Jonathan Phouminh

WRTG 1150-0077

Essay 3

Word Count: 2438

**Implementation of AI : Are Humans Ready?**

**Abstract**

Artificial intelligence is a branch of statistics and computer science that focuses on the integration of automated and self learning machine technologies that are intended to serve towards human interests and improve overall convenience of life. This literature review will analyze scholarly research on the question of whether humans are ready for the full integration of artificial intelligence at its current stage in development. We will ultimately see that research around artificial intelligence is pointing to the idea that although there have been successful implementation of technologies featuring artificial intelligence, humanity is not ready as scholars believe that humanity is ill-prepared for the dangers that come along with the technology and further studies are required. Lastly, with how relatively new the study of artificial intelligence currently is there are many unanswered questions regarding the focus of future research. Also, the impact on societies ethics and standards when AI technologies are implemented into society is still left unanswered.

**Introduction**

Official artificial intelligence research studies began in the 1956 when computer scientists and mathematicians began discussing the possibilities of creating an artificial brain for the purpose of creating technologies to serve human interests and improve convenience. One of the most discussed questions around artificial intelligence is whether or not humanity is ready for the integration of artificially intelligent technologies in society. Examples of artificially intelligent technologies include self-driving vehicles, self-piloted military drones, and companion / caretaker robots.This is a popular question among the science community because of the science community with the current levels of artificial intelligence still pose high probability of potential dangers that are unknown if humanity is ready to face. The debate among this question is split in two ways for those who fear the integration of artificially intelligent technology and those who are for the technology as they believe the benefits outweigh potential risks. Looking at both sides of the debate, few researchers have argued that current levels of artificial intelligence knowledge is sufficient to be implemented into technologies which shall be used in society while others argue that current levels of artificial intelligence is not. Researchers that are for integration for artificial intelligence have pointed to early implementations of the AI technologies that have already proven to be beneficial such as Tesla’s self driving vehicles that more secure in terms of safety in comparison to human drivers (Marr 1). Conversely, researchers who oppose integration of artificial intelligence have acknowledged that although humanity’s current level of knowledge around the technology has proven to show its benefits in early applications of AI technologies, people are not fully prepared for the economic dangers and ethical problems that will arise if the technology goes wrong. This literature review will analyze current research that has been done around artificially intelligent technologies to show the debate of whether or not humanity is ready to fully implement artificial intelligence into society. By analyzing common methodologies used to study artificial intelligence, why some researchers believe humanity is ready for artificial intelligence technologies and why some do not, and understanding who scholars believe is at fault for when the technology goes wrong.

**Trends In Methodology of Research**

Most researchers employ similar methodologies when it comes to collecting long term data reports surrounding artificially intelligent technology. Thibault De Swarte, author of the article, “Artificial Life and Robotics,” conducts research on artificial intelligence’s impact on humanity by studying two case studies where artificial technology is introduced in two different applications in society: military drones and caretaker robots for the elderly. He considers both immediate impact and also potential future outcomes by exploring different scenarios where AI technologies could failed to perform their task. After collecting information, Swarte concludes that humanity not fully ready for integration of artificial intelligence and states, “Artificial intelligence in its current state is very primitive in its decision making,” (Swarte 8) Swarte alludes to the idea that artificial intelligence decision making skills are not advanced enough to make an impact on society when considering applications in military drones and robotic caretakers. Swarte concludes that there are still too many variables that artificial intelligence unaccounted when making decisions which results in more potential harm towards humanity than good (Swarte 8). James Dewey takes a similar approach to conducting research by examining how existing artificial intelligence technologies have already made impacts on humanity. Similar to Swarte, Dewey also acknowledges that a few early implementations of artificial intelligence has proven to be beneficial by showing statistics, but states, “Machines and robots that outperform humans across the board could self-improve beyond our control,” (Dewey 1) Dewey emphasizes that humanity is not ready for the technology by discussing the pitfalls that must be handled if it were to go wrong. Dewey is able to conclude this statement by demonstrating how artificial intelligence learning improves exponentially. Clearly, Dewey and Swarte were able to draw their conclusions by analyzing long term data collection and justifying their position by discussing potential pitfalls from the data.

Next, in the article, “Tesla Explains How A.I. Is Making Its Self-Driving Cars Smarter,” the author also used a similar method to Swarte and Dewey by analyzing long term data of effectiveness to draw conclusions around the artificial intelligent technologies and their impact on society. This article examined the process of how Tesla vehicles are able to use data collection to make the quality of autonomous driving vehicles better for humans. With the abundance of data collection, corrections to bad driving patterns can be made from new data and be installed to the vehicles on the road making for a safer and higher quality ride experience (Ryan 2).

