CalsMUN 2020

Historical Influences

Research Report

Forum: **Human Rights Council**

Issue: **Ethics and rights in regards to artificial intelligence**

Chairs: **Sal van der Schaar and Daksh Khanna**

Personal Introduction

Daksh Khanna

Dear delegates,

My name is Daksh Khanna and I will be co-chairing with Sal at CALSMUN! I live in the Hague and I attend Gymnasium Haganum. I like debating and that's also why I love Model United Nations. CALSMUN will be my tenth MUN, but also my first time chairing in a Human rights committee. In my free time I play the piano, I love to read (I'm a nerd ;)), I learn languages, and I enjoy hanging out with friends. I am very excited to meet you all at the MUN and I am sure we are going to have a very, very fruitful debate!

Sal van der schaar

Hi! I'm Sal van der Schaar, I am 16 years old and I will be serving as your chair this weekend together with Daksh. I go to the Stedelijk Gymnasium Haarlem, in Haarlem, where we organize HMUN, although I live in Zandvoort. CALSMUN 2020 will be my tenth MUN conference in total, of which this will be my second chairing experience. In my free time I enjoy playing tennis and hockey, although I won't have much of that the upcoming week(s). I am looking forward to meeting you all!!

Introduction

Often as kids, we hear on TV that robots can take over the world. Most of us neglect this, but it can mostly be true one day. Many scientists around the world fear that without any ethical frameworks and laws on robot ethics, they could become smarter themselves and one day, be smarter than humans. For instance, there was a study in 2009 in Switzerland where robots were programmed to work together. They were programmed to search for a beneficial recourse and not a poisonous one. They learned to lie to each other in order to stock on the beneficial recourse for themselves.

When it comes down to robot morality, in other words, ethics or roboethics, we also encounter the problem of who should be held accountable for specific actions taken by robots. For instance, on March 19, 2018, a woman called Elaine Herzberg in Arizona, who was jaywalking, got killed by a self-driving Uber Car. Whom can we blame for the accident? The pedestrian, the driver, Uber, the government or the car company? In this case, the car was in the driver, but the control was in the hands of the car. However, the car was not programmed to stop these kinds of accidents (jaywalking pedestrians), so it was not prepared to see people on the street and stop the car. These are some issues regarding robot ethics.

The Committee

The Human Rights Council is an inter-governmental body within the United Nations system made up of 47 States responsible for the promotion and protection of all human rights around the globe.

The body can discuss all thematic human rights issues and situations that require its attention throughout the year.

Definition of Key Terms

Robot ethics

The term 'robot ethics', points to the way robots are designed, constructed, used and treated by humans.

Artificial Intelligence Arms Race

An artificial intelligence arms race is a competition between two or more states to have its military forces equipped with the best "artificial intelligence" (AI)

General Overview

A scientist called Joseph Weizenbaum explained in 1976 that Artificial intelligence should never replace humans in positions which require respect such as; customer service representatives, therapists, nursemaids for the elderly, soldiers, judges and police officers. The scientist explains that people who have those positions, especially therapists, need to have empathy for people in order to help the clients the best they can. If AI replaces the jobs above, clients/customers will probably not have a safe feeling around the AI, as the 'empathy' from robots will not feel real to us. The scientist argues that customers/clients will feel frustrated instead of safe. Moreover, lastly, the scientist says that the use of artificial intelligence for positions mentioned before is a threat to human dignity.

At the same time, Andreas Marcus Kaplan, a professor at ESCP business school, points out that Artificial intelligence is only as smart as one makes them. He describes them as machines where someone can put information into, in other words, programming, and that they will perform the actions they were programmed to do. With this in mind, he also shows that using AI as judges can be very problematic. When we use AI in court rulings, the AI will base its ruling from previous rulings and profiles created of specific groups of people which biases can form.

Some academics and experts attended a conference, which the Association for the Advancement of Artificial intelligence (AAAI) organized, to discuss the potential dangers of AI and talking about what would happen if AI would become self-sufficient and able to make their own choices. The AAAI showed that some machines acquired some level of autonomy by, for instance, finding power sources by themselves and being able to decide who their enemies are, and eliminating them. With this information, they pointed out that it will be unlikely that AI will become self-aware, which film producers often portray in science fiction, but that the decision making by robots could be dangerous.

Major Parties Involved

Organisations

Association for the Advancement of Artificial Intelligence (AAAI)
The AAAI is an international scientific society dedicated to promoting research and responsible use of artificial intelligence. AAAI also aims to increase public understanding of artificial intelligence, improve the knowledge of AI practitioners by training and teaching, and guide research planners and funders concerning the importance and potential of current AI developments/threats and future directions.

Centre for the study of existential risk

The Centre for the Study of Existential Risk (CSER) is an interdisciplinary research centre within the University of Cambridge dedicated to the study and mitigation of existential risks.

