Source: *The Business Times,* p16
Date: 3 August 2021

when it comes to ethics, machines cannot correct and take charge to make better and more ethical decisions than humans can. Just because AI is called 'intelligent' does not mean it can be more ethical than humans.
IMAGE: PIXABAY

Both moral and digital upskilling equally vital

AI cannot differentiate between right and wrong and can never be responsible for the decisions that it takes. BY DAVID DE CREMER AND GARRY KASPAROV

WORLDWIDE, there's an emphasis on gaining digital skills. Data science and artificial intelligence (AI) courses are gaining in popularity. There are even coding courses for preschoolers. We are getting ready for a world that increasingly uses AI in our work and lives.

The digital upskilling initiative is a good one and needed. We feel, however, that the almost obsessive focus on acquiring digital skills has led to a situation where we may fail to understand the need for moral upskilling too.

As AI gains importance in decision making in various aspects of life, there is greater scrutiny on whether AI is "ethical". Will the AI system that helps to filter job resumes be biased against a minority race, because previous position holders were not from that race? Will facial recognition technology lead to the wrongful arrest of some people? Will the algorithm recommend certain beauty products to the consumer because they bring in higher profits for the company?

Because AI gets involved in decision-making, a consensus has grown that machines need to be taught ethics so their decisions will be rendered ethically. It is our view, however, that in this discussion on AI ethics, we have been misled if we think that AI can develop its own moral compass and subsequently can choose to be ethical.

Why is it that we think that AI can act and reason ethically on its own? The Big Tech industry is known to use a narrative that emphasises the idea that technology can be used to solve most problems that we encounter in society and business – referred to as a "techno-solution" mindset. As a result, this typical Silicon Valley mindset has penetrated governments and businesses that ethical dilemmas can also be solved if one has the right technology.

SOLVING ISSUES

For example, in the 2018 Congressional hearings in the United States, Facebook CEO Mark Zuckerberg's response to most of the lawmakers' questions was that AI can be used to solve issues ranging from hate speech and discriminatory ads to fake accounts and terrorism content.

Because of this "techno-solution" mindset, we have come to see ethics almost as being synonymous with transparency and intelligibility, which, interestingly are exactly features that can easily be optimised by modifying technological features with self-learning algorithmic solutions. Take, for example, Google's ethics-as-a-service message, which conveys to business leaders the idea that algorithms revealing unethical decisions can be fixed by working on specific technology features.

This has led to the mindset that we can expect no less from AI – that it can differentiate between right and wrong. And because of this ability we reason that AI therefore is also responsible for the decisions that it takes. For example, a recent report from the UN Security Council revealed that an autonomous drone attacked people, without receiving the specific order, in Libya last year. When the news broke, the image of "killer robots" was conjured in the minds of many. This idea underscores the idea (and fear) that we believe that AI is capable of making decisions in autonomous ways and thus is the one in charge to act in either good or bad ways.

However, in our view, this kind of logic is tantamount to saying that a gun fired itself after a person pulled the trigger. It shows that we clearly have forgotten that the drone was initially designed by humans to launch attacks. Relevant information that was keyed into the drone system was included by human beings. The ethical choice to design and adopt the use of these drones lies in the hands of the human, not the algorithm. Therefore, because AI did not decide to intentionally commit a bad deed, but simply acted upon decision-making rules coded by humans, it cannot be labelled as a bad machine that we can blame.

> When the usage of AI increases, it is the moral compass that we humans are relying on to guide us in making decisions. Without one, both machines and humans would be at a loss.

So, when it comes to ethics, machines cannot correct and take charge to make better and more ethical decisions than humans can. The reason is simple: AI acts as a mirror to our biases. It reflects bias when the humans show bias. If datasets include human bias, machine learning will act upon these biases – as the drone had learned – and even optimise those biases in its actions. Just because AI is called "intelligent" does not mean it can be more ethical than humans.

A recent illustration of how biased data leads AI to act in unethical ways was the decision to employ AI to predict the results of A-level students in the United Kingdom based on the historic performance of individual secondary schools. The result, however, was that many students' results were downgraded, particularly those from poorer schools. In an ironic twist, the use of AI, meant to reduce teachers' bias in predicting the students' results, thus created a new bias and revealed outcomes that we regard as unethical.

All of this means that we need to stop thinking that we will be able to trivially design machines that are more ethical than we are in the same way a programmer can create a chess program that is far better at chess than they are. It is not from machines that we can expect more responsible behaviour, but from the choices that people make with respect to intelligent technologies.

ETHICAL BUSINESS DILEMMAS

For this reason, we believe that as managers seek technological improvements that make data more easily interpretable, they should also be trained to be more aware of, and able to deal with, ethical business dilemmas.

Where AI reveals unethical outcomes, managers should be trained to recognise what human bias was underlying the machine decisions. This way, AI that amplifies our own biases can be used as a learning tool that can help managers recognise blindspots within their own organisation.

The benefit of promoting awareness of the company's ethical challenges and learning together with AI the potential biases underlying organisational decisions is that it will lead to a more ethically-aware company while at the same time enhancing its ability to use technology in more responsible ways.

When the usage of AI increases, it is the moral compass that we humans are relying on to guide us in making decisions. Without one, both machines and humans would be at a loss.

David De Cremer is director and founder of the Centre on AI Technology for Humankind. He is also a provost chair and professor of management and organisation at the National University of Singapore Business School. He is the author of 'Leadership by Algorithm: Who leads and who follows in the AI era?' and editor of a new book 'On The Emergence And Understanding of Asian Global Leadership'. Garry Kasparov is chairman of the Human Rights Foundation and founder of the Renew Democracy Initiative. His famous matches against the IBM super-computer Deep Blue in 1996 and 1997 were key to bringing AI, and chess, into the mainstream. His latest book on AI and the future of human-plus-machine is 'Deep thinking: Where machine intelligence ends and human creativity begins'.
The opinions expressed are the writers' and do not represent the views and opinions of NUS.