

Ethical Concerns and Issues in Artificial Intelligence

Term Paper Assignment

By

GIRISH KUMAR RAMACHANDRA

CWID 888253630

CPSC 548-02 (14087)

Fall, 2019

Professor: Michael D. Falkow

Department of Computer Science

California State University, Fullerton

November 2019

Abstract

The capacity of Artificial Intelligence (AI) and Autonomous Systems (AS) has been highly advanced since its term and concept was coined in 1956. Ever since these Intelligent agents have been in a race to automate and handle tasks in almost all aspects of human life in order to reduce human efforts of doing tasks that are redundant, time consuming , ultimately striving to be effective and efficient. AI in healthcare industry have been so much advanced that there are robots which could perform minimally invasive but a professional level surgery. Autonomous self-driving cars are no more an imagination anymore. Meanwhile, Robotic Process Automation is an art of scripting a bot in order to work at the application level, which could automate tasks that millions of humans do redundantly every day. Having said that, the growing use of AI in daily human routines, the question of ethics in their decision making, and thinking capacity need to be addressed in the first place. And the terrible impacts they cause to human life on the flip side of a coin is the problem that is paramount. But, how do you think these values or ethics can be implemented in these systems? This Paper tries to portray a few of these ethical concerns, where AI impacts human life.

Key Words: Artificial Intelligence, Autonomous cars, Autonomous Systems, Healthcare, Robotic process Automation, Values, Ethics, Ethical concerns, Ethical issues.

Introduction

Artificial Intelligence as a term was coined in a conference that was organized at Dartmouth college in 1956, (Xiuquan Li and Tao Zhang, 2017) says that it has been more than 60 years since then, and the improvement in the technology has been tremendous. Especially in last few years the progress has been remarkable. In 2011 IBM introduced Watson, which is a cognitive system that was successful in defeating the Jeopardy's champions, Brad Rutter and Ken Jennings. It was a tremendous moment as in the history of the universe, a non-human intelligence dominated and won humans. Similarly, in 2016 , Alphago , which is google's AI system, dominated a top world's chess player Lee se-dol. As technology improves, there will be more powerful machines and intelligence involved in order to dominate human intelligence and also become mainstream human assistants who can do works like cleaning the house, driving cars, etc. just like humans do (Xiuquan Li and Tao Zhang, 2017).

Time is so dynamic that science fiction can become real facts beyond imagination, and within a short span of time, internet usage has been matured to redefine the way we communicate with others and society (J. Banks, 2018).

Meanwhile, auto manufacturers and MNC's are in a technological race to build autonomous vehicles with state of the art driving technologies , where each company has their own build and operate models which affect the final prices of the product making it a question of affordability.

As the question of safety and security about autonomous vehicles and systems are being answered and assurances promptly pouring in , the society , in no time will start to adopt the technology in its day to day life to such an extent that there will be very few options left in order

to slow down the process especially when it becomes affordable at the grass-root levels (N. Hutchins, Z. Kirkendoll, and L. Hook, 2017).

The health care industry has also quickly embraced AI technology, and it has started to reap the benefits from the technology. It is no wonder now that the application of robotics and autonomous machines and sensors that carry out automated tasks are highly equipped with intelligence that can challenge human work effectiveness. Hence, machines that can adjust to the environment implement the AI are being keenly explored (J. Banks, 2018).

To be honest, tech giants have already started putting these technologies to use at the ground level, and their usage is now not at all a dream. Robotic machine, da Vinci surgical system, can operate minimally invasive surgeries successfully. Several robots help people rehabilitate from their post-stroke situations. Several robots are in use , helping the elderly lead their lives better by assisting them in day to day lives (J. Banks, 2018).

Meanwhile, Robotic Process Automation (RPA) , technology , the world of bots , where we can create software robots in order to perform tasks that are redundantly performed by humans. In other words, RPA automates human tasks, which can work 24/7 efficiently, ultimately being cost effective for the companies. Previously, automation of cross border applications was a difficult task, and now, with the intervention of RPA, things have completely changed.

