Yadav, Rohit Kumar (yada6101@vandals.uidaho.edu)
University of Idaho

## Abstract

AI (Artificial Intelligence) The theory and development of computer systems able to perform tasks that normally require human intelligence such as Visual perception, speech recognition, decision making and translation between languages. As technologies are emerging enormously in today's fast running technology world and Artificial Intelligence(AI) and robots are one among them which is outperforming everywhere whether its comes to helping our daily lives or in industries. so far the greatest danger of Artificial Intelligence is that people conclude too early that they understand it. The field of AI has a reputation for making huge promises and then failing to deliver on them. This paper generally discusses about the positive and negative factor in global risk in which most people having misconception about the artificial intelligence and Robotics which makes erronous beliefs about AI and robotics as benevolent and malevolent.

## Literature Review

Yudkowsky, E. (2008) who mainly focus the positive and negative factors of Artificial Intelligence in global risks. The field of AI has a reputation for making huge promises and then failing to deliver on them. Most of the observers and researchers concluded that AI is hard and I think It is hard indeed. The critical inference is not that AI is hard, but that, for some reason, it is very easy for people to think they know far more about Artificial Intelligence than they actually do. In Eliezer's one of his research he opened by remarkinig that "few people would deliberately choose to destroy the world; a scenario in which the Earth is destroyed by mistake is therefore very worrisome." (Eliezer). Few people would push a button that they clearly knew would cause a global catastrophe. But if people are liable to know and confidently believe that button does something quite different from its actual consequence that would be an actual alarm.

Eliezer also compared his research on AI with cognitive biases; "Cognitive biases are settled science. Artificial intelligence is not settled science; It belongs to the frontier, not to the textbook." (Eliezer) and due to this reason he conducted the research and discussed on the topic of global catastrophic risks.

# ARTIFICIAL INTELLIGENCE

Yadav, Rohit Kumar (yada6101@vandals.uidaho.edu)
University of Idaho

This paper proposed a friendly AI study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. (McCarthy et al 1955) think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer. (McCarthy et al. 1955)

Mostly paper discussed about the AI prediction and its design and depicted it as anthropomorphic bias which means having an attribution of human traits and emotions in non human entities like robots. It says that It would be very hard to say anything and predict exactly about the AI and its design where people easily believes through anthropomorphism that they can make predictions which makes an errornous beliefs on AI design to be friendly or enemy. It discussed about the technical and philosophical failures which happen by the peoples erroneous belief about the AI that they did not completely understand AI which leads to the failure where when you try to build something which failed to understand the working of your own code It best described through the examples of neural network of overtraining. Philosophical error is trying to build the wrong thing which means even if you build it successfully you will fail the benefit of humanity. It discussed generally with the examlpes of communism where people are so keen towards and want the robots that they forgets the catastrophe can be done by AI and robots. It also discussed about the threats and promises and the intelligence enhancement of AI vs Human.

Hislop D.,Coombs C., Stanimira T, and Barnard S. (December 2017) summarizes the findings of a rapid evidence review that was conducted between October and December 2016 into the impact of various emerging technologies (artificial intelligence, robotics, and automation technologies) on knowledge and service work, relevant professions, and society, The review focuses on academic literature (peer-reviewed journal papers and conference papers) that had been published since 2011" (Hislop et al).

This paper reviewed on various topics and outlined various questions related to the impact of AI and robots on work for example what are the work related outcomes and mediators from the utilization of AI and robots and It covered both the impact of workers and organization, also what are the ethical issues related to the contemporary utilization of AI, robotics and automation technologies.

Yadav, Rohit Kumar (yada6101@vandals.uidaho.edu)
University of Idaho

Nitto, Hiroyuki, Daisuke Taniyama, and Hitomi Inagaki (February 1, 2017) paper's data shows of the three countries "the US has the highest level of robot utilization in home and in retail stores whereas germany have large number of users used for industrial purpose. Japanese consumers generally have strong perspective and positive attitudes towards robots however survey findings revealed that they still lack sufficient of knowledge about robots and therefore think the rise of robots society far in the future, causing them not to carefully consider the correlation between human and robots " (Nitto et al). this paper outlined that as a emerging science and technologies we should accept the potential of robots and needs to understand the intelligence and proper use of robots in industrial fields, which may change our way of living and daily life. This paper outlined the general customer survey on autonomous driving cars and smartwatches and also highlighted the percentage of people expecting self driving cars to hit the roads by 2025 also It discussed about the difference in attitudes towards and acceptance of robots in Japan, US and germany.

Pineau, Joelle, Michael Montemerlo, Martha Pollack, Nicholas Roy, and Sebastian Thrun ( March 2003) describes about the mobile robotic assistant which is developed to give assistant to elderly individuals with mild cognitive and physical impairments as well as their support nurses in their daily activities. The authors of this paper presented three modules relevant to ensure successful human robot interactions: an automated reminder system; a people tracking and detection system and finally a high level robot controller that" performs under uncertainity by incorporating some knowledge from low-level modules, and selecting appropriate courses of actions" (Pineau et al).

