

Ethical machines

WRONG RIGHT



\$400 – amount paid by Google for Deep Mind

2018 – date Eric Schmidt believes the Turing test will be passed

Ethical machines

Automation spreads beyond trading and managing systemic risk. As we approach technology singularity, autonomous robots and smarter algorithms make ethical judgments that impact life or death.

In ten years time, driverless cars will fill our roads, machine-learning algorithms combat disease and drones will deliver our shopping. Rapid advances in machine learning, visual and voice recognition and neural network processing mean that computers are getting better at perception tasks. This puts Artificial Intelligence (AI), once the mainstay of science fiction writing, at the forefront of next generation computing.

AI brings extraordinary benefits, particularly around disease diagnosis, while it also does a lot of boring but useful stuff, like recommending a specific book for online shoppers. Despite this, many thoughtful people, Stephen Hawkins, Bill Gates and President Obama amongst them, are concerned about the impact that deep learning computing will have, not only from a social and economic perspective, but also on the future of humanity itself. Hawkins, for example, warns of the “technological catastrophe” that could follow if artificial intelligence vastly exceeds that of its human creators.

The Internet of Things will generate exponential amounts of data and intelligent machines, as they crunch their way through it all, will increase their accumulation of knowledge exponentially. Algorithms are now designed to learn from raw perceptual data, understand language and recognise images, meaning computers build on the knowledge they amass, learn more skills, understand nuance and ultimately gain what we term common sense. As they become more adept, they can also self-improve, building better versions of themselves without human involvement. The impact of this is not lost on Google, which paid USD 400 million for Deep Mind, a UK-based AI start up. Facebook and Amazon are also making huge investments in this area.

In the short term, concerns about AI surpassing its creators are low but what seems certain is that it will soon take over some of the repetitive tasks that until now have formed the basic activity of many traditional professions such as accountants, lawyers, pharmacists and doctors. For a long time now computers have been better at analysing complicated data more effectively than people. Supermarket and factory workers have already found this to their cost. These days computers can read handwritten notes, write and translate reports and even respond to conversations. Despite initial installation costs, they have the added benefit of not getting tired or fed up and are unlikely to demand a pay rise - small surprise that they are gradually replacing their less flexible, human colleagues. Some argue AI is not replacing people, rather it is augmenting their abilities and in so doing making them more effective. Certainly more efficient computers will make some firms much more productive, but most likely at the expense of human capital.

AI is not replacing people, rather it is augmenting their abilities.

Data revolution



AI's transformational role is already impacting the automation industry. "Swarm intelligence," the collective behaviour of decentralized, self-organized systems, is currently being used to improve safety, when one car's brake sensors register icy conditions, for example, the information is shared with others through the cloud. Cars are therefore becoming more intelligent and the next generation vehicles will have thousands more sensor-connected computers to build on this. When all this is taken collectively, cars will be able to monitor themselves, their environment and even keep a weather eye out for passengers.

Widening this out across the transport sector, mesh networks and ubiquitous mobile connectivity will soon be able to offer totally automated highways that will improve safety, increase road capacity and reduce congestion. Pretty soon driverless cars will become an accepted norm. We are almost there; Google's autonomous vehicles have already covered 2 million miles and with only 14 accidents in that time, have an impressive safety record. Safer and more efficient roads will in turn change how risk is managed and shared, as insurance shifts from the individual and their car to whole fleets and, ultimately, the entire system.

Unfortunately however, accidents do and will continue to happen, particularly in built-up areas. Autonomous vehicles will therefore have to make some difficult ethical decisions and there are questions around how they will do this. Getting it wrong has huge legal implications that may well vary across jurisdictions. In Germany, for example, it is illegal to weigh up the value of one life against another, making it almost impossible to programme human life to a value that could be processed by an algorithm – whereas the US is not so rigorous in this regard. Should rules governing autonomous vehicles emphasize the greater good, the number of lives saved, without putting value on the individuals involved? At the moment there seems to be more questions than answers.

Once created it will be impossible to pull the plug on AI weapons.

Perhaps what is most concerning is the use of AI in warfare. Remote drones are already distancing fighters from the fighting; soldier robots and autonomous weapons are perhaps the next step. In the next decade or so algorithmic intelligence may well have the potential to surpass that of its human creators, identifying who to kill and why; the implications of this are frightening. Some argue that AI might be a better judge than say, extremists. As with the Internet, once created it will be impossible to pull the plug on AI weapons. It is also impossible to foresee where the development of such autonomous weapons will end.

The United Nations now convenes regular meetings to discuss this and the matter is so concerning that over 1,000 AI experts have already called for the development to stop. This is unlikely to happen. Perhaps the best we can hope for is a postponement while more consideration is given to regulation and constraint. One proposal is to ensure that the first generally intelligent AI is "Friendly AI" which will be able to control subsequently developed AIs. It seems rather fanciful today but perhaps in ten years this will be normal.

Historically, technology ethics has been mainly about the responsible and irresponsible use of technology by human beings. In the future, deeper consideration will be given to the behaviour of machines towards human users and other machines. Trust in the system will increasingly drive success, so organisations will seek to make data ethics a focus. In the short term, regulation is needed to determine whether the designer, the programmer, the manufacturer or the operator is at fault if something goes wrong – whether that is a car on a motorway or robot in a warzone.

It all seems a bit sci-fi. But science fiction is often the precursor of reality. What will happen when machines become smarter and more adaptable than their creators? Perhaps we should tread carefully.

Deeper consideration will be given to the behaviour of machines towards human users and other machines.

Related insights

Autonomous transport



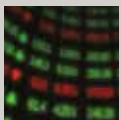
The shift to fully autonomous transport is an evolution via truck platoons on highways and small urban delivery pods. Connected cars create the network and test the technologies for the eventual revolutionary driverless experience.

Everything connected



Over 1 trillion sensors are connected to multiple networks: everything that can benefit from a connection has one. We deliver 10,000x more data 100x more effectively but are concerned about the security of the information that flows.

Dynamic pricing



The algorithms of Amazon and Uber cross over to affect more businesses, from energy use to parking. Real-time transparency allows better purchasing at the same time as margins and yields are automatically enhanced.