Artificial Intelligence should benefit society, not create threats

January 16 2015, by Toby Walsh

Science fiction has plenty of tales of AI turning against society including the popular Terminator movie franchise, here depicted on brick wall art. Credit: Flickr/Garry Knight, CC BY-SA

Some of the biggest players in Artificial Intelligence (AI) have joined together calling for any research to focus on the benefits we can reap from AI "while avoiding potential pitfalls". Research into AI continues to seek out new ways to develop technologies that can take on tasks currently performed by humans, but it's not without criticisms and

concerns.

I am not sure the famous British theoretical physicist [Stephen Hawking](#) does irony but it was somewhat ironic that he recently welcomed the arrival of the smarter predictive computer software that controls his speech by warning us that:

The development of full artificial intelligence could spell the end of the human race.

Of course, Hawking is not alone in this view. The serial entrepreneur and technologist Elon Musk also warned last year that:

[…] we should be very careful about artificial intelligence. If I had to guess at what our biggest existential threat is, it's probably that.

Both address an issue that taps into deep, psychological fears that have haunted mankind for centuries. What happens if our creations eventually cause our own downfall? This fear is expressed in stories like Mary Shelley's Frankenstein.

An open letter for AI

In response to such concerns, an [open letter](#) has just been signed by top AI researchers in industry and academia (as well as by Hawking and Musk).

Signatures include those of the president of the [Association for the Advancement of Artificial Intelligence](#), the founders of AI startups [DeepMind](#) and [Vicarious](#), and well-known researchers at Google, Microsoft, Stanford and elsewhere.

In the interests of full disclosure, mine is also one of the early signatures

on the list, which continues to attract more support by the day.

The open letter argues that there is now a broad consensus that AI research is progressing steadily and its impact on society is likely to increase.

For this reason, the letter concludes we need to start to research how to ensure that increasingly capable AI systems are robust (in their behaviours) andbeneficial (to humans). For example, we need to work out how to build AI systems that result in greater prosperity within society, even for those put out of work.

The letter includes a link to a document outlining some [interdisciplinary research priorities](#) that should be tackled in advance of developing [artificial intelligence](#). These include short-term priorities such as optimising the economic benefits and long-term priorities such as being able to verify the formal properties of AI systems.

The AI threat to society

Hollywood has provided many memorable visions of the threat AI might pose to society, from Arthur C. Clarke's [2001: A Space Odyssey](#) through [Robocop](#) and [Terminator](#) to recent movies such as [Her](#) and [Transcendence](#), all of which paint a dystopian view of a future transformed by AI.

My opinion (and one many of my colleagues share) is that AI that might threaten our society's future is likely still some way off.

AI researchers have been predicting it will take another 30 or 40 years now for the last 30 or 40 years. And if you ask most of them today, they (as I) will still say it is likely to take another 30 or 40 years.

Making computers behave intelligently is a tough scientific nut to crack. The human brain is the most complex system we know of by orders of magnitude. Replicating the sort of [intelligence](#) that humans display will likely require significant advances in AI.

The human brain does all its magic with just 20 watts of power. This is a remarkable piece of engineering.

Other risks to society

There are also more imminent dangers facing mankind such as climate change or the ongoing global financial crisis. These need immediate attention.

The [Future of Humanity Institute](#) at the University of Oxford has a [long list of threats](#) besides AI that threaten our society including:

- nanotechnology
- biotechnology
- resource depletion
- overpopulation.

This doesn't mean that there are not aspects of AI that need attention in the near future.

The AI debate for the future

The [Campaign to Stop Killer Robots](#) is advancing the debate on whether we need to ban fully autonomous weapons.

I am organising a debate on this topic at the next [annual conference](#) of the Association for the Advancement of Artificial Intelligence later this

month in Austin, Texas, in the US.

[Steve Goose](#), director of Human Rights Watch's Arms Division, will speak for a ban, while [Ron Arkin](#), an American roboticist and robo-ethicist, will argue against it.

Another issue that requires more immediate attention is the impact that AI will have on the nature of work. How does society adapt to more automation and fewer people needed to work?

If we can get this right, we could remove much of the drudgery from our lives. If we get it wrong, the [increasing inequalities](#) documented by the French economist [Thomas Piketty](#) will only get worse.

We will discuss all these issues and more at the first [International Workshop on AI and Ethics](#), also being held in the US within the AAAI Conference on Artificial Intelligence.

It's important we start to have these debates now, not just to avoid the potential pitfalls, but to construct a future where AI improves the world for all of us.

This story is published courtesy of [The Conversation](#) *(under Creative Commons-Attribution/No derivatives).*

Source: The Conversation