ABC NEWS

The new laws of robotics — building on Asimov's science fiction legacy in the age of AI

ABC Radio National / By Antony Funnell for Future Tense

Posted Wed 9 Dec 2020 at 11:30pm, updated Thu 10 Dec 2020 at 7:04am

The new laws are designed to ensure AI works for humans, not against them. *(Unsplash: Andy Kelly)*

Way back in the early 1940s — long before smart phones, Siri and semi-autonomous weapons — the late great sci-fi writer Isaac Asimov drew-up a set of principles for the development of advanced robotic systems.

Asimov was essentially an optimist, but he realised that future AI devices, and their designers, might need a little help keeping on the straight and narrow.

Hence his famous Three Laws, which have influence in science and technology circles to this day.

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

Now, almost 80 years later, legal academic and artificial intelligence expert Frank Pasquale has added four additional principles.

He's given RN's Future Tense podcast the lowdown. Here's what you need to know.

What did Asimov get wrong?

Professor Pasquale says while Asimov's ideas were well founded, they assumed a certain technological trajectory that no longer holds — innovations are not always for the good of humanity.

"Whereas Asimov is talking about how we have robots help people and not harm them, I'm talking about how we democratise the direction of technology," he says.

And he draws an important distinction between AI and what he terms IA — intelligence augmentation.

"The artificial intelligence goal is often to replace human beings, to create, for example, a robot doctor or a robot teacher. The intelligence augmentation goal is to help human beings."

So, what are the new laws and why are they necessary?

Professor Pasquale says jobs that involve judgement and deliberation over choices should be kept for humans. *(Getty: Donald Iain Smith)*

Professor Pasquale says the stakes are high.

The optimal mix of robotic and human interaction is far too important to be determined only by corporations and engineers, he argues.

"If we go for an artificial intelligence vision of replacement, that's going to create a jobless future in many areas," he says.

"Whereas if we emphasise intelligence augmentation, that should actually increase both productivity and the value of labour."

New law 1: AI should complement professionals, not replace them

Professor Pasquale says some areas of the economy, particularly in manufacturing, will continue to see rapid automation, but jobs that involve judgement and deliberation over choices should be kept for humans.

"In areas like teaching, medicine, a lot of professional fields, you want to have people able to explain options to their clients, patients and students, rather than having some large tech firm just assume what is best and automate the result."

But he emphasises the need to recognise and help those made redundant by technology.

"Part of making it easier is to invest in people and to emphasise that everyone can have some role in society that involves work with judgement, with their expertise being valuable to that work."

New law 2: Robotic systems and AI should not counterfeit humanity

Devices should not be developed to mimic human emotions.

Professor Pasquale says while personal assistants like Siri might seem benign, they risk deceiving people and manipulating our feelings.

The future of human-computer interaction is going to involve tough judgement calls about how seamless personal interactions with robots should be, and whether some friction should be involved to create a level of distance.

Professor Pasquale also argues that we need to be more disciplined in the language we use around robotics.

"I resist even novelists or fiction writers calling the robot 'he' or 'she' like a person. I think that 'it' should always be the pronoun because I think that 'it' sort of denotes the fact that this is a machine and it's not a person."

Go deeper: What do our virtual assistants say about sexism? Quite a lot, as it turns out.

New law 3: Robotic systems and AI should not intensify zero-sum arms races

The unchecked development of smart robotic weapons systems risks spiralling out of control, Professor Pasquale warns.

And given our track-record with other military spending, there's every reason to suggest an arms race will develop over the development and deployment of AI weaponry.

"Very early on I think we have to say how we get societies to recognise that this is destructive, it's not providing real human services, it's just investing in the history of destruction," Professor Pasquale says.

He also warns against technology companies such as Google and Facebook engaging in an "arms race for attention" both among advertisers and platform users.

Hyper-competitiveness, he warns, is the road to "technological dominance and monopolisation".

Related: The Australian Defence Force is investing millions in 'killer robots' research.

New law 4: Robotic systems and AI must always indicate the identity of their creator(s), controller(s) and owner(s)

Greater levels of transparency are needed to increase accountability and deter inappropriate and illegal activity by both the owners and developers of technology.

Just as cars have license plates, so too should robots, says Professor Pasquale.

He believes no robotic system should ever be made fully autonomous. The risks, he says, are too high.

"It's important because we know how to punish people, we don't know how to punish robots or AI, and that's really important to the future of legal enforcement."

You decide: **Would you let a car determine who dies?**

But won't that put a brake on innovation?

The short answer is yes.

Professor Pasquale acknowledges the widespread application of his new laws would stall the development of certain technologies, but that, he says, would be for the public good.

"We've seen so many fields where technological advances have led to very troubling and harmful consequences like global warming. We've got to get ahead of these things," he says.

"We can't just look at something like global warming and say, well, we'll just get a technological fix out of it.

"We have to think very deeply about directing innovation, and innovation itself can't just be a watchword to stop regulation."

Read more from Future Tense

- How do you make a rule book for a space war?
- What COVID reveal about rising urban inequality
- Whatever happened to the idea of geoengineering the planet?
- How to help kids use technology safely

These laws are detailed in Professor Pasquale's book New Laws of Robotics: Defending Human Expertise in the Age of AI. To hear the full interview, listen to the Future Tense podcast.