

Is AI a danger to humanity?

Pranesh Prakash

2018-03-02

The Hindu

Abstract

We need to debate what AI ethics and regulation should look like

It's complicated

We need to debate what AI ethics and regulation should look like

We are told that AI is working magic, and also that it may lead to humankind's ultimate destruction.

Strong and weak AI

While we are far from “strong AI” (the idea of ‘thinking’ machines) we already have “weak AI” all around us — from translation Apps to facial recognition on social networks. But for most marketers AI has just become a buzzword for any form of algorithmic decision-making or usage of big data combined with self-improvement. Weak AI builds on mathematical techniques that have been developed since the 1940s, but have only more recently been computationally feasible.

Apart from computational power, AI requires copious amounts of data to learn. This data can either be generated by the machine itself — imagine a machine being instructed in the basic rules of chess and what constitutes “success”, and then playing millions of games against itself and using that as the basic data for improving itself — or it has to be provided data. If the data being provided have not been cleaned (whether in terms of accuracy or bias), then the resultant learning will also exhibit the flaws in the data.

By using AI to create closed captions on videos on YouTube, Google is helping all persons with hearing impairment (but currently in a restricted number of languages); by using AI for real-time image recognition, visually impaired persons are provided a chance to have the world in front of them narrated to them. And it is not just in rational “thinking” that AI can aid humankind, but also by performing emotional labour (as movies like Her highlight). These beneficial uses of AI cannot be denied. Despite the beneficial uses of AI, scientists and

leading thinkers like Stephen Hawking, Nick Bostrom, and Elon Musk warn us about the dangers of AI and the coming technological singularity.

Ethics and regulation

While it may sound trite, the greatest promise of AI is that of beneficial change at a faster rate than ever before, and accelerating. The greatest challenge of AI is the same, except with harmful change. While technological capabilities — and with it human capabilities to use technology — are changing at a faster pace than ever before, our ability to arrive at ethical norms regarding uses of AI and our ability to regulate them in an intelligent and beneficial manner have not nearly kept pace, and are not likely to. That is why we need AI researchers to actively involve ethicists in their work. Some of the world's largest companies are cornering the market for AI researchers with backgrounds in mathematics and computation: Baidu, Google, Alibaba, Facebook, Tencent, Amazon, Microsoft, Intel. They also need to employ ethicists.

Additionally, regulators across the world need to be working closely with these academics and citizens' groups to put brakes on both the harmful uses and effects of AI. Some parts of this will involve laws regulating data which fuel AI, some will involve empowering consumers and citizens vis-a-vis corporations and governments which are using AI, and some other parts will involve bans on certain kinds of uses of AI. While some of the most difficult legal and ethical questions around AI — involving liability for independent decisions made by AI — might not be questions we need to answer as of now, given that we are still far from strong AI, we still have difficult questions to be asked about harms caused by AI, everything from joblessness to discrimination when AI is used to make decisions. But for governments to regulate, we need to have clear theories of harms and trade-offs, and that is where researchers really need to make their mark felt: by engaging in public discourse and debate on what AI ethics and regulation should look like. And we need to do this urgently.

Pranesh Prakash is policy director of the Centre for Internet and Society, Bengaluru