RPA has opened its gates to not just secondary sectors like manufacturing firms, wholesale and retail businesses, but also to tertiary sectors like banking, law firms, IT sectors , automating their redundant tasks like data entry, business processing units, and knowledge processing units. Big 4 companies have already implemented RPA in their day to day operations. (D.Schatsky, C.Muraskin, and K.Iyengar, 2017).

We could witness that most of the daily affairs tilting towards automation, in the near future, many of those jobs which were performed by humans would be taken over by machines, and we have already started witnessing it. Maybe one-day individual jobs might disappear and give birth to fundamentally different types of jobs, which would require extreme knowledge to design these bots, which many would fail to do, due to their level of exposure, which keeps people in the state of instability striving for better knowledge every day.

Having said that, AI has already started replacing humans in many aspects in fields, not just healthcare, autonomous vehicles, and Robotic process automation but also several other fields that we will come across in this paper. By reviewing the upsides of AI, and the tasks performed by them in our life we could notice that they are becoming influential, with them arises the question of ethical concerns as these systems have started impacting us (V. Vakkuri and P. Abrahamsson, 2018).

Likewise, (A. R. Honarvar and N. Ghasem-Aghaee, 2009) say that every coin has another side, AI too has its downside, which could impact our lives if we do not consider several aspects while implementing them into our day to day tasks. Privacy and security are paramount. They could cause instability in people's lives if we fail to identify them.

AI could end up producing destructive machines which can harm mankind and behave immorally, as technology advances AI/AS play a prominent role in people's life. And the key factor which decides robot and human nexus is trust while designing. Autonomous systems, when designed, should be capable of making decisions in the key situations, and it could be flawed too sometimes; humans are not immune from blames, will machines be?

Hence more the freedom given to a machine more will be the human responsibility to design moral aspects in them. Artificial moral agents (AMA) are systems that are capable of making

moral decisions, and designing them is a challenging task, perhaps most importantly relevant to today's situation when machines are replacing humans in several key aspects (A. R. Honarvar and N. Ghasem-Aghaee, 2009).

1. AI in Healthcare

Introduction of AI and ML to the field of health care industry has taken it to a new level of the most advancing technology where it is grasping all the focus of attention to be in the headlines.

These technologies are huge in the market, and the firm Tractica forecasts that shipments and orders related to healthcare would increase threefold up to 10k units by 2022 (J. Banks, 2018).

Another research suggests that robots used for surgeries and their market would double up from 3 billion dollars in 2013 to 6 billion dollars by the end of 2020 (J. Banks, 2018).

For the planning of several deadly treatment programs and diagnosis of critical illnesses, Artificial Intelligence and Machine learning algorithms play an extremely important role.

In a very short time, this expert system scans the quickly growing body of scientific literature and assists the medical practitioner in achieving more accurate results in their assessments of critical cases of illness in a very short time. (V. Vakkuri and P. Abrahamsson, 2018) states an example, AI is used to gather and analyze symptoms in new GP at Hand service launched in London, which checks symptoms using a mobile application and schedules a video consultation appointment with a doctor within the coming two hours.

1.1 For Good or bad?

Every action has its advantages and disadvantages, so does AI and robotics. It looks like the world can differentiate R2-D2 and the relative Terminator. It seems that there are enormous concerns about the perspective of the technologies, even though we have shifted from science fiction to science facts. According to the researches done by several professors and scientists such as Prof Stephen Hawking and Prof. Nick Bostrom of Oxford University stated that AI could spell the end of the human race and brought forward the sentiment that AI could overtake human capabilities by 2045 (V. Vakkuri and P. Abrahamsson, 2018).

Moreover, with the advancement of technologies, several patients are willing to adapt to the latest sophisticated medical treatments. A research performed by PwC (formerly, PricewaterhouseCoopers), found that 39% of the patients from the United Kingdom are willing to participate in the healthcare system involving AI, Machine learning, and robotics (V. Vakkuri and P. Abrahamsson, 2018). Their people are mostly aged 18 to 24; this figure rises to 55%.

The latest machines in the health industry with the implementation of AI and machine learning, such as brain pacemakers, cardiac pacemakers, smart pharmaceuticals, artificial organs with build-in defibrillators, have covered the market. (J. Banks, 2018) says to enhance the accuracy in the medical field, the Da Vinci operating a robot is an extraordinary machine with the implementation of AI and machine learning algorithms as its base. AI has also helped in the implementation of the production of digital images and holograms with high resolution.

