scanning of a specific brain and the implementation of identical algorithmic structures as on a computer to persevere memory and personality) and the creation of a fully realistic virtual reality are two feasible products of a highly advanced artificial intelligence. In the article “Will Robots Take over the World?,” three renowned professors from the University of Cambridge discuss their opinions on the future of artificial intelligence development and the possibilities of these developments causing detrimental outcomes for the human race. These researchers have identified two major global uses of artificial intelligence that provide the greatest positive societal impact. The first involves the use of robotic machines in places where humans cannot safely travel or reach. These places include collapsed mines, nuclear reactor sites, oilrigs, and even outer space or other celestial bodies. The second use involves implementing robots for the care of the elderly, such as helping them tie their shoes and cut their toenails. These robots would greatly benefit society and allow for more scientific research to be conducted safely; however, the development of artificial intelligence for these uses presents an existential threat to the safety of our society. In addition to the two mentioned above, a third possible use relies on the creation of robots on a miniaturized scale – called the nanoscale – that would allow these machines to be placed within the human body in order to monitor and report data regarding daily bodily processes. This concept, known as nanotechnology, would provide the greatest benefit for society as it would aid in medical research and allow humans to monitor their health at the molecular level. Deepak Srivastava, a professor of pediatrics at the University of California at San Francisco, has analyzed the topic of computational

---

nanotechnology for its technological, medical, and environmental applications in the next century. The most influential progress that has been made in the last ten years has been the complementary discoveries of nanotubes and fullerenes – materials that can be used on the nanoscale to produce sensors and capacitors. These nanotubes and fullerenes allow scientists to create devices that can analyze items on the molecular level. Nanotechnology can now develop and eventually advance to the level of intentional DNA manipulation. Srivastava maintains the perspective that nanotechnology, due to the advances of nanotubes, can now evolve into a more defined field with applications in cancer eradication and genetic engineering. His views on the benefits of computational nanotechnology range from the prevention of genetic and destructive cardiovascular diseases to the implementation of microscopic sensors. These sensors would be used to study the interactions of molecules in the human body. On the other hand, a handful of Srivastava’s peers view computational nanotechnology as a dangerous field that could harm humans and computer systems. The perspective that nanotechnology will be a detriment to the human race contains flaws relating to the difficulty of operating molecular regeneration devices and nanotubes. In regards to the question of when artificial intelligence will surpass human intelligence, this answer relies on the current definition of intelligence along with how AI develops in order to achieve that intelligence. As each generation changes the definition of intelligence, artificial intelligence comes along and meets that level of “smartness.” This idea worries many scientists since they believe that we are on the verge of annihilation, due to the rapid successes of the world’s greatest – yet most dangerous – creation, artificial intelligence.

Most long-range forecasts of technological abilities in the future dramatically underestimate the power of future technology because they are based on the “intuitive linear” view of technological progress rather than the “historical exponential view.” To express this another way, it is not the case that we will experience a hundred years of progress in the twenty-first century; rather we will witness on the order of twenty thousand years of progress (at today’s rate of progress). This logarithmic plot demonstrates the exponentially shrinking gaps between key events in human intellectual history. Ray Kurzweil, a world-renowned computer scientist, has created "The Law of Accelerating Returns", according to which the rate of change of technological evolution tends to increase exponentially. If applied to Moore's Law, the exponential shrinking of transistor sizes on an integrated circuit, this law can describe exponential growth of diverse forms of technological progress. Currently, every consumer processor is constructed using silicon — the second most common element found in the Earth’s crust. However, silicon is not the perfect conductor, since there are limitations to the movement of the electrons, which affects the maximum density of these silicon transistors. In addition, an effect known as quantum tunneling can bring about issues of containing electrons past a specific “thickness threshold.” Besides research facilities, modern silicon transistors are not reduced to

