An Analysis and Report on the Ethics of Artificial Intelligence

Jordan Mason

November 2018

COM 203

Introduction

From the time when artificial intelligence was almost exclusively science fiction, the consequences of AI have been constantly in question. As this concept has become more and more real with the passing of time, so too have the consequences. Controversy has surrounded the use of Artificial Intelligence, especially in cases of AI replacing human workers. Even so, it has become more and more prominent in modern society. In every application, however, the use of AI has been viewed in different ways, with different consequences and different opinions in situations of law, military weaponry, life-and-death decisions, allocation of responsibility, and many other situations.

Background

While there is no easy solution to the controversy surrounding Artificial Intelligence, it can be broken down into a series of categories where machine-oriented decision-making is most heavily debated. All of these situations have very real consequences, many of which are permanent or life-changing. This report will focus on controversy surrounding the use of Artificial Intelligence in decisions of law, how responsibility for the decisions of AI can be allocated, the risk of errors in programming, the risk of cyberattacks, human error in programmed decision-making, misunderstandings occurring between AI and human workers, the economic consequences, the use of AI in military weaponry, and the effects that AI implementation can have on medical science.

Five Major Risks

While the implementation of AI comes with a series of risks, the most fundamental is the chance that it might not work as intended. Bugs are common in most computer programs, and are surely inevitable in the development of new Artificial Intelligence programs. As AI replaces human workers and human decisions, these bugs could have very costly or dangerous consequences, which has led to plenty of controversy from the side of skeptics. Sometimes commands that are not necessarily bugs can still be very disastrous, as well. For example, with self-automated vehicles, there are many factors that come into the decision-making process of the vehicle's programming. There are many parameters that must be met for an action to be taken, or for inaction to be decided as the most valuable option. Because many decision-making AI programs learn from previous experiences, it is possible that their learning can confuse the machine when more complicated decisions are presented. This, of course, can result in catastrophic ends if the wrong action is taken, or no action is taken when action is needed.

While cybersecurity is something to consider with any computer program, it is especially important when the consequences of security breaches could be matters of life and death. Cyberattacks on AI systems could result in altered learning, which would affect the behavior of machines that rely on these systems. When these attacks could result in the misuse of automated vehicles and weapons, or the displacement of large amounts of money, faults in cybersecurity could be catastrophic to the lives of people affected. "Before we put AI algorithms in control of high-stakes decisions, we must be confident these systems can survive large-scale cyberattacks" (Dietterich).

The next risk is different but comparable to the concept of failed learning. It is possible that the programmer of the Artificial Intelligence might fail to foresee situations where the AI would have to make decisions it isn't prepared for. It must be appropriately taught on how to respond to every command in every way. While not a "bug," a human programmer failing to appropriately instruct an AI in how to carry out its decisions in a nuanced and safe way is very possible and very dangerous. It is important that intentions and instructions do not get confused, because machines do not understand information in the same way that humans do. Artificial Intelligence programs must be programmed with this in mind.

A fourth major risk in the implementation of AI is the confusion of collaborative efforts between AI and human workers. Introducing Artificial Intelligence to a large number of workers, or into a situation where workers have to take over the job of AI at various intervals, can result in confusion and miscommunications due to a lack of understanding regarding what the machine has done. "For example, airline accidents have been linked to misunderstandings arising when pilots took over from autopilots" (Dietterich).

For a long time, workers have been afraid of losing their jobs to machines. As Artificial Intelligence programs are essentially created to do the jobs that human workers once did, this is a realistic fear. However, it replaces those jobs with new ones – specifically, the need for programmers and officials to maintain and operate those AIs. This does, however, result in a wage gap, due to the fact that maintaining an Artificial Intelligence program is a more complicated and higher-paying job than the work that the machines replace. In order for the implementation of Artificial Intelligence to be

economically responsible, it is important that the consequences be fully understood and dealt with. Eloquently put, “These questions move beyond computer science into the realm of economic policies and programs that might ensure that the benefits of AI-based productivity increases are broadly shared” (Dietterich). While these Artificial Intelligence programs are not designed with the purpose of redistributing wealth, it is important to understand that this could be a very real consequence as AI continues to grow.