1.2 Technology review

Development of a personal remote assistant using AI and machine learning algorithms are common nowadays. According to the researchers, these projects were to find the advanced and new optimized processes for robot and human conversations, which focus primarily on the social aspects. As we all know that these technologies deal with human lives; therefore, there advanced technical solutions need more attention and certification to ensure safety for the common human (M. B. Hoeschl, T. C. D. Bueno, and H. C. Hoeschl, 2017). Safety is still a weakness in this in the health care industry even after the improvement and development of the robotic prototypes devoted to the patients. In this world of competition, it is extremely difficult to achieve a product that is safe and inexpensive at the same time.

According to Eriksson, there is some disinclination to use the technology in the area of health care as it involves the interaction between humans, for various reasons such as ethical, moral and needs for personal contact. Moreover, the advanced systems such as exoskeletons for people who cannot walk for themselves and electronic pets for patients with Alzheimer's will be in the market very soon. Additionally, as a human cannot keep up with the eruption of scientific literature, our new generation will see advanced algorithms that will search the huge databases and cloud services to find out the best practice and best solution for one particular treatment (J. Banks, 2018).

1.3 Ethical Concern

The important thing to focus on now is who is accountable for any decision made using the latest AI and machine learning technologies in the field of treatment. Moreover, we cannot rely on a technologically driven society. We need to get involved in the process of ethical

decision making, in addition to the advancement of technology. It will be wrong if we leave the new emerging ideas which are developing globally in the hands of scientists and engineers, rather focus on their ethical obligation we need more people in the field of nursing and social science (V. Vakkuri and P. Abrahamsson, 2018).

If monitored ethically, when it comes to the fact that robots are performing certain tasks, human beings get paranoid by the thought. Who will we blame or who will be actually responsible in case if something goes awry? We all know though the human is good at analyzing data and recommending a creative solution for health treatment robots could also help us for the same. Systems must be plotted in a way that the machines are not deciding on their own. Though human beings are more aware nowadays regarding the ethical issue, in spite, they will need the assistance of the machines and robots to perform advanced tasks and automate them (V. Vakkuri and P. Abrahamsson, 2018). To guarantee the ethical issue in technological development is addressed before time, several steps have already been taken. The role of the AI and the robots should be determined well in advanced in this health care profession. The critical part is that health care technology must be developed by keeping the human context in mind and not just the final agenda.

2. Autonomous Vehicles and AI

The amount of time and research efforts are put in autonomous vehicle technology to prevent causalities in our vehicle's post-implementation. The computerized vision on personal or public transport vehicles performs a driving function without human intrusion. The research to combine electronic and optic sensors for accuracy and a better understanding of the driving environment via machine learning (ML) algorithm and computer vision is enormous. An

inclining interest from big corporations to invest in autonomous vehicle research keeps increasing in number with a time to propose a technique or algorithm that provides insights on the driving environment (N. Hutchins, Z. Kirkendoll, and L. Hook, 2017). The most widely researched fields for better functioning of autonomous vehicles are not limited to but include the detection of a 3D object, lane, and edge. The like Tesla, Waymo, Ford is investing time and effort to develop algorithms for vehicle drive automation. The advancement in sensor technologies helps to gather data that include detecting and identifying signs with advanced deep learning techniques that are making the impossible come true with time (Li, B., Ouyang, W., Sheng, L., Zeng, X., & Wang, X., 2019).

2.1 Ethical and Moral Concerns

The concerns with artificial intelligence or autonomous systems, be it ethical or moral, need to be conveyed in an understanding term that's easy and crystal clear before the society could accept them. A debatable topic by philosophers and scientists on which can be the pure form of ethics or combination of ethics, that benefit AI/AS decision making with time. In our society, there is no constructive answer to this debatable topic that would stop the implementation or slow down the creations of those systems (Chen, X., Kundu, K., Zhang, Z., Ma, H., Fidler, S., & Urtasun, R., 2016).