---

any size smaller than 14 nanometers. Even though certain chips of 10 nanometers might be constructed in the future, it is obvious that in order to advance Moore’s Law over the next century, technology will have to develop and new, more capable materials will need to be used as the foundations for next-generation computational devices. An example of such a material is graphene – specifically rolled tubes of graphene known as carbon nanotubes. Since graphene is “atomically thin,” sometimes referred to as having two dimensions, it could be the solution to the continued miniaturization of transistors predicted in Moore’s Law. Unfortunately, the bandgap of graphene — or the energy difference required to move electrons across the material between the conducting and non-conducting bands — is not as effective compared to that of silicon. For silicon transistors, the mode of computation relies on the switching on and off of the conducting and non-conducting bands. Without this issue being solved, a graphene computer would have to initiate a completely novel, logical method for computing. Yet even with these difficulties, a graphene computer chip created by IBM demonstrated incredibly fast processing power, with a rate of over 10,000 times faster than the average silicon chip — however the chip cannot be used for normal everyday processes. Due to graphene’s inability to be switched on and off effectively, it would be unfeasible to replace silicon with graphene in large quantities while maintaining contemporary chip designs. In addition, Quantum computing is a potential solution to the difficulties of Moore’s Law; however, research is still in the beginning stages therefore it is still not confirmed. While certain computer scientists theorize that quantum computers will include an enormous and instant upgrade over current processors, it is evident that computer encryption will be impaired by the utilization of these new systems. Contrary to popular belief, quantum computing will not come to market as a programmable, digital computer immediately; original quantum computers will not be capable of operating with Windows, regardless of if the processor is thousands of times faster than theoretically necessary. In relation to all of the potential solutions for the impending issues of Moore’s Law, quantum computing is most likely the least economically and computationally feasible. While it may have great potential for very precise uses, this style of highly advanced computing for PCs is still too unexplored to be fully considered. Whenever a technology approaches some kind of a barrier, new technology will be invented to allow us to cross that barrier. Paradigm shifts have and will continue to become increasingly common, leading to technological change so rapid and profound that it represents a rupture in the fabric of human history. This rupture is commonly known as the technological singularity. The Law of Accelerating Returns implies that a technological singularity will occur before the end of the 21st century, around the year 2045.

As robots get smarter, humans just will not be able to keep up. While simpler tasks may be outsourced to robots at first, it is predicted that by 2045 machines will be billions of times smarter than un-augmented human beings. Robots have already replaced factory workers, and self-driving cars are just around the corner. Still other computer-science technologies, such as the ultrafast stock trading programs that cause "flash crashes," are being developed without considering how they could damage people or put them out of work. Unemployment due to artificial intelligence taking over certain jobs will cause economic difficulties and future
economic decline as automated systems replace human beings. Robotic and Artificial Intelligence machines have already invaded many of the world’s industrial manufacturing sectors. Manufacturing of automobiles, electronics, pharmaceuticals, and many other products are almost completely controlled by robotic systems that do not need much human interaction. While the United States ranks 7th in the world for robotic density, Japan and Korea come in at the top of the list with 339 and 347 robots per 10,000 people respectively. In the next decade, artificial intelligence-enhanced products are projected to create trillions of dollars in economic value. A large fraction of that should be invested in the ethics of autonomous machines, solving the A.I. control problem and ensuring humankind’s survivability. Pieluigi Paganini, Chief Information Security Officer at Bit4Id, depicts the perspectives of revolutionary computer scientists and experts such as Elon Musk, Stephen Hawking, and Bill Gates in his article, “Cybersecurity and Artificial Intelligence: A Dangerous Mix.” It is possible, and extremely dangerous, for a generation of superintelligent drones to believe that the human race is a detriment to its own survival, resulting in that generation declaring war on human beings. The successful advancement and implementation of artificial intelligence would be one of the most important achievements of the 21st Century; however, it is extremely imperative that the human race does not neglect to consider the possible consequences that this power entails. On the other hand, Google CEO Larry Page believes that the dissemination of these superintelligent machines will only lead to positive outcomes in the long run. He believes that AI will help to create more jobs and allow humans to maximize the efficiency of resource use. In this situation, artificial intelligence would actually benefit the global economy rather than negatively affect unemployment. Paganini provides a few possible solutions to counter the possibilities of malicious artificial intelligence systems. In order to prevent these systems from emerging, it is important for computer scientists to develop intelligent machines with a Hippocratic emphasis – meaning that the robots maintain a “first, do no harm” mentality. In addition, the author believes that computer scientists must supervise the development of AI and its use of natural resources.