Comparing these authors methods for research we see that the popular method for determining if artificial intelligence is ready for full integration into society is by looking at historical impacts of artificial intelligence and thus discussing pitfalls that still exist or highlighting positive trends in history that show humanity is ready for AI.

**Benefits of Artificially Intelligent Technologies**

The researchers that are for the integration of artificial intelligent technologies also share similar the same approach for arguing for the technology. Researchers, Nabi and Marr, highlight areas where AI technologies have proven beneficial and also address potential benefits. For example, in the article, “How Bioethics Can Shape Artificial Intelligence and Machine Learning,” written by Junaid Nabi, she states, “Empirical research has demonstrated that the latest generation of AI and machine learning applications cannot only identity at‐risk patients but even achieve better diagnostic performance than expert physicians can,” (Nabi 2) Nabi argues that implementation of artificial intelligence in hospital settings will improve efficiency of the hospitals and also create a better work-life balance for health practitioners. She discusses how modern hospitals and health practitioners are already performing at a high level, but their performance can be greatly elevated if artificially intelligent technologies were implemented. Practitioner performance can be greatly improved by allowing the technologies to perform the small tasks that eat up most of their time such as making diagnoses (Nabi 9).

Similarly, Marr focuses on how artificially intelligent cars have proven to be beneficial to humanity already since autonomous cars are statistically shown to drive safer then most man piloted vehicles (Marr 1). Autonomous vehicles are able to drive safer with more data collection surrounding driving patterns which allow the artificially intelligent vehicles to make quicker judgements in driving situations (Marr 1). Marr goes further to explain that artificially intelligent cars will only get safer as more driving pattern data is collected thus improving the quality for humanity’s use, showing that humanity is ready for full integration of artificial intelligence. Marr takes an almost identical stance as the same author of, “Tesla Explains How A.I. Is Making Its Self-Driving Cars Smarter,” as both scholars highlight the benefits of self driving cars to justify their positions in the debate. \*\* clean up the conclusion

All in all, these researchers share the general consensus for the implementation of artificial intelligence by highlighting aspects in areas of society where the technologies implementation has been proven to be beneficial towards society and or can be justified to being beneficial.

**Fear of Artificial Intelligence and Ethics**

Conversely, there are researchers that have come to a general consensus by considering the pitfalls or dangers of artificial technology if they were too arise. Researchers have expanded on pitfalls of artificial intelligence including economic instability such as job loss and also acknowledging risks that these underdeveloped technologies have on people.

To start, in the article, “Anticipating Artificial Intelligence,” the author begins to develop their idea that humanity isn’t ready for artificial intelligence technology by discussing how far artificial intelligence research has come in since the 1950’s. The author acknowledges how implementation of the technologies have already proven to beneficial towards humanity, but establishes his viewpoint by stating that people are not ready to burden the economic dangers that will come with the technology (Dewey 5). By pointing to past examples in history, the author furthers the idea of economic instability due to artificial intelligent technologies by showing when artificial intelligence has taken jobs from humans in factory settings to supermarkets. Thus, the author concludes that since economic instability in the job sector for humans is already a issue the introduction of AI technology that potentially furthers economic instability is not a problem humans are ready to face (Dewey 8).

Similarly, Swarte takes a similar stance to Dewey as he argues that humanity isn’t ready for full integration of artificial technology. Swarte demonstrates his stance by examining applications of artificial intelligence in military drones and robotic caretakers. First, Swarte looks at the application of artificial intelligence in military drones that are meant to be used in combat and concludes that humanity isn’t ready to allow this technology. Swarte believes that current level of artificial intelligence isn’t capable of distinguishing between hostile combatants and civilians which would result in potential false casualties or failure of critical assignments (Swarte 9). Next, Swarte looks at the application of artificial intelligence in robotic caretakers. He defines robotic caretakers as robots capable of taking care of humans who are incapable of coping psychologically or performing certain kinds of physical activity. Swarte states that if the technology were to go wrong in the role of caretakers then the robots could potentially worsen the condition of patients if incorrect decisions were made (Swarte 10). Swarte concludes that further algorithmic studies are needed in artificial intelligence before full implementation can be considered.