In the absence of proper guidelines, corporations will create their own goals that serve a specific purpose. It need not necessarily mean that in this situation, harmful ethics will be designed to AI/AS. However, there are chances for dangerous ethics to exist, and no matter how minute the problem is, it may represent the society in which AI/AS exists. In today's world, natural human language does show any effect on the working of an autonomous

machine, so a simple and basic idea of ethics can't be easily fed and come to an assumption that autonomous systems will function by the command by humans.

We must create such an approach that stresses on logic and formal ethical rules to accurately interrupt when needed by a machine. Questions on quantitative results and finding solutions to a well-defined problem can be answered correctly by a scientific or engineering approach, whereas issues such as how to make the world a better place cannot be explained entirely by just science or engineering alone (Chen, X., Kundu, K., Zhang, Z., Ma, H., Fidler, S., & Urtasun, R., 2016). It is quite challenging to differentiate between an indirect and a direct consequence in AI/AS systems. For example, outcomes such as accurate time, place, events that happen in sequence might be required to be considered by AI/AS systems, which no human could interpret. As per human instinct, there should be a difference in ethical values between various consequences.

2.2 Tesla Crash Report



Figure 1: Tesla crash (medium.com, 2018)

Recent Tesla Model S crash with a fire truck which was parked on California freeway which was responding to a situation, we can analyze the scene where an autonomous driving car made an accident, and the reason can be many, be it a coding flaw, logical error in manipulations, image processing problem, etc. As the NTSB investigation, it concluded that the system detected and followed two different vehicles 15 seconds before the crash (extremetech.com, 2019). from the case, the lead vehicle was switching lane to the right and noticed fire truck very near 3-4 secs before impact. In research and development stages, these are generally referred to as “cut-out scenario.” The tesla’s autopilot system failed, and the car had accelerated to 30.9 mph before the impact. Tesla's autopilot system was able to detect the fire truck when the distance to contact was merely 40 to 45 feet, and a car traveling at 30mph can cover 44 feet of distance in one second (extremetech.com, 2019). As we know Such cases, where there are fundamental issues in the design, can be rectified but not at the cost of sacrificing lives and livelihood.

3. Robotic Process Automation

RPA is an art of scripting bots at the application level, which can automate tasks that are redundantly performed by millions of humans redundantly every day. The way RPA is performing in the market we can clearly see a boom in the trend that companies are adopting it in order to replace the routine and redundant tasks performed by humans in order to make the task time-efficient and cost-efficient and the boom in the robot market is rewriting the way humans and robots interact with (B. Kehoe, S. Patil, P. Abbeel, and K. Goldberg, 2015).

Tech giants are creating a nexus together with RPA and other cognitive technologies like NLP, Speech recognition, ML in order to automate decision and judgment related tasks that are traditionally performed by humans.

The symbiotic relation between RPA and these cognitive technologies is helping automation to dwell deeper to get into areas that had previously never experienced automation, which is making the experience efficient and agile as they are strategically becoming digital firms.

RPA technology automates redundant, repetitive and boring tasks that humans invariably perform every day, this is done by interconnecting different applications through Modern RPA tools and developing a script through which, it will perform the task just like humans do.

Potential advantages are many as they can be cost effective and time-saving, meanwhile being efficient, which can 24/7 without complaints and takeoffs. And we can scale the resources as and when necessary and scale down when the traffic or workload is low.

“Attended automation” is a trending automation technology under RPA where it could assist humans in front office works, a potential virtual call center agent who could take call summary notes and much more useful stuff (D.Schatsky, C.Muraskin, and K.Iyengar, 2017).

Royal Bank of Canada redesigned its banking operations deployed several hundred bots in its operations for a few banking backend processes, which handled almost two million operations per year, and as a result, it was efficient in cutting down the redundant workforce and becoming time-efficient boosting its revenue incredibly.

3.1 Fourth Industrial Revolution

In the past centuries, first world countries, which were first to revolutionize industrially, happened strategically switch their operations towards secondary and tertiary sectors and switch their labor and manufacturing operations to third world countries where they found cheap labor and raw materials. The phase of Industrial revolution 4.0 will also do the same with the current scenario; a robot operation will cost 90% less than a human worker per

hour, which makes it cheaper than work outsourced even from any third world country (M. B. Hoeschl, T. C. D. Bueno and H. C. Hoeschl, 2017).