Technology has always been a double edged sword, bringing us longer and healthier life spans, freedom from physical and mental drudgery, and many new creative possibilities on the one hand, while introducing new and salient dangers on the other. We still live today with sufficient nuclear weapons to end all human life on the planet. Bioengineering is in the early stages of enormous strides in reversing disease and aging processes. However, the means and knowledge will soon exist in a routine college bioengineering lab (and already exists in more sophisticated labs) to create unfriendly pathogens more dangerous than nuclear weapons. As technology accelerates toward the Singularity, we will see the same “intertwined potentials: a feast of creativity resulting from human intelligence expanded a trillion-fold combined with many grave new dangers” that must be considered. In the book Global Catastrophic Risks, Swedish philosopher Nick Bostrom presents the global risks and rewards of a future engulfed with artificial intelligence and highly advanced computational technology. One of the major

---

dangers of AI involves the misguided confidence that humans have regarding their understanding of this technology. Many people conclude that they comprehend the complex developments of artificial intelligence, yet the advancements occurring in the next 20 years will drastically affect how quickly the concept adapts and grows. Unfortunately, when humans constantly employ a certain technology in their everyday lives, they start to become accustomed to the technology and eventually they forget that it even exists. Bostrom, a professor at the University of Oxford, discusses three points of reasoning involving why artificial intelligence must be controlled and monitored within the next 15 to 20 years.\(^7\) He states that a powerful Artificial Intelligence source could defeat the defiance of the human race, eventually leading to the termination of humanity. In addition, the economic consequences of an AI takeover of the job market are considered in their global context. On the other hand, the author comments on the positive possibilities of AI within the medicinal community. He recognizes that the development of AI in the future will potentially save many human lives, yet Bostrom also realizes the difficulties that the planet Earth will need to overcome in order to sustain an enlarged population with reduced resources. In the article "Thoughts on Robots, AI, and Intelligence Explosion," Tomasik, a computer scientist from Microsoft, discusses how artificial intelligence will drastically alter society within the next 100 years. The institution of superintelligent beings will have a “soft takeoff” during which portions of industries implement these creations, as AI continues to develop, yet there is still a great chance for a “harder takeoff” that results in one AI machine exponentially advancing and causing financial/cultural chaos.\(^8\) The author states that if any type of rogue artificial intelligence were somehow able to take over the Earth, it could possibly create an end goal of attempting to colonize the galaxy while also continuing research and achieving many scientific or engineering feats. As this occurs, these beings would not show decent respect for the human race, including the prevention of suffering of less powerful beings – animals. The question of what a rogue-AI situation would consist of is analyzed and critically viewed for the possibilities. Overall, the ethical questions regarding artificial intelligence and its potential uses are very important realms in which philanthropists can make a beneficial impact in many regions of the world. In order to fully support the development of AI, it must be controlled, through a system of checks and balances, as a method of ensuring that only positive outcomes arise. Overall, the article focuses on how AI appears to be beneficial in the development stages; however, we have yet to see the consequences of this technology when it becomes smarter than its own creator.

For many computer scientists, Bill Gates being one, there is a fear that artificial intelligence programs will take over the job market and ruin the financial security for many people. “The potential of Artificial Intelligence is enormous and in fact a 2013 study by Oxford University

estimated that it could take over nearly half of all jobs in the United States in the near future.\textsuperscript{9} Independent of the future development of artificial intelligence and computation, almost all academics with knowledge in this area understand the ramifications of a world filled with AI along with how the technological world will undoubtedly change within the next few years. While many people fear the physical effects of an AI takeover, all of these sources reaffirm the belief that a “Terminator: Rise of the Machines” occurrence would be physically and technologically impossible within any reasonable amount of time. These statements directly negate the claim that a physical takeover of Intelligent Machines is impossible. The conclusion of the article, "Artificial Intelligence Is Changing the World, and Humankind Must Adapt.\textsuperscript{10}, touches on a question that has recently gained importance, technology vs. education. As technology advances rapidly, secondary education and other professional schooling must also improve in order to compete with the increasing competition from artificially intelligent machines.\textsuperscript{10} Robots might not ever threaten the life of a human being, yet they will most likely take the jobs of a thousand individuals. In "How to Avoid a Robot Takeover: Political and Ethical Choices in the Design and Introduction of Intelligent Artifacts.\textsuperscript{11}, the unlikelihood of a detrimental artificial intelligence takeover is analyzed as it relates to the potential real-world consequences. University of Sussex Professors Whitby and Oliver discuss how the idea of a robot takeover is possibly a journalistic ploy to scare the public away from a reliance on technology. By constantly reporting on the possibilities of an artificial intelligence “attack”, the public image of AI becomes dirtied by the negative publicity. The authors state that AI domination over the human race is a simplistic assumption that does not take into account the motivation of these robotic systems. Even if computational intelligence advances past the level of the human race, there is no tangible evidence that suggests that these systems will desire to dominate over humans. Another questionable assumption involves the statement that an increase in intelligence does not necessarily correlate to an increase in political power in our society. Furthermore, Oliver and Whitby conclude that the fate of the human race, as it relates to Artificial Intelligence, ultimately relies on the choices made by humans as they develop and implement these AI systems in the real world. The article uses logical reasoning in order to negate certain arguments regarding the inevitability of an Artificial Intelligence takeover.