**Ethics of Artificial Intelligence**

Next, scholars have also justified their position in the debate by discussing current ethics of society and how they play a role when AI technologies failures and successes are considered.. In the article written by James Overden, his research focuses on the ethics side of artificial intelligence and considers three cases for who to blame if the technology fails to perform its intended task. Overden discusses three cases in detail for either blaming the buyer, the creator, or the actual technology itself. After considering all three cases for deciding who is to blame when the technology fails, Overden concludes that the answer will remained unanswered until we figure out how morals and ethics of society will have to change when living with an artificially intelligent world (Overden 5). Therefore, Overden concludes that humanity is not ready for AI technologies based on his understanding of societies current set of ethics. Conversely, in the article, “Ethics of Artificial Intelligence”, written by Stuart Russell, Russell discusses his research around artificial intelligence by focusing on how society reacts to the concept of artificial intelligence rather than the actual technology. Russell states that currently there is a negative connotation that arises when people hear artificial intelligence due to exaggeration of the dangers of artificial intelligence in media (Russell 7). The most significant point that Russell emphasizes in his research is understanding that humanity shouldn’t shy away from artificial before understanding the true benefits towards society that can be made if implemented correctly. Russell then suggests that there are more of the benefits of artificial intelligence than dangers and people are actually ethically accountable for doing what is best for humanity. (Russell 11).

**Future Studies**

Because of how relatively new the field of study of artificial intelligence still is, future studies encouraged by researchers have ranged from focusing on actual development of artificial intelligence to focusing on how implementation of artificial technology will immediately impact society.

To start, some researchers suggest that future studies will be focused on actual development of artificial intelligence. In the article, “How Bioethics Can Shape Artificial Intelligence and Machine Learning,” Nabi conducts her studies around implementation of artificial intelligence technologies into the healthcare system. Nabi states that if artificial intelligence can successfully be implemented then efficiency in services will increase dramatically (Nabi 6). Therefore, Nabi suggests that future research should be focused on how artificial intelligence can be implemented into the healthcare system and also, at the same, time be able to be understood by those who are not technologically savvy (Nabi 8). Similarly, in another article, “Tesla Explains How A.I. Is Making Its Self-Driving Cars Smarter,” Ryan discusses how Tesla vehicles are making smarter driving decisions and becoming smarter through long term big data collection (Ryan 2). Ryan explains that through engagement of big data collection, Tesla can make improvement in algorithms used by the vehicles that judge driving scenarios and are able to predict events such as slow moving traffic and possible accidents (Ryan 4). Thus, by highlighting where artificial intelligence is flourishing, Ryan suggest that future research should also be focused on actual development and studies around long term data collection and machine learning. \*\* maybe restructure ethics

**Conclusion**

The question if humanity is ready for the integration of artificial intelligence into society is still highly debatable between researchers, but the majority of researchers seem to lean towards holding off on actual integration of the technology until the technologies are able to perform at a higher level and also until humanity understands how integration of the technology will change society. Based on trends of researchers discussed in this review it seems that future research around artificial intelligence will branch off in the direction of development and ethics. A few questions that still remain around the issue are; at what level of knowledge and capability will the artificial intelligence need to be to be considered being implemented into society, and how will the integration of artificial intelligence influence career paths of young adults if job loss in certain sectors occur. With technology capable of changing the way society functions there are certainly many great opportunities to be reaped, but it also comes with dangerous expenses if implemented incorrectly.

**Bibliography**

Vamplew, Peter. “Ethics and Information Technology.” *Human-Aligned Artificial Intelligence Is a Multiobjective Problem*, vol. 20, no. 2, ser. 1, 4 Oct. 2017, pp. 27–40. *1*

Tusinski, Kati. “Journal of Media Ethics .” *The Ethics of Artificial Intelligence: Superintelligence, Life 3.0 and Robot Right*, vol. 3, no. 23, 25 July 2018, pp. 151–153.

Russell, Stuart. “Anonymous . Nature.” *Ethics of Artificial Intelligence*, vol. 521, no. 7573, ser. 1, 28 May 2015, pp. 415–418. *1*.

De Swarte, Thibault. “Artificial Life and Robotics.” *Artificial Intelligence, Ethics and Human Values: the Cases of Military Drones and Companion Robots*, vol. 1, no. 13, 28 Jan. 2019, pp. 1–6.

Dewey, James. “Tech Principles.” *Anticipating Artificial Intelligence*, vol. 534, no. 7746, 28 Apr. 2015, pp. 1–1.

Nabi, Junaid. “The Hasting Center Report.” *How Bioethics Can Shape Artificial Intelligence and Machine Learning*, vol. 48, no. 5, 11 Oct. 2018, pp. 40–55.

Ovenden, James. “Innovation Enterprise.” *Who Pays When AI Goes Wrong?* , vol. 3, no. 24, 15 Sept. 2017, pp. 1–1.

Ryan, Kevin J. “Big Data.” *Tesla Explains How A.I. Is Making Its Self-Driving Cars Smarter*, 13 Sept. 2016, pp. 1–1.