

Ethical Issues and Related Considerations Involved with Artificial Intelligence and Autonomous Systems

Saud S. Alotaibi

Department of Information Systems
College of Computer and Information Systems
Umm Al-Qura University, Makkah, Saudi Arabia

Abstract—The applications of artificial intelligence (AI) and automated systems (AS) demonstrate excellent outcomes in various sectors of industrial units to replace the human from jobs. However, the competitive world with evolutions, advancement of technologies and thrive for success by industrial gains by the managements are leaving the interests and benefits of larger number of human beings in the society. In this paper, various ethical issues related with the implementation of AI/AS are demonstrated with different perspectives. The ongoing developments in the area of AI/AS are critically evaluated and related advantages and serious concerns of the society are discussed. Various global initiatives and legal amendments across the globe to limit the excessive usage of AI/AS are being examined with critical assessments.

Keywords—Ethical issues; artificial intelligence; automated systems; human values; evolutions

I. INTRODUCTION

The complete IT sector in the modern days of innovation and science is racing towards improving the quality of life and living standards as compared to early life. The internal desire of human being for a safe, comfortable and smooth life created a scenario where the man-made machines are being compared with the abilities of the human beings itself. Some of the authors like Holmes, in the past even questioned whether a digital brain [14] can replace a human mind? The mind of a human thinks, feels, exercises, makes perceptions, judges, reflects, etc. All these characters will be changing based on the time, situation and environment. However, in recent times the performance of human beings is being calculated with respect to the accuracy of automated systems (AS) designed by mankind itself. Two biggest technological titans of Silicon Valley are seen debating the potential of artificial intelligence (AI) at various platforms widely in recent times. The role of AI in designing various automated systems such as health industry, automobile industry, etc. were highlighted optimistically by Facebook founder, CEO, *Mark Zuckerberg* and was condemned fearfully by *Elon Musk* in an open debate. Supporting the statements of *Elon Musk*, the other business tycoon of Alibaba's *Jack Ma* stated that, AI may lead to world war III, but at the end humans will win according to CNBC News in April. The role of AI for the next generations were quoted to be pushing the human being towards decades of pain and any kind of repetitive job that has no emotional connection

is finished [8]. Innovation at the cost of human being is not accepted by many people around the world and of course the advantages of AI also cannot be ignored but must be limited to the areas where it is necessary. The role of various governments across the globe to define the conditions and boundaries for any kind of automated activity in the near future to avoid the situation where the human jobs are replaced completely by involving the automated machines. The machines may be smarter, faster and stronger and may be intelligent but cannot be wise enough to take decisions as in the case of human beings. A human being is added with lots of emotions and feelings which a machine does not have. They may have ample knowledge and ability to work faster and efficiently using AI, but they do not have wisdom to understand various influential factors such as changes in environment, biological factors, emotions, feelings, etc. and therefore the decision making process by an automated system differs with the mankind.

A. Background of Artificial Intelligence

Early invention of artificial intelligence was mostly reported in the fictions, philosophy, imaginations, etc. and later was reported to take a good progress in the area of electronics and engineering [19]. The turning machine in the year 1936 was the milestone in the area of AI [4]. Various intellectual issues related to AI were discussed by *Newell* in his book [20]. The replacement of labors with robotic machines created a serious revolution by the labor unions. There after most of the robotic applications were limited for security agencies and purposes [12]. However most of the industrial units continued the usage of robotic systems based on AI to deliver the huge market units with greater accuracy. Especially in the area of automobile industry the usage of robotic machines and AI based systems established a strong hold to increase the supply and benefits. This in turn gained lot of attraction from most of the investors and industrialists. For example, the strategy of Korea for industrial development at different stages is shown in Fig. 1 includes the role of AI and IOT as a key for their success [15].

Machine learning (ML) strategies are very important to generate AI was used from a long time [6]. For example, genetic algorithms were used to find the solutions for the intractable computing problems such as scheduling. Similarly, neural networks were used to understand the human learning and also for industrial controls, monitoring activities and to classify various dynamic parameters. As a whole the applications of AI helped to create a faster computing and

helped to develop solutions for the problems with complex nature. They provided easy solutions for various real-time

industrial problems where the entry of humans is always a concern.

