AI Ethics – Backgrounder

The ethical implications of technology are vitally important and often overlooked. Artificial Intelligence (AI) is at the forefront of the Fourth Industrial Revolution, so there is an urgent need for AI ethics

Related topics: Artificial Intelligence, Internet of Things, Robotic Process Automation, Data Anonymization, Virtual Personal Assistant, Gamification

INTRODUCTION

Artificial Intelligence (AI) is already common in our everyday lives. With every Google search, social media login, tap on the smartphone or view of a weather forecast, we are using AI to communicate and make decisions. We are also contributing data to improve these AI systems that give us insights and make recommendations. As the use of AI grows, overlooking its ethical implications can have serious consequences. Factors like bias, transparency, security and privacy show the need for responsible use of AI.

This backgrounder explains AI ethics, why it should be a concern and which ethical issues may require discussion and governance.

WHAT IS AI ETHICS?

AI ethics is typically defined as 1) the moral behaviour of humans as they design, construct, use and interact with artificially intelligent beings/agents,[1] 2) The moral evaluations of the outcomes of AI-based systems.

Some issues on AI ethics are of immediate significance, others are only speculative. This document explores the following AI issues: effect on security and privacy, workforce automation, transparency, unintended bias and socio-economic impact.

[1] THE ETHICS OF ARTIFICIAL INTELLIGENCE. Nick Bostrom, Eliezer Yudkowsky.
Link: https://nickbostrom.com/ethics/artificial-intelligence.pdf

WHY IS AI ETHICS AN ISSUE?

With the widespread use of AI in building software, human influence such as bias is becoming evident in the results these systems produce. Research from the University of Washington[2] found that a Google Images search for "C.E.O." in the United States produced a result comprising 11 percent women, even though 27 percent of their chief executives are women. This is just one example of how AI can produce biased results that may skew the views of the general population. The rapid growth of this emerging technology has created new ethical challenges as more companies are using it (see figure 1) to build solutions

Biased AI is one of the key ethical issues. Humans design the algorithms and train AI agents with data that is primarily created by human to human or human to system interactions. As a result, AI systems inherit people's statistical biases.[3]

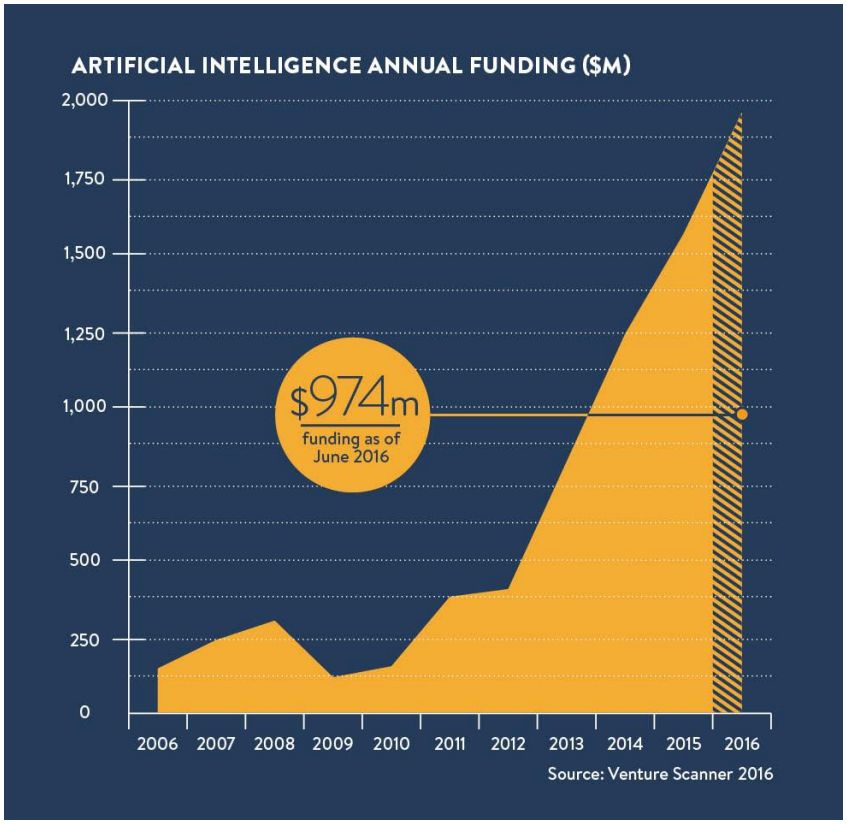

Figure 1: Annual global funding on AI systems (2006-2016)