This is causing the project origin countries to bring back the work to their home country

However, this does not happen to boom employment in the home country but rather still employs bots and automation dominating the human workforce.

Pew research center interviewed 1896 experts' opinions about advantages in AI and Robotics and their impact on jobs in society, 48% of them believed that the technology would displace the human workforce and cause dis-harmony, causing income inequality and disruption in the social fabric. While the other half of them predicted that AI would, in turn, create efficient jobs strategically shifting the workforce towards efficiency and more expertise jobs (M. B. Hoeschl, T. C. D. Bueno and H. C. Hoeschl, 2017). Despite fears since the emergence of this industrial revolution mattering loss of jobs by the strategic technological shift, progress regarding the same is yet to be settled. Clear examples state that fields, where digital technology has made enormous impact be it in any way, are yet to fully make the major changes that are impending in the near future.

4. Ethical Design Considerations

Deontological ethics helps us understand objectives considering rightness or wrongness of action that are more important than its consequences, more like the intentions are important than the end result. Ultimately intentions of the action set the standards for what ought to be good or bad. However, applying a non-consequentialist approach to certain consequential

matters is also difficult as they are based on consequences again (N. Hutchins, Z. Kirkendoll, and L. Hook, 2017).

A few actions are inherently considered bad unless it explicitly harms someone. This approach is too broad as of now to bring it in the sphere of AI/AS. For society, the important consideration for AI/AS should be a consequence of the perceived actions. A machine capable of behaving intelligently may still be doing the right things, but for society, all it matters is the end result, which could be unharmed to human life.

Utilitarian ethics, be it whatever it looks at, benefiting as many people as possible or reducing the loss as much as possible irrespective of actions and consequences.

This approach helps the AI/AS, to be a bit clear about ethics as they can be working towards reducing the damage as much as possible (N. Hutchins, Z. Kirkendoll, and L. Hook, 2017).

And once considering the statistical results, we can be clear that the system does not take into account the emotions in considering when making decisions, and their decision making is a purely logical and rational basis.

In the sphere of utilitarian agents, this approach can be practical; meanwhile, when these agents have to interact with several others in different scenarios, maybe with humans, this approach can be unfair.

Ethics has a few branches, of which normative ethics, applied ethics, and meta-ethics play a key role (M. B. Hoeschl, T. C. D. Bueno, and H. C. Hoeschl, 2017).

Normative ethics can be categorized into teleological, virtue and deontological ethics while teleological considers the consequences, utilitarian ethics which we discussed above is a popular form of teleological ethics, virtue considers character traits of individuals in order to

evaluate them and in deontological, focus is purely based on actions as stated in the starting of this section .

Meta-ethics considered of few things to deal with possibility, if at to have a moral knowledge then what this knowledge is and how can we make it possible While applied ethics looks after surgical domains like environmental issues, medical issues, etc.

With a proper structural setup, converging several ethical approaches to build and implement would be the best way to overcome at least several ethical problems related to the industry. Letting the users involved in the design of approach as which would best suit the society based on purchasing power that can meet the AI/AS considerations, laws, and rules that are set by the governing agency (M. B. Hoeschl, T. C. D. Bueno and H. C. Hoeschl, 2017).

Certain assumptions in ethics are necessary when building an AI/AS model. In that way, at least there is a ground/ opportunity to discuss the issue and principles together. Any approach for that reason depends on the ethical guidelines that dictate staying unbiased, being inclusive, consistent, fairness, etc. that build nexus between society and AI/AS.

5. Security, Privacy and Safety concerns with AI

AI/AS protection should be the prime interest. **Safety** has become the leader of the generation of technology and is also relevant to autonomous systems for a very good reason. AI/AS can be easily compromised or hacked. This may influence the general control safety of the systems and the misrepresentation of its ethical decision-making methods. If the ethics of an autonomous agent can be jeopardized, then that could be managed to cause harm to various people (Xiuquan Li and Tao Zhang, 2017). This problem will probably be a brainstormer among the manufacturers and the one who is ready to manipulate AI/AS.