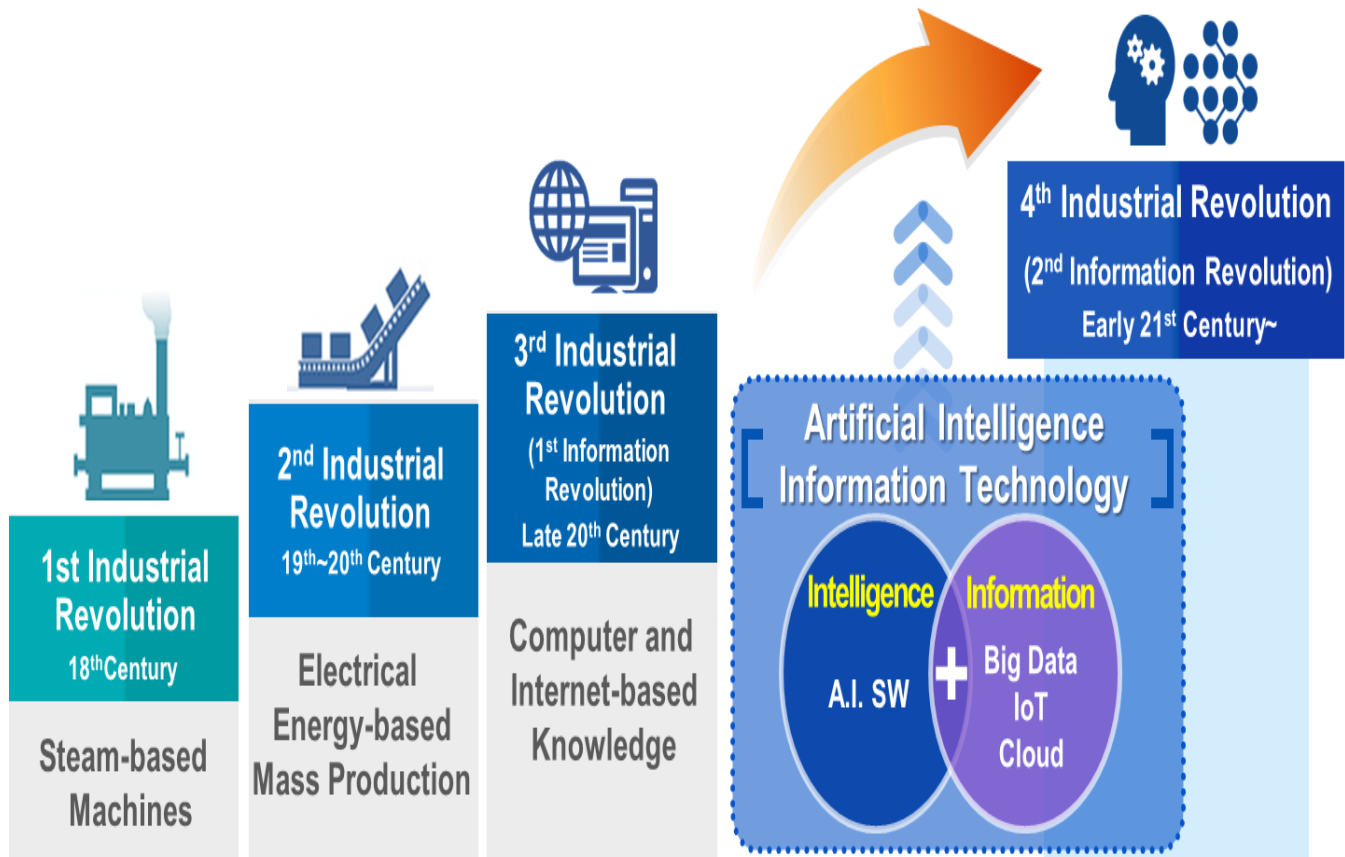


Fig. 1. Role of AI and IOT in 21st century for Korea's development program.