Future of Humanity Institute

FHI is a multidisciplinary research institute at the University of Oxford. Academics at FHI bring the tools of mathematics, philosophy and social sciences to bear on big-picture questions about humanity and its prospects. Founding Director Professor Nick Bostrom leads the Institute.

Countries

Saudi Arabia
In October 2017, Saudi Arabia granted a robot called Sophia 'citizenship' in the country, though some people saw this more as a publicity stunt, rather than an action to honour the rise of AI, while other saw it as a denigrating human right.

China
The state council plans to spend around $2.1 Billion on artificial intelligence and plans to become the global leader of artificial intelligence by the year 2030.

United States of America

The USA shares the same goal as China when it comes to wanting to become the global leader of artificial intelligence. They have spent more than China on AI, around $10 billion.

Timeline of Key Events

Date	Description of Event
1942	Isaac Asimov writes three laws of robotics. First law: A robot may not injure a human being or, through inaction, allow a human being to come to harm. Second law: A robot must obey the orders given it by human beings except where such orders would conflict with the First Law. Third law: A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.
1950	Alan Turning proposes the imitation game. This is an exercise in which a human "interrogator" is challenged to make the difference between the text-only responses of a machine and a human being. It helped give shape to a philosophy of artificial intelligence.
1956	Dartmouth holds an AI conference. Dartmouth invited many of the top advanced science researchers of the day to Dartmouth Hall for the event in summer 1956. The scientists discussed numerous potential areas of AI study, including learning and search, vision, reasoning, language and cognition, gaming (particularly chess), and human interactions with intelligent machines such as personal robots. The general idea from the discussions was that AI had great potential to benefit human beings. The scientists yielded a general framework of research areas where machine intelligence could have an impact. The conference organized and energized AI as a research discipline for years to come.
1970's	At this time, AI has only been in the research stage. Government agencies like the U.S. Defence Advanced Research Project Agencies, pumped money into AI and did not pay attention into getting anything in return, while AI researchers kept exaggerating about AI so that they could keep their funding. However, when two reports got published by the U.S. and the UK, describing the little potential of AI, different agencies around the world demanded that AI researchers should provide detailed programs in return for funding.

Previous Attempts to Resolve the Issue

There have been no previous attempts to resolve this issue.

Possible Solutions

The biggest question we need to ask in this situation is, do AI deserve laws of their own, and until what level should we make robots and computers self-sufficient/able to learn more through themselves. Also, smaller questions like, why does China and the US want total control over AI and for what purposes will they use AI, in comparison to other nations? Furthermore, should AI makers/researchers be transparent in the work that they do and to what extent should they be transparent? For instance, they could share entire codes online, but for someone who does not know what those codes mean, it is not that useful.

What we could work on are temporary legal frameworks regarding AI, as their level of self-sufficiency/general abilities will shift in various speeds throughout the century. As well as trust and knowledge from humans towards artificial intelligence, especially in lower developed nations. Lastly, we could also work on the work that AI researches do and how transparent they have to be.

Bibliography and Further Reading

Further reading (watching)

1. https://www.youtube.com/watch?v=WSKi8HfcxEk
2. https://www.youtube.com/watch?v=DHyUYg8X31c&t=

Bibliography

1. Evolving Robots Learn To Lie To Each Other, Popular Science, August 18, 2009, http://www.popsci.com/scitech/article/2009-08/evolving-robots-learn-lie-hide-resources-each-other
2. Artificial intelligence arms race - Wikipedia. https://en.wikipedia.org/wiki/Artificial_intelligence_arms_race
3. Who is responsible when a self-driving car has an accident?". *Futurism*. Retrieved 2019-07-26. https://futurism.com/who-responsible-when-self-driving-car-accident
4. Wakabayashi, Daisuke. "Self-Driving Uber Car Kills Pedestrian in Arizona, Where Robots Roam". *The New York Times*. March 22, 2018. https://www.nytimes.com/2018/03/19/technology/uber-driverless-fatality.html
5. General overview on the opinion of Joseph Weizenbaum - Joseph Weizenbaum, quoted in McCorduck 2004, pp. 356, 374–376
6. Kaplan Andreas; Michael Haenlein (2018) Siri, Siri in my Hand, who's the Fairest in the Land? On the Interpretations, Illustrations and Implications of Artificial Intelligence, Business Horizons, 62(1)". November 27, 2018. https://web.archive.org/web/20181121191205/
7. Scientists Worry Machines May Outsmart Man By JOHN MARKOFF, NY Times, July 26, 2009. https://www.nytimes.com/2009/07/26/science/26robot.html?_r=1&ref=todayspaper

8. OHCHR | HRC Home. https://www.ohchr.org/EN/HRBodies/HRC/Pages/Home.aspx
9. Association for the Advancement of Artificial Intelligence. https://aaai.org/home.php
10. About us - The Centre for the Study of Existential Risk. https://www.cser.ac.uk/about-us/
11. Future of Humanity Institute. https://www.fhi.ox.ac.uk/