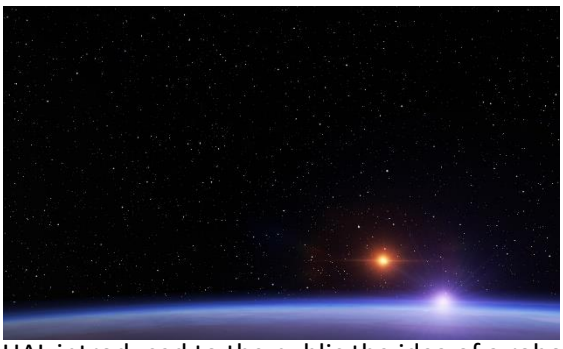
Ethics and robotics



In the classic film *2001: A space Odyssey,* the supercomputer HAL 9000 faces a moral dilemma: to reveal the true nature of the spaceship's mission that he was tasked to keep a secret or lie to the crew. This seemingly kind computer, full of human emotions decided to kill the crew to protect the mission's secret.

HAL introduced to the public the idea of a robot that is not built just for factory work, but that makes decisions, understands natural language, and has a vision.

Robots are becoming more and more independent and sophisticated, and the concept of AI facing ethical dilemmas is moving from science fiction into the real world.

Robots nowadays serve a variety of roles. They optimise logistics, translate texts, compose art, conduct research, make predications, detect fraud, and much more. It's indisputable that as AI becomes more capable, our lives become easier, more efficient, and safer. For instance, driverless cars are considered safer than ordinary cars, the way autopilots have made airplanes safer.

Robot ethics, or roboethics or machine ethics, is a growing interdisciplinary research effort that deals with the pressing ethical questions about developing and implementing robotics in the society covering many areas from medical care to military robots, from search and rescue missions to entertainment robots. Initially, military robots where the main focus of the roboethics discussion. A continuous debate has been going on about the development of military AI and autonomous weapons and whether and when autonomous robots should be allowed to use lethal weapons and whether they should be allowed to make those decisions.

In 2015, 25000 AI and robotics researchers wrote an open letter requesting the ban of autonomous weapons without human supervision in order to avoid an international military AI arms race. In the recent years, other types of robots and AI have been included in the debates.

When talking about roboethics we might mean 3 different things:

- The ethics of the systems of the robots and the ethical problems that may occur, whether for example they pose a threat to humans, in the short or long run, which uses of robots are problematic and unethical (for example in healthcare or in war)
- The ethics of the people who develop and use the technology and how they design them in order to be "ethical".
- The ethics of how people treat robots.

The discussions about the morality of robots are not new. In 1942, the science fiction writer Isaac Asimov, in his short story "Runaround", developed the Three Laws of Robotics to serve as a moral guide for robots. According to his laws, when robots face conflict, they must:

- Never harm a human being, or, through inaction, allow a human beings to be harmed.
- Obey the instructions given by humans, except where such orders conflict with the First Law.
- Protect their own existence as long as they are not violating the First and Second Laws.

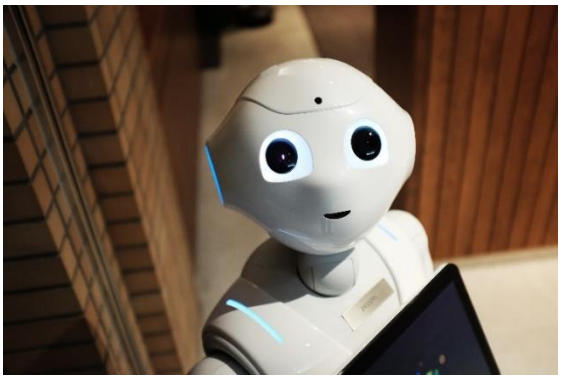
80 years later, Asimov's laws offer little help to the development of a robotic code. The writer established his founding principles as a counter to the Frankenstein story on the basis that semi intelligent human creations are jealous of the human nature and will inevitably turn against their creators. At the time, people were imagining robots with arms and lasers, not as autonomous vacuum cleaners or social media advertising algorithms. We now have a very different view of what robots can look like and how we interact with them and Asimov's laws don't apply to that.

However, although it is not very likely that Apple's Face ID will murder you, an unethical programmer could use the technology to invade your privacy using your iPhone's camera. Therefore, a set of rules for the design and use of the technology that prioritise predicting harm, reducing risks, and ensuring safety is important. Apart from military robots usually the ethical questions of roboethics don't have to do with killer robots. So, *don't kill humans* is not a good enough rule.

We should prevent robots from doing harm to people but also prevent people from avoiding responsibility for their actions. Laws are needed to determine whether the designer, the engineer, the manufacturer or the operator are responsible for a system malfunction or in case a robot harms humans. In other words, creators and operators must be held morally accountable for what is brought out into the world and how it's used.

Facial recognition is an example of a system that may not recognise the facial characteristics of people of colour, or those of women because their data is based on a limited set of data.

Such biased AI algorithms can cause problems such as missed flights, tense police encounters, or false arrests, but could we accuse the machine of discrimination? Bias in these algorithms can be reduced by using more inclusive training data. Companies and engineers in this case have the responsibility to examine their creations for bias in the way they check for software bugs.



So, could robots be moral agents? Could they be trustworthy partners, caretakers, educators, and members of the human communities?

At the moment robots are far from being fully independent moral agents. However, as they develop greater capabilities and ethical sophistication, they might become more autonomous bringing different ethical and policy challenges.

Each new technological advance pose new troubling questions. AI is no different. Mankind shall answer the questions in order to fully enjoy the benefits of the new technology.

References

Asimov, I., Warrick, P. S., & Greenberg, M. H. (1984). Machines that think: the best science fiction stories about robots and computers. London: Allen Lane.

Bossmann, J. (n.d.). Top 9 ethical issues in artificial intelligence. Retrieved from https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence/

Bradbury, D. (2017, July 7). Why, Robot? Understanding AI ethics. Retrieved from https://www.theregister.co.uk/2017/07/04/ai_ethics_and_what_next/

Fourtané, S. (2019, September 27). Roboethics: The Human Ethics Applied to Robots. Retrieved from https://interestingengineering.com/roboethics-the-human-ethics-applied-to-robots

Greene. (2018, February 23). Are Asimov's Laws of Robotics still good enough in 2018? Retrieved from https://thenextweb.com/artificial-intelligence/2018/02/23/are-asimovs-laws-of-robotics-still-good-enough-in-2018/

Malle, B. F. (2015). Integrating robot ethics and machine morality: the study and design of moral competence in robots. Ethics and Information Technology, 18(4), 243–256. doi: 10.1007/s10676-015-9367-8

Morals and the machine. (2012, June 2). Retrieved from https://www.economist.com/node/21556234

Open Letter on Autonomous Weapons. (n.d.). Retrieved from https://futureoflife.org/open-letter-autonomous-weapons

Robot ethics. (2020, February 8). Retrieved from https://en.wikipedia.org/wiki/Robot_ethics

Rouse, M. (2016, November 28). What is roboethics (robot ethics)? - Definition from WhatIs.com. Retrieved from https://whatis.techtarget.com/definition/roboethics-robot-ethics

Scope. (n.d.). Retrieved from https://www.ieee-ras.org/robot-ethics