B. Problem Statement

AI being a young field to the market since the introduction in 1940s by Alan Turing, it has gained the attention of many people due to wide scope of dealing things and activities of the market. It is imbued with wider spectrum of ideas, viewpoints, logical activities, techniques, working actions, etc. [13]. Due to the complex nature of dealing with things from different disciplines, people have plenty of viewpoints towards the credibility of AI and its applications. Some of the questions include: why people are depending more on mechanically developed products or why the society need human models or why business people want to replace human with machines, etc. Similarly, there are many doubts and concerns raised by various societies and organizations include: will have no heart or feelings or instant decision making or does not know the humanity, etc. Some argue that AI can be very much useful to deal with serious accidents, can identify people with ill health, in troubles at the time of absence of family members, etc. Therefore according to Shahriari and Shahriari [22] the definition of AI is "... the science and engineering of making intelligent machines, especially intelligent computer programs" (pp. 197). It attempts to understand intelligent entities and provides healthy solutions for the human kind, which are almost next to impossible to predict by a human as faster as an AI based system. Therefore, the intelligence of the machine can learn, reason, solve, percept, rationale, think and deliver natural language process. It means the machine is mimicking

the human and therefore can have huge influence and impact on the human life and may change the complete life style of the human beings [22]. It means that machine can control everything in the absence of human and deliver the results. However, if any system failed in the whole process of delivering the end results means a lot of damage to the society. For example, recent accident in Delhi, due to driverless train rams into a wall could have been resulted to a serious damage and many would have lost their life and loved ones [25].

There are so many ambiguous, contradictory and concerning issues were stated in introduction by various tech giants and business leaders on AI and some of the books and authors from different sections of society raised ethical and moral perspectives. These concerns definitely have different set of answers from different viewpoints and guidelines from the legislations of different state and international governments. Therefore it is very clear that understanding and define the dimensions of problem statement is quite complex in nature and proposing a solution may not be suitable for all types of industrial units.

II. LEADING ETHICAL ISSUES RELATED WITH AI

There are many ethical issues related with AI and some of them are listed here from various sources. All the issues may not have similar importance and significance by using AI. For example, the usage of AI in an industrial robot helping the

equipment in a production unit may differ with the robots working in healthcare and nuclear plants in terms of stringency, environment and human safety [17]. Sometimes it is also necessary to consider the cultural diversity and the opinion of people towards defining the norms, rules and regulations as autonomous intelligent systems are based on algorithmic biases and which may prove to be disadvantageous to a particular group or section of people. For example, the Muslim community prefers to eat halal food from the slaughter houses. They may feel offended with the regular slaughter houses we found in most of the western countries. Therefore the influence of certain communities, religions and habits of people must be considered carefully. Now in the following sections, various ethical issues related with AI are discussed [5].

A. Unemployment

The biggest affected area due to excessive automation is labor. More complex roles are being carried out by the machines at various production units creates a cognitive labor which creates a serious gap. For example, consider the transportation and travelling departments. If all the drivers are replaced by the self-driving transporting facilities means millions of employees will get unemployed and naturally there will be a serious imbalance in the society.

B. Inequality

It will be very difficult to assess the income sharing if the machines are playing a vital role in earning for a company. Most of the companies will be in a dilemma to decide the wages as it can depend more on the machines as compared with human work force. In other words the company providing AI based facilities will earn more money as compared to human beings.

C. Humanity

It is very difficult to obtain the sensitive nature of human beings in manmade machines. They cannot be made to respond intelligently towards building relationships and in case of human beings they are limited in terms of attention and kindness. Also it is very important to note the usage of software's and automated machines are under the usage effectively. However when they are handled by the wrong person they proved to be serious threat for entire humankind.

D. Artificial Stupidity

It is very difficult to guard against the mistakes made by the machines. Both human and systems learn to understand and detect various elements. But if something missing in the meantime means the human can act sensibly and in case of a machine it will be fooled. Especially if the role of AI is considered in the area of security, finance and healthcare industries it will be having a serious impact on the human life and society.