Chad Jenkins, Associate Professor of Computer Science at Brown University, and Alexandra Peseri discuss the real-world possibility of a robot “take over” of the world through the gradual invasion of robots in everyday processes and the enhancement of the quality of life that will eventually occur due to this invasion. Rather than “sentient Skynet robot overlords.\textsuperscript{11} as depicted in movies, there will be more helpful robotic applications that work to increase overall human productivity and reduce inefficiency through an active human-robot collaboration effort.

This article is written as a way to inform the public about the specific questions that they should be considering when regarding the topic of robotic involvement in the future. The idea that computers will turn into killing machines that take over and rule the Earth as the superior being is a pure opinionated perspective. The answers to these questions analyze the future of robotics and artificial intelligence technology along with the possible societal impacts that these concepts may cause. The authors of this piece specifically address the problems of artificial intelligence in regards to labor, as more and more automated machines start to replace real human beings in the job market. This paradigm shift towards more efficient and cheap labor will eventually prove negative when unemployment rates skyrocket and the financial markets become seriously damaged. Furthermore, facial recognition and pattern recognition has been added to the functionality of artificial intelligent beings. This recognition software enables them to differentiate between everyday items just like a normal human being. The question of bandwidth capabilities and processing power of current technology always developing at an exponential rate depends on the resource issues and natural disasters that could prove harmful to advancement. In addition, deep learning will help to improve machine translation, which will allow computers to begin speaking, and interacting much more frequently. The improvement and development of deep learning on computing systems is a reason why artificial intelligence will become a dominant information source. This reasoning behind the argument that AI should be controlled significantly, allows for the research behind robots and artificial intelligence being detrimental to job security and efficiency in factories and companies across the globe. The claim that computers will take over the world could be negated by the fact that there is a greater fear being placed on the future of jobs for the white-collar working class, skewing the opinion of the public as a whole. The strengths of the claim that AI will be detrimental to the economy due to the computing power taking over job opportunities are that artificial intelligence is developing at an exponential rate, meaning that each year computing power is developed and increased at 10x, 100x, and finally 1000x the rate of the previous year. The weakness of the argument deals with the future of resource sustainability with regards to battery power, memory, bandwidth, and deep learning abilities.

Fortunately, the main claim that artificial intelligence will result in a final situation of robots becoming aggressive and violent may be negated by the three (most recently four) laws of robotics that will be used to control robotic AI creations in the near future. This law code ensures that humans are not affected by these intelligent beings, neither physically, economically, or culturally. In addition, the claim that computing power will reach a point of singularity in 2045 has support due to the computer science studies demonstrating the developments in the past 20 years and the exponentially increasing rate at which these advancements are created. In the article, “Is Artificial Intelligence Really an Existential Threat to Humanity?,” Edward Moore Geist – a MacArthur Nuclear Security Fellow at Stanford University's Center for International Security and Cooperation – discusses the ideas of super

