Artificial Intelligence (AI) and Its Associated Ethical Issues

Shahino Mah Abdullah*

The adoption of AI in many industries has been regarded by some as a threat to low- and middle-skilled workers, as it will drastically cut down reliance on the human workforce. Besides unemployment, there are also concerns about rising economic inequality caused by AI-driven companies. With fewer employees, these companies could gain a disproportionate advantage over conventional companies that still depend on normal, shift-based systems. There is also the issue that some AI bots have achieved the capability to interact with humans and build relationships through conversations. This influential communication could eventually enable these bots to affect human behaviour and possibly trigger certain actions. Significantly, therefore, such intelligent machines are not immune to mistakes and confusion since not all possible examples of real world interaction are covered during their training; this weakness could be manipulated to fulfil certain ends. Also, AI cannot be expected to be entirely fair and neutral, since it is dependent on human programmers, who have their own interests and whims.

Recent developments have revealed AI's ability to recreate images, imitate voices and even generate fake videos of a person, all of which could be potentially used for fraud. At this level, AI can make a phone call and impersonate a human voice to the point that even voice recognition software can no longer identify such fraud. Also, the ability of AI to collect huge amounts of data without consent and put humans under surveillance could be a threat to human privacy. Even in cases where AI could potentially eradicate disease or disasters, its approach and method could cause undesirable consequences that humans may not approve of. AI raises great security concerns, especially in the form of autonomous cyber-attacks in international conflicts and war. Even though AI is currently under human control, the growth of AI's ingenuity is exponential, with AI experts fearing that it might surpass human intelligence and potentially lead to a technological singularity, which then becomes a threat to the whole of humanity.

Therefore, the problem before us now is how to program AI with morals. Previously, the 'Laws of Robotics' were introduced in 1942 by Isaac Asimov, the modern father of robotics, to outline the interactions between humans, intelligent robots, and morality. However, these have since proven insufficient as there are more tricky questions surrounding what counts as ethical behaviour. It will be difficult to use these laws to answer questions such as 'Should a self-driving car

just hit another person to save the driver?' 'Is it permissible to program military robots to kill terrorists?' or 'Should a trading bot inform everyone of a predicted market disaster, or just keep it a secret and earn money from the prediction?' For that reason, holistic ethical codes must be developed to protect humans from harm. An AI framework that can take into account human desires, goals, preferences and ethical codes is indispensable. Within this context, one of the leading companies in AI research, Google DeepMind, expressed an interest in studying the real-world impact of AI and help technologists put ethics into practice. With the launch of its new 'ethics and society' research team in October 2017, Google DeepMind's co-founder and chief product officer, Mustafa Suleyman, stated that he hopes technology will address the world's challenges, i.e. help to address the effects of inequality, injustice and bias.

Collaborative work between Google DeepMind and OpenAI, a non-profit AI research company founded by Elon Musk to develop artificial general intelligence that is safe and beneficial to humanity, has outlined a new method of machine learning: deep reinforcement learning (RL). This takes its cue from human preferences based on a reward system. With this method, the actual goal of AI is to maximise total self reward, but while entailing human coaches and feedback. As this method relies heavily on the quality of its reward function, it could prevent AI from thinking on its own, which could lead to undesirable consequences. Currently, the AI framework that adopts RL is being extensively studied and developed by researchers to ensure that artificial agents act ethically. Even though this method requires a lot of human feedback, making it less efficient, it could direct AI towards satisfying our preferences while also acting more ethically when achieving that goal. To further increase its efficiency, an 'ethics shaping' framework has been proposed. Utilising a 'reward shaping' method, this would allow ethical decision making to be made independent of RL goals, as long as they are rooted in ethical human behaviour. Work is still in progress to improve this 'technology of ethics'.

To ensure that the rise of AI in all industries brings advantage to the people, we need to subject it to an ethical code that does not compromise humankind's potential or survival. Recently, the UK government has published a report entitled "AI in the UK: ready, willing and able?", commissioned by the House of Lords AI Select Committee, in which five core principles (the AI code) are outlined: 1) that AI should be developed for the common good and benefit of humanity; 2) that AI operates within the parameters of intelligibility and fairness; 3) that AI should not be used to lessen the privacy or data rights of individuals, families or communities; 4) that all people should have the right to be educated and to flourish alongside AI; and 5) that the use of AI as killer robots with the autonomy to hurt, destroy or deceive be opposed. The UK government believes it is ready

to take advantage of AI based on the resources it has and anticipates taking an important lead role in AI's ethical development.

Ultimately, however, the global community can develop a shared set of ethical principles when forming AI ethical codes. The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, for instance, has launched a crowd-sourced global treatise entitled, 'Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems'. This report, which has been translated into many languages, aims to align AI systems with moral values and ethical principles that prioritise human well-being. It provides recommendations for academics, technologists, and policymakers to utilise immediately. The report takes into account various well-established 'classical ethics', including religious and culturally-based ethical systems. However, there are ethical principles from some religions that are not included in the report, perhaps due to the absence of viewpoints representing those particular religious perspectives amongst the report's authorship. But, as the global community continues to work together on the ethics of AI, vast opportunities remain to offer additional ethical input from a variety of perspectives, including based on Islamic teachings.

In Islam, ethics is referred to as *akhlaq* (virtuous character traits). In the Arabic language, *akhlaq* is sometimes employed interchangeably with *adab*, which means manner, attitude, behaviour, and the etiquette of putting things in their proper place. Islamic ethics covers all the legal concepts, ranging from shariah (Islamic law), *fiqh* (jurisprudence), *qanun* (ordinance), and *'urf* (customary practices). Adopting and applying moral values based on Islamic ethical concepts or applied Islamic ethics could be a way to address various issues in today's society. At the same time, this approach is in line with the higher objectives of shariah (*maqasid al-shariah*), which aim to conserve human benefit by the protection of human faith (*hifz al-din*), life (*hifz al-nafs*), lineage (*hifz al-nasl*), intellect (*hifz al-'aql*), and property (*hifz al-mal*). This approach will be very helpful when addressing contemporary issues like AI. As there are so many ethical questions arising from AI, we desperately need to discuss and address the relevent issues using various perspectives from different disciplines. It is hoped that, through this approach, both scientific experts and religious scholars will work together to address issues in harmony and end their current working practice (in the context of ethics) of operating in two separate universes.

Notes

*   *Shahino Mah Abdullah* is Research Fellow, International Institute of Advanced Islamic Studies (IAIS) Malaysia.