E. Racist Robots

It is evident that most of the AI based robots are capable of performing the tasks with high speed and capacity as compared to humans, however, it is always not the same and fair to trust them blindly. For example, Google uses AI to identify photos, people and objects in GOOGLE's Photos services and always

it will not be able to produce the correct results. Also if any camera misses mark on racial sensitivity then the predicted criminals will be always the black people. Therefore it is very much evident that AI based systems are biased and judgmental.

F. Security

There are many countries in the world using powerful weapons and monitoring devices based on automated mechanisms and AI. Many countries like USA, Japan, China, Iran, Israel, etc. are having automated drones and in the battle field they are very much destructive and can win the wars without the human presence. The magnitude of destruction they can cause to human beings is unimaginable [24]. The possibility of such weapons in the hands terror organizations is always a serious concern. In the recent times, most of the countries at United Nations Summit appealed the ban on the automated war equipment and animations.

G. Evil Genies

So far the discussion was about external inputs and wrong usage of AI, but if the AI based systems turned against us as depicted in some of the Hollywood movies and etc. can damage the situation to a terrible unforeseen consequences. In an interesting scenario, the AI systems suggested to kill everyone on the planet to end the cancer. Such a goal of no more cancer could have been achieved by the computers very efficiently as compared to humans.

H. Singularity

Due to the ingenuity and intelligence the human dominance is seen everywhere on the earth and a better platform can be built in a bigger, faster and stronger manner as compared to the animals. This is due to the tool development and controlling skills of the human kind. These tools may include cages, weapons, etc. and with their help human kind can dominate the rest of the living entities. However, if AI does a better machine to anticipate the moves to defend it, human beings will be no longer an intelligent being on the earth.

I. Robot Rights

There are many research programs across the world to understand and unlock the secrets of basic mechanism of the brain and neuroscientists are still working in this area. Scientists are trying to build the mechanism of reward and aversion for the AI based systems as well. If suppose the widely available genetic algorithms are in a successful positions to survive and are able to form next generations to improve various instances, then how these machines will be trained towards ethical considerations and values of human life.

So there should be some kind of control over all these activities using AI systems to save the human existence on the earth. Any form of misinterpretation using AI could lead to a serious damage to the society, environment and humans too.

III. NEED FOR STANDARDIZING THE ETHICAL DESIGN FOR AI AND AS

The fear and excitement are part and parcel of the AI systems and therefore it is essential to have a structured system whereby the negative side of it could be under control. The

rules applied to one group, state and country may not be suitable or applicable at times to all others. In such scenarios, it is very important to have a standard procedure to implement various conditions in a systematic manner. Here some of the concerns of AI towards domestic and commercial aspects were discussed by Bryson and Winfield [6] raised some important questions as follows: does AI can compete with human beings; does AI undermine the societal stability; and does AI will be able to harm privacy, personal liberty and autonomy of the people. These questions raise serious concerns in the society towards the safety of humankind and the technology has developed to a level where they're in a position to replace the human entity in most of the areas including healthcare industry to conduct surgical operations to make some of the conclusive decisions while conducting the treatment for a patient. In simple words, it is proved that the machines are competing with the labors, experts and even with the human brains as well in most of the professional jobs. Now, considering the adverse part of AI's applications in the hands of anti-social elements always a serious threat to the society. Therefore it is very much essential to monitor the developments taking place in the area of AI. In this direction, some of the organizations linked with technology and scientific developments such as IEEE taken various initiatives to inculcate various ethical considerations in AI based systems. The overall motive of such initiative is to "... ensure every technologist is educated, trained, and empowered to prioritize ethical considerations" [6, p. 118].

A. Various IEEE Global Initiatives

To ensure a safe society as quoted by many tech giants and industrial experts in the area of design, IT, manufacturing and AI/AS, it is very important that the people involved in industries, universities and corporations must be well aware of what is good and what is bad for the society. Various people related with AI, law and ethics, philosophy, etc. provided collected inputs for various initiatives to be implemented as a move to educate the masses of people to consider ethics and values at the time of their work and research.