The Solitude for AI/AS acts as a new sample of potential 4th act violations in terms of **Privacy**. The right of the people to be safe in their homes, papers, and results, against irrational searches and illnesses, shall not be disrupted. Mass production and execution of AI/AS will give enough possibilities for the entity's stress for the use of potential in a sustained state of being observed through data gathering and storage (Xiuquan Li and Tao Zhang, 2017). UAVs provide a good tool for authorized monitoring through the use of zoom lenses, thermal image, facial recognition, and performance profiling. Yet, those tools could readily undergo a transition to public monitoring for keeping track of individuals, groups, with potential for prying and unfair targeting and profiling.

Once mass approval of AI/AS has occurred, **Safety** plays a key role, society's civil rights and secrecy could easily and quickly be at stake, without ample assurance or overlooking. Protection for AI/AS is another area of concern for ethical strategies. By what metrics are AI/AS supposed to be safe, or safe just for public use? By what ethical guidelines will the society decide that an autonomous system has made an honestly good decision under definite conditions? If an employee is threatened by an autonomous system at his work, would the accountability be credited to a human superintendent, the company running the autonomous machines in the workplace, or the real operator? These are complicated situations that require new views on what is supposed to be the best ethical way. At present, most autonomous systems still need guidance by a human, but in the future, that wouldn't be the same (Xiuquan Li and Tao Zhang, 2017).

6. Conclusion

It is of no doubt that Artificial intelligence technology is dominating the human life, the growing use of AI/AS has a profound impact on human life, and the speed in which it is embracing the lives of people is remarkable. As this paper discusses the case studies, it is important for every stakeholder in this field to take the problems associated with it seriously. Especially the ethical concerns arising out of it. Investment of time and research into setting the guidelines for AI/AS, be it IEEE standards or the German government ethics guidelines for autonomous vehicles. It is certain that the future of AI/AS development will need more and more ethical rules and guidelines in order to make those tough decision making by machines.

References

- J. Banks, "The Human Touch: Practical and Ethical Implications of Putting AI and Robotics to Work for Patients," in *IEEE Pulse*, vol. 9, no. 3, pp. 15-18, May-June, 2018.
- N. Hutchins, Z. Kirkendoll and L. Hook, "Social impacts of ethical artificial intelligence and autonomous system design," 2017 IEEE International Systems Engineering Symposium (ISSE), Vienna, 2017, pp. 1-5.
- A. R. Honarvar and N. Ghasem-Aghaee, "An artificial neural network approach for creating an ethical artificial agent," 2009 IEEE International Symposium on Computational Intelligence in Robotics and Automation - (CIRA), Daejeon, 2009, pp. 290-295.
- V. Vakkuri and P. Abrahamsson, "The Key Concepts of Ethics of Artificial Intelligence," 2018 IEEE International Conference on Engineering, Technology, and Innovation (ICE/ITMC), Stuttgart, 2018, pp. 1-6.
- M. B. Hoeschl, T. C. D. Bueno and H. C. Hoeschl, "Fourth Industrial Revolution and the future of Engineering: Could Robots Replace Human Jobs? How Ethical Recommendations Can Help Engineers Rule on Artificial Intelligence," 2017 7th World Engineering Education Forum (WEEF), Kuala Lumpur, 2017, pp. 21-26.
- Xiuquan Li and Tao Zhang, "An exploration on artificial intelligence application: From security, privacy and ethic perspective," 2017 IEEE 2nd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), Chengdu, 2017, pp. 416-420.
- D.Schatsky, C.Muraskin, and K.Iyengar "Robotic process automation A path to the cognitive enterprise " Deloitte university press, 2017.
- B. Kehoe, S. Patil, P. Abbeel, and K. Goldberg, "A Survey of Research on Cloud Robotics and Automation," in *IEEE Transactions on Automation Science and Engineering*, vol. 12, no. 2, pp. 398-409, April 2015.
- Chen, X., Kundu, K., Zhang, Z., Ma, H., Fidler, S., & Urtasun, R. (2016). Monocular 3D Object Detection for Autonomous Driving. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). doi: 10.1109/cvpr.2016.236
- Li, B., Ouyang, W., Sheng, L., Zeng, X., & Wang, X. (2019). GS3D: An Efficient 3D Object Detection Framework for Autonomous Driving. IEEE Conference on Computer Vision and Pattern Recognition (CVPR).