AI in the Court of Law

With Artificial Intelligence replacing more and more jobs in the workplace, it comes as no surprise that a robotic lawyer has been in development. Named “ROSS”, this AI has been programmed to scan cases of law and make decisions based on the instructions it has been taught (Nunez, 2017). Controversy has surrounded the theoretical use of this machine, as its decisions have very heavy consequences for the people involved in those cases. This robot analyzes past cases as well as the given input regarding the current case, trying to assess the information as a human does. However, the person responsible for programming the robot has to decide how the AI determines the results of a case. This is where the controversy begins. In situations of the law, verdicts are not as simple as a “yes” or “no” response – there is not a pre-determined solution to every specific case, or else a jury and a court hearing would not be necessary. This is why it is handled by humans, and multiple humans, who can use reasoning and evidence to come to conclusions. A robot, naturally, is only as good as it is programmed to be, which also means that the programmer’s biases and faults will carry over into the decision-making AI.

The Use of AI for Military Weaponry

As progress has been made in the development of Artificial Intelligence programs, talk of AI-operated military weaponry has become a source of controversy. The concept of weapons that automatically select targets, including humans, with the intent to destroy or kill has – reasonably – been met with many questions. This is different from a human-operated weaponized drone, as the AI would operate completely independent of human control. Germany and Japan have made claims that they have no interest in weapons such as these (Russell, 2015). Agreements within the Geneva Convention are too vague to fully ensure that Artificial Intelligence programs fulfill the requirements, and so a series of meetings have been had to determine the morality in using these machines. Despite this, multiple countries have outlawed their use.

The Use of AI in Medicine

Despite the controversy surrounding various other applications for Artificial Intelligence programs, the use of AI in the field of medicine has been especially important in more recent years. It has grown and been generally accepted, becoming a popular replacement for other methods and procedures. In surgery and radiation diagnosis, machines are much more precise in their methods and calculations (Gill-Cox, 2018). Many procedures in the fields of medicine have been taken over by machines and artificial intelligence, to make up for the flaws in humans and to allow for quick machine-learning.

Allocation of Responsibility

In addition to the consequences of utilizing AI over human work, there are questions circulating about how responsibility for these consequences would be allocated. The Canadian federal government, for example, has already lent their focus to the task of creating regulations for AI, and holding companies responsible for the use of AI. "Consumers will hold companies responsible for how AI is used and operated, states Sage, and that accountability will lie with C-suite executives" (Simpson, 2018). An effort has been made to both regulate and normalize the use of Artificial Intelligence in British companies, working to build future industries involving the responsible use of AI.

Conclusion

The various risks of the implementation of Artificial Intelligence programs appear in many various situations where machines have replaced human workers, and the most prominently feared of these risks can be categorized simply. Bugs, cyperattacks, failed machine learning, confusion due to human interference, and a potential wage gap or other economic issues are topics that fuel controversy the most. However, their applications in medical science, military weaponry, and decisions of law all face different consequences, assuming these risks become real, observable problems. Regulations on the allocation of responsibility for these programs has also been a topic of focus for multiple countries.

Works Cited

Deng, B. (2015). Machine ethics: The robot's dilemma. Nature, 523(7558), 24–26. https://doi-org.saintleo.idm.oclc.org/10.1038/523024a

Dietterich, T. G., & Horvitz, E. J. (2015). Rise of Concerns about AI: Reflections and Directions. Communications of the ACM, 58(10), 38–40. https://doi-org.saintleo.idm.oclc.org/10.1145/2770869

ELLIOTT, S. W. (2018). Artificial Intelligence, Robots, and Work: Is This Time Different? Issues in Science & Technology, 35(1), 40–44. Retrieved from https://saintleo.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=133097078&site=ehost-live&scope=site