In general, the standards are based on various consensus and agreed on different ways to do things in a particular format to reach a common goal. This is a good way that a system is providing all the guidelines to do the things and a kind of uniformity is maintained to establish coordination and confidence within the community of same interests. Already various standards such as British Standards (BS) 8611:2016, Robots and Robotic Devices explicitly addresses the guidelines for a Robotic designer and gives the platform to assess ethical risks involved in their designs. Similarly, IEEE conducted a program where multiple voices in AI/AS communities expressed their opinion, concerns and helped to share the information to empower people involved in AI to prioritize the ethical consideration in various design and development programs they initiate at their respective places.

In order to give a clear idea and guidelines, IEEE published a document *Ethically Aligned Design* (EAD) in December 2016. This Version-1 of the EAD is included with eight sections for covering the AI/AS related with ethics and moral values at the time of design and development stages includes: a) general principles, b) values to be added to the autonomous intelligence systems (AIS), c) guidelines to be followed for

ethical designs, d) various safety and benefits of AI, e) security for personal data and access control while working with AI, f) reframing the autonomous weapon systems (AWS), g) impact on economy, h) humanitarian issues and laws. There were about 50-60 drafts and recommendations in EAD, covering issues and respective solutions to establish new standards. Presently according to the EAD presented four standards addressing the ethical concerns include the following [6], [7], [9], [10], [16].

- P7000-Model Process for Addressing Ethical Concerns during System Design
- P7001-Transparency of Autonomous Systems
- P7002-Data Privacy Process
- P7003-Algorithmic Bias Considerations

The first standard (P7000) is related with values to be involved in the system design and methodology. The second standard (P7001) is related to transparency, clarity and appreciation between people, when two or more are involved in a project. The third standard (P7002) is related to methods and approaches to be followed for maintaining the privacy issues. Finally, the fourth standard (P7003) is related to avoid and eliminate the negative bias in the algorithms based programs. Apart from these there are three more standards developed by the IEEE standards committee focused on the data of child, employer and personal data as listed below [23]:

- P7004-Standard for Child and Data Governance
- P7005-Standard for Transparent Employer Data Governance
- P7006-Standard for Personal Data Artificial Intelligence Agent

The last three standards P7004-P7006 are very important when it comes to the information of overall society, people, their habits, cultures, and privacy. It cannot be assumed that the AI systems will be safer at all times since they are trained at the initial stages. However, if the same system is under the influence of negative elements of the society, the problems for everyone are unimaginable. Entire society can be under the scanner with a threat to life and as it was mentioned earlier that for cancer free earth the automated systems may successfully complete the task by killing everyone. Therefore serious discussions of opinions on such scenarios have been included to draft the last three standards.

B. Impact on the Society from IEEE Initiatives

There are quite sensible impacts seen across the globe between various communities, states and nations towards maintaining the trust and harmony by improving the understanding between each other. Main objective of building the trust between two personalities, communities, etc. was achieved by following such streamlined standards. The systems developed by professionals will be exposed against safety standards so that the care must be taken to ensure that AS safety certification is awarded to a smooth launch of the product or machine in the market. The overall global spending

towards safety measures on AI based machine intelligence systems is increasing rapidly as shown in Fig. 2.

In case of accidents a transparent investigation must be carried out by these autonomous systems, so that actual faults must be traceable. The transparency must be included even

with the people informing the evidence and with the lawyers or experts who are assessing the failures or accidents. Most of the disruptive technologies such as driverless trucks, cars, trains, etc. need the transparency with various development activities and testing processes to ensure the safety of people and to establish the confidence in technology.