[2] When Algorithms Discriminate. The New York Times. Link: https://www.nytimes.com/2015/07/10/upshot/when-algorithms-discriminate.html

[3] A review of possible effects of cognitive biases on interpretation of rule-based machine learning models. Link: https://arxiv.org/abs/1804.02969

AI also brings socio-economic impact. Mass AI automation may reduce time-consuming, repetitive labour and free up more time from people's day to day activities because they will need to work less. However, it may also increase the unemployment rate if fewer people are required to complete the same tasks.

Artificial stupidity, a term coined by 'The Economist'[4], refers to deliberately dumbing down computer programs to introduce errors in their responses so that they become more human-like. While it has already proven beneficial for some uses (e.g. chatbots), it has also introduced concerns about deception if we fail to distinguish the AI from a real person.

Personality traits such as compassion and empathy as well as human behaviour and social interactions have also been impacted. Widespread reliance on AI-based search engines, like Google, and social media websites, like Facebook and Twitter, has drastically changed the way we interact with each other both socially and professionally.

WHICH ETHICAL ISSUES SHOULD WE ADDRESS?

Perhaps the most important issues are security and privacy. It is becoming increasingly difficult to understand how AI works due to its unparalleled speed and ability to find sophisticated patterns. How do we safeguard sensitive information from malicious, third-party AI attacks when we cannot predict what will happen? The research is still in its infancy[5].

Another important issue is the evaluation of automated decision making. Tasks like recruitment and hiring, fraud detection, credit risk scoring, health risk prediction and future sales recommendations[6] are some of the avenues benefiting from AI's decision-making capabilities. However, with benefits come risks. There is concern about how much influence human bias has on AI's decision-making capabilities.

Workforce automation is a common practice. Which tasks should be automated and how much automation is good for an organization? AI automation may yield better performance and efficiency for an organization. It may enable people to enjoy a better lifestyle with more time for personal interests. However, it may also reduce job opportunities.

[4] Artificial stupidity. The Economist. Link: https://www.highbeam.com/doc/1G1-12504583.html

[5] The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation. Link: https://arxiv.org/ftp/arxiv/papers/1802/1802.07228.pdf

[6] See examples: Applications of artificial intelligence. Wikipedia. Link: https://en.wikipedia.org/wiki/Applications_of_artificial_intelligence

As AI spreads throughout the social and commercial world, new work roles will open up and demand will increase for people with the necessary skills.[7] The focus will shift to training the existing workforce to develop and maintain AI systems ethically.

The transparency of neural networks, which are the driving force behind today's AI, is a pressing concern. According to Geoffrey Hinton, who is considered the father of Deep Learning, humans currently have very little understanding of how a deep neural network recognizes sophisticated patterns to make decisions.[8] Therefore, it may be difficult to justify the implementation of such a system to process public data.

There are many other ethical issues that have recently come to light such as facial recognition tools for government or police use, weaponization of AI using drones, superintelligence, robot rights and singularity. Until AI becomes a mature technology, it will be difficult to address these issues.

CONCLUSION

Every disruptive technology requires discussion of its ethical implications. While AI enables systems to make complex decisions quickly and effectively, there is a global consensus that comprehensive policies must be in place to govern the ethical design of these systems. The leading technology companies have created 'The Partnership on AI'[9] to study and formulate best practices on AI technologies. Many governments around the world, including Canada, are also creating policies and standards to harness machine intelligence responsibly.

[7] AI will create new jobs but skills must shift, say tech giants. Link: https://techcrunch.com/2018/02/28/ai-will-create-new-jobs-but-skills-must-shift-say-tech-giants

[8] Geoffrey Hinton - The Neural Network Revolution. Link: https://www.youtube.com/watch?v=1JO1Pcr5rYA

[9] The Partnership on AI. Link: https://www.partnershiponai.org/