Ganapathy, K., Abdul, S. S., & Nursetyo, A. A. (2018). Artificial intelligence in neurosciences: A clinician's perspective. Neurology India, 66(4), 934–939. https://doi-org.saintleo.idm.oclc.org/10.4103/0028-3886.236971

Geist, E. M. (2017). (Automated) planning for tomorrow: Will artificial intelligence get smarter? Bulletin of the Atomic Scientists, 73(2), 80–85. https://doi-org.saintleo.idm.oclc.org/10.1080/00963402.2017.1288435

Giannetti, W. (2018). Artificial Intelligence: Myths and Realities. Air & Space Power Journal, 32(3), 92–95. Retrieved from https://saintleo.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=131615636&site=ehost-live&scope=site

GILL-COX, A. (2018). Consumer Health Applications, Machine Learning, and Systems Neuroscience: The Use of Artificial Intelligence Algorithms in Clinical Medicine and Healthcare Delivery. American Journal of Medical Research, 5(2), 46–51. https://doi-org.saintleo.idm.oclc.org/10.22381/AJMR5220186

Giuffrida, I., Lederer, F., & Vermeys, N. (2018). A Legal Perspective on the Trials and Tribulations of Ai: How Artificial Intelligence, the Internet of Things, Smart Contracts, and Other Technologies Will Affect the Law. Case Western Reserve Law Review, 68(3), 747–781. Retrieved from https://saintleo.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=130372095&site=ehost-live&scope=site

Gordon-Murnane, L. (2018). Ethical, Explainable Artificial Intelligence: Bias and Principles. Online Searcher, 42(2), 22–44. Retrieved from https://saintleo.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=128582745&site=ehost-live&scope=site

Hauer, T. (2018). Society and the Second Age of Machines: Algorithms Versus Ethics. Society, 55(2), 100–106. https://doi-org.saintleo.idm.oclc.org/10.1007/s12115-018-0221-6

Lu, W., Tong, Y., Yu, Y., Xing, Y., Chen, C., & Shen, Y. (2018). Applications of Artificial Intelligence in Ophthalmology: General Overview. Journal of Ophthalmology, 1–15. https://doi-org.saintleo.idm.oclc.org/10.1155/2018/5278196

Nunez, C. (2017). Artificial Intelligence and Legal Ethics: Whether AI Lawyers Can Make Ethical Decisions. Tulane Journal Of Technology & Intellectual Property, 20189-204.

Ravid, S. Y., & Xiaoqiong (Jackie) Liu. (2018). When Artificial Intelligence Systems Produce Inventions: An Alternative Model for Patent Law at the 3A Era. Cardozo Law Review, 39(6), 2215–2263. Retrieved from https://saintleo.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=130946294&site=ehost-live&scope=site

Russel S., Hauert S., Altman R., Veloso M.. Robotics: Ethics of artificial intelligence. (2015). Nature, 521(7553), 415–418. https://doi-org.saintleo.idm.oclc.org/10.1038/521415a

Simpson, M. (2018, August 16). Keeping ethics in artificial intelligence. Retrieved September 17, 2018, from https://www.itworldcanada.com/article/keeping-ethics-in-artificial-intelligence/408090

Stern, S. (2018). Introduction: Artificial Intelligence, Technology, and the Law. University of Toronto Law Journal, 68, 1–11. https://doi-org.saintleo.idm.oclc.org/10.3138/utlj.2017-0102

Toward More Ethical AI: STANDARDS WILL HELP INTEGRATE TECH INTO SOCIETY. (2018). BizEd, 17(5), 66. Retrieved from https://saintleo.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=131675248&site=ehost-live&scope=site

Vardi, M. Y. (2016). The Moral Imperative of Artificial Intelligence. Communications of the ACM, 59(5), 5. https://doi-org.saintleo.idm.oclc.org/10.1145/2903530

Wei, F., Sheng, D., & Lili, W. (2018). Evolutionary Model and Simulation Research of Collaborative Innovation Network: A Case Study of Artificial Intelligence Industry. Discrete Dynamics in Nature & Society, 1–13. https://doi-org.saintleo.idm.oclc.org/10.1155/2018/4371528