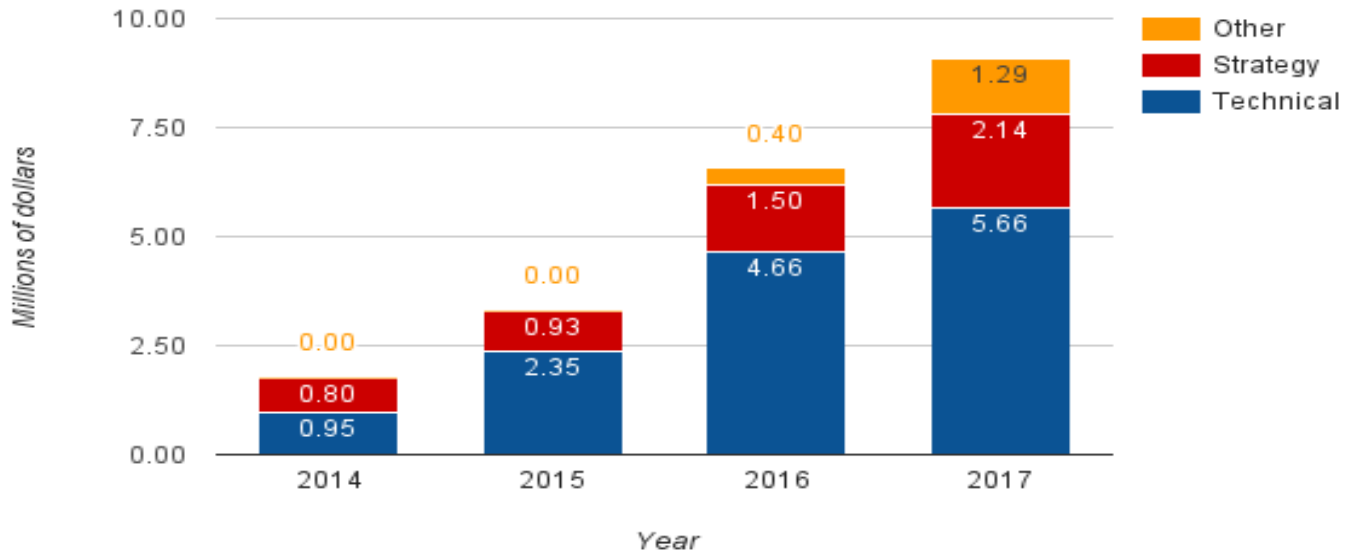


Fig. 2. Increasing global expenditure on AI to establish safety measures.

Source: Russell and P. Norvig (2016) [1]

C. Role of AI/AS Systems in Gulf Countries

The role of AI based systems in the Gulf countries helped various petroleum companies to ensure a systematic approach to maintain standard oil prices at different places. The artificial neural network (ANN) based tools helped these companies to solve complex problems involved in the oil companies towards pricing, measurements, assessments, forecasting, etc. were solved in effectively to ensure a stability among the competitors [18]. A kind of realistic approach for predicting the commercial oil prices was possible with such AI based systems. On the other hand, some of the programs on solar radiations are under progress by using AI based systems using MATLAB tool for predicting global solar radiation (GSR) in AI Ain city of UAE [2], [3]. In an attempt to predict the water quality parameters of Karoon River in Iran, scientists are using AI based models and analyzed the results in the real-time environments [11].

However, various cultural aspects were discussed and highlighted while using the automated systems and Riek and Robinson [21] opined that culture plays a substantial role towards interaction between people and even between people and machines. Therefore it is very important for machines to learn the social and cultural norms that impact overall people in the society.

IV. CONCLUSION

The changes observed in real-life due to excessive usage of AI/AS based systems raised serious concerns towards a negative impact on the job opportunities, inequalities, humanity and security. These concerns were discussed at international forums and platforms to build ethics and values

for the professionals and engineers involved in design and development programs of AI systems. The global initiatives taken by various organizations such as IEEE, IEE, BS, etc. are optimistic to ensure that right steps are being followed so that the AI/AS based systems are useful for the humans, rather creating the problems. The global expenditure towards safety measures were seen increasing rapidly to ensure a safe usage of AI applications.