It discussed about the hardware and software structure and architecture autominder and their plan management. The system has been tested successfully by authors on autonomously providing reminder guidance to elderly residents. The experiments were successful in two main dimensions. First, they provided some evidence towards the feasibility of using autonomous mobile robots as assistants to nurses and institutionalized elderly. This was demonstrated in part by the robot's ability to complete the assigned task, but also by the fact that the response from the elderly participants was uniformly positive. Second, this project also demonstrated the effectiveness of probabilistic tracking and decision making for interactive robots. "Pearl is one of a few robots to use POMDPs, and the first to apply POMDP planning to the highest level of decision making. The ability to represent the uncertainty inherent in a person's behavior, and formulate plans

Yadav, Rohit Kumar (yada6101@vandals.uidaho.edu)
University of Idaho

accordingly, allowed the robot to robustly handle difficult situations, including noisy communication and crowded environments." (Pineau et al)

Parsons, H. Mcllvaine, and Greg P. Kearsley(October 1981) described the human factors community to the field of robotics including current and future applications and needed areas of research. It also explained and described why human factor engineering is important to the field of robotics. It discussed relationships between the human and robots. Generally these discussion just dwell on the organizational impact of robots. It discussed how the person machine interfaces and interactions in these be arranged so the joint activities of robots with people is as effective as possible. The objective of this paper is recognizable as the essence of human factor engineering though for some it may be unclear by more explosive issues. Human factors has been very less involved in few years till date in the field of robotics and some aspects is described in this article. There have been a number of reasons for this lack of attention. "Most robot applications at present and in the near future are in the domain of manufacturing, an area in which few human factor's scientists have worked extensively. Further, much of the basic research in robotics has come from the field of ,artificial intelligence, another domain in which few human factors researchers have been significantly involved. Perhaps the most important reason is that until recently there were only a handful of actual robot installations to present problems or issues to study" (Parsons et al)

"Due to drooping productivity in many nations during the past, few years, interest has increased greatly in industrial automation, with resulting boom in robot applications. Advances in the electronic industry especially LSI circuits are leading to robots that are compact, powerful and affordable" (Parsons et al). It described how the performances of robots and people should be interrelated and tried to explain that we must have to understand the capabilities and limitations of each, For that It examined the human factors engineering issues related investigations and potential for human factors engineering applications and research in robotics.

Yadav, Rohit Kumar (yada6101@vandals.uidaho.edu)
University of Idaho

**Conclusion:**

All scientific ignorance is hallowed by ancientness. Each and every absence of knowledge dates back to the dawn of human curiosity; and the hole lasts through the ages, seemingly eternal, right up until someone fills it. Yudkowsky, E. concluded that It is possible for mere fallible humans to succeed on the challenge of building Friendly AI. But only if intelligence ceases to be a sacred mystery to us. Intelligence must cease to be any kind of mystery whatever, sacred or not. We must execute the creation of Artificial Intelligence as the exact application of an exact art. And maybe then we can win. Paper concluded on the research with introducing friendly AI and its possible for mere fallible humans to succeed on the challenge of building Friendly AI. But only if intelligence ceases to be a sacred mystery to us, as life was a sacred mystery to Lord Kelvin. Intelligence must cease to be any kind of mystery whatever, sacred or not. We must execute the creation of Artificial Intelligence as the exact application of an exact art. (Yudkowsky, E. 2008)

Yadav, Rohit Kumar (yada6101@vandals.uidaho.edu)
University of Idaho

**Reference:**

Hislop D.,Coombs C., Stanimira T, and Barnard S. (December 2017). "Impact of Artificial Intelligence, Robotics and Automation Technologies on Work." CIPD https://www.cipd.co.uk/Images/impact-of-artificial-intelligence-robotics-and-automation-technologies-on-work_2017-rapid-eveidence-review_tcm18-35319.pdf.

Nitto, Hiroyuki, Daisuke Taniyama, and Hitomi Inagaki (February 1, 2017). "Social Acceptance and Impact of Robots and Artificial Intelligence." *NRI Papers*. https://www.nri.com/~/media/PDF/global/opinion/papers/2017/np2017211.pdf.

Parsons, H. Mcllvaine, and Greg P. Kearsley(October 1981). "Human Factors and Robotics: Current Status and Future Prospects." *Human Resources Research Organization*,. https://files.eric.ed.gov/fulltext/ED206911.pdf.

Pineau, Joelle, Michael Montemerlo, Martha Pollack, Nicholas Roy, and Sebastian Thrun ( March 2003) "Towards Robotic Assistants in Nursing Homes: Challenges and Results." *Robotics and Autonomous Systems* 42, no. 3–4 271–81. https://doi.org/10.1016/S0921-8890(02)00381-0.

Yudkowsky, E. (2008). "Artificial Intelligecnce as a Positive and Negative Factor in Global Risk." Edited by Nick Bostrom and Milan M. Cirkovic. *Global Catastrophic Risks*, n.d. https://intelligence.org/files/AIPosNegFactor.pdf.