---

intelligence in computers and the “existential risk” of artificial intelligence becoming dangerous to humans. The major idea behind the article rests in the hope that the author possesses that artificial intelligence beings will be created with a friendly relationship and law code with human beings. This law code is more commonly known as the three laws of robotic ethics. According to the original code, “A robot may not injure a human being or, through inaction, allow a human being to come to harm. A robot must obey the orders given to it by human beings except where such orders would conflict with the First Law. A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.” While these three laws could effectively be programmed into AI in order to prevent a disaster for the human race, it is impossible to predict if future computer systems will be able to violate these laws or if they will fall into the hands of a terrorizing force. Furthermore, this code of ethics also introduces the question of how a robot defines harm. Harm can be classified into emotional, physical, and psychological categories. However, for a superintelligent being, any event, activity, or product that causes (or has the potential to cause) these types of harm would become a reason for robotic intervention. This could unfortunately mean the termination of the sales of unhealthy foods, fast sports cars, firearms, and even soda. Television and video games can cause laziness and eventually lead to personal physical harm; therefore, Artificially Intelligent systems may discontinue access to these human-made entertainment programs for the benefit of society as a whole. Allowing this technology to operate under these laws and with a state of autonomy could bring about major restrictions to the freedoms and pleasures of human life – imposed by the very computers necessary for access. In addition, the creator of the law code – Isaac Asimov – established a modification to his established set of rules. This modification created a fourth law known as the “zeroth law” which precedes all others in order of importance: “A robot may not harm humanity, or, by inaction, allow humanity to come to harm.” The specific diction of this law is dangerous, in and of itself, as it only exacerbates the issue of robot-imposed restrictions as discussed above. If an artificially intelligent system were to analyze the freedom that humankind possesses, and notice that a more restricted world would not “allow humanity to come to harm,” then it is possible that this system could attempt to increase the limitations and constraints on the public. Geist states in the article that “the risks of self-improving intelligent machines are grossly exaggerated and ought not to serve as a distraction from the existential risks we already face, especially given that the limited AI technology we already have is poised to make threats like those posed by nuclear weapons even more pressing than they currently are.” This perspective states that the idea of AI taking over and affecting the human race is an exaggeration of the human mind. The other perspectives deal with famous computer scientists Nick Bostrom and Ray Kurzweil who theorize about AI reaching a point of singularity during which superintelligence will be so advantaged that humans will become subordinate. These opinions are backed by the research of the rate of increase of artificial intelligence development in the last

two decades. This rate is exponential; therefore, it is possible that we have a dangerous future ahead. The claim that artificial intelligence will not reach a point of singularity is an opinion since every major computer scientist has researched the possibility of the exponential rate becoming a detriment to human beings. The argument against the ideas of Kurzweil is not supported by fact since there is no evidence that the point of singularity will ever exist or that we will ever reach it. Since the evidence is nonexistent, this opinion is difficult to consider due to the major theoretical question. The author suggests the AI will not be like “summoning the devil”, however it will lead to unemployment and the intelligence level of the human race could be affected by the increased reliance on AI. The conclusion that AI posing an existential threat to the human race is false implies that the future economy and job market will continue to dwindle as time goes on. If jobs are created in abundance and AI is used in conjunction with the employees, AI will never threaten the economy. In addition, the argument that singularity does not exist since there is no evidence of a possible event implies that the exponential rate of development will level off in the near future and computer systems will not be advanced. The strength behind the claim that AI will only threaten the intelligence of the human race due to the reliance on computers for information is that computing power is developing in order to meet the demand of humans requiring information quicker and cheaper. The weakness to this argument is that the computers are being used to increase information capabilities; therefore, the human race is technically increasing in intelligence rather than losing intelligence by employing AI. The valid conclusion, stated at the end, is that AI will not be an existential threat, due to a lack of physical harm or time for development. However, this conclusion is weak because it ignores the possibility of terrorist forces taking control of AI systems and the worldwide movement to store every piece of necessary and private information on the cloud – an easily accessible vault. More evidence will be needed to find out whether or not technology will be able to develop at this rate with all of the resource depletion and data usage.

Artificial intelligence, if it continues to advance exponentially in the near future, has the potential to make lives easier through the understanding of human desires along with the ease of transportation, but if left uncontrolled, the technology will pose a serious security threat to our society as a whole. There are many questions to consider when it comes to the secure development of this technology. As intelligence increases, will these holds continue to work or will they be faulty due to the faster, smarter computational programs of AI? Since this technology develops rapidly, the current security measures would not be sufficient to prevent an advanced cyber attack/takeover. If/when Artificial Intelligence becomes self aware, will they become a new race? It is important to consider the ethical situation regarding the classification of self-aware, artificially intelligent beings. Artificial Intelligence could be the greatest invention in the history of humankind; however, it is necessary for researchers and analysts to determine the existential risks of the development of advanced intelligent beings. Overall, artificial intelligence must be controlled through a system of checks and balances in order to prevent economic volatility of financial markets, massive unemployment, and ethical/cultural tension.
Works Cited