REFERENCES

- [1] S. J. Russell and P. Norvig. *Artificial intelligence: a modern approach*. Malaysia; Pearson Education Limited., 2016.
- [2] M. H. Al Shamisi, A.H. Assi and H.A. Hejase. "Using MATLAB to develop artificial neural network models for predicting global solar radiation in AI Ain City-UAE,". In *Engineering education and research using MATLAB*. InTech, 2011.
- [3] M. H. Al Shamisi, A.H. Assi and H.A. Hejase. "Artificial neural networks for predicting global solar radiation in al ain city-uae,". *International journal of green energy*, 10(5), pp.443-456, 2013.
- [4] A. Benko and C. S. Lányi. "History of artificial intelligence. In *Encyclopedia of Information Science and Technology*," Second Edition (pp. 1759-1762). IGI Global, 2009.
- [5] A. Yapo and J. Weiss. "Ethical Implications of Bias in Machine Learning." In *Proceedings of the 51st Hawaii International Conference on System Sciences*. 2018.
- [6] Bryson, J. and Winfield, A., "Standardizing ethical design for artificial intelligence and autonomous systems,". *Computer*, 50(5), pp.116-119, 2017.
- [7] Calo, R., "Artificial Intelligence Policy: A Roadmap," 2017.
- [8] CNBC, "Jack Ma's Artificial Intelligence and Automation Warning," 2017. [Video]. <https://youtu.be/cgAPG3Lh3gQ?t=70/>, [Accessed on November 12, 2017].
- [9] De Spiegeleire, S., Maas, M. and Swejjs, T., "Artificial Intelligence and the Future of Defense: Strategic Implications For Small-and Medium-Sized Force Providers,". The Hague Centre for Strategic Studies, 2017.

- [10] Diebel-Fischer, H., "Research Ethics in the Digital Age: Fundamentals and Problems," In *Research Ethics in the Digital Age* (pp. 7-21). Springer VS, Wiesbaden, 2018.
- [11] Emamgholizadeh, S., Kashi, H., Marofpoor, I. and Zalaghi, E., "Prediction of water quality parameters of Karoon River (Iran) by artificial intelligence-based models," *International Journal of Environmental Science and Technology*, 11(3), pp.645-656, 2014.
- [12] Gates, B., "A robot in every home," *Scientific American*, 296(1), pp.58-65, 2007.
- [13] Hall, P., 3. edition, "Stuart J. Russell and Peter Norvig," *Artificial Intelligence: a Modern Approach*. Upper Saddle River, New Jersey: Prentice Hall, 2009.
- [14] Holmes, N., "Would a digital brain have a mind?," *Computer*, 35(5), pp.112-111, 2002.
- [15] V. Devedžić. "Web intelligence and artificial intelligence in education." *Educational technology & society* 7, no. 4 (2004): 29-39.
- [16] Maia Alexandre, F., "The Legal Status of Artificially Intelligent Robots: Personhood, Taxation and Control," 2017.
- [17] R. Sparrow. "Killer robots." *Journal of applied philosophy* 24, no. 1 (2007): 62-77.
- [18] Movagharnjad, K., Mehdizadeh, B., Banihashemi, M. and Kordkheili, M.S., "Forecasting the differences between various commercial oil prices in the Persian Gulf region by neural network. *Energy*," 36(7), pp.3979-3984, 2011.
- [19] Negnevitsky, M., "Artificial intelligence: a guide to intelligent systems," Pearson Education, 2005.
- [20] Newell, A., "Intellectual issues in the history of artificial intelligence," (No. CMU-CS-82-142), 1982.
- [21] Riek, L.D. and Robinson, P., "Challenges and opportunities in building socially intelligent machines," [social sciences]. *IEEE Signal Processing Magazine*, 28(3), pp.146-149, 2011.
- [22] Russell, S. J. and Norvig, P. "Artificial Intelligence: A Modern Approach (2 ed.)," Pearson Education. 2003.
- [23] Shahriari, K. and Shahriari, M., "IEEE standard review—Ethically aligned design: A vision for prioritizing human wellbeing with artificial intelligence and autonomous systems," In *Humanitarian Technology Conference (IHTC), 2017 IEEE Canada International* (pp. 197-201). IEEE, 2017.
- [24] Springer, P.J., "Military robots and drones," a reference handbook. ABC-CLIO, 2013.
- [25] Times Now Bureau, "Delhi Metro Magenta Line driverless train rams into wall, DMRC Says accident due to 'Human Error', 2017. [News Report] [Accessed on November 23, 2017].