



The Skynet Paradox

Published in "The Malta Independent on Sunday", 1st April 2018
By Dr Ian Gauci

"The unknown future rolls toward us. I face it, for the first time, with a sense of hope because if a machine, a Terminator, can learn the value of human life, maybe we can too." (Sarah Connor in the *Terminator 2: Judgement Day*).

Skynet is a central theme and a revolutionary artificial intelligence system built by Cyberdyne Systems in the Terminator movies. Accidentally, when we first think of artificial intelligence the first port of call is science fiction and sentient beings or robots. Even the famous Law of Robotics is the result of *Runaround*, one of Asimov's novels in the collection 'I Robot'.

But what is artificial intelligence (AI) exactly? Its origins can be traced back to the 18th century, from Thomas Bayes and George Boole to Charles Babbage, who constructed the first electronic computer. In his classic essay *Computing Machinery and Intelligence*, Alan Turing also imagined the possibility of computers created for simulating intelligence. In 1956, John McCarthy coined the definition for AI as 'the science and engineering of making intelligent machines'.

Despite this definition, we are still somewhat at a loss to bring out a clear and harmonised taxonomy of two simple words umbilically linked in AI.

The first limb, 'artificial', can mean something not occurring in nature or not occurring in the same form in nature. If we pause a little here and think of the current and envisaged advancements of 3D printing to construct human organs, technologies like Crisper and the advancement of science, this definition is a conundrum. The artificial is no longer linked to a programming output or a synthetic material, and without delving into the moral and ethical dilemmas, this is slowly but gradually blurring the legacy concept of what is natural and what occurs in the same form in nature.

This leads us to the second limb: intelligence. From a philosophical perspective, 'intelligence' is a vast minefield, especially if treated as including one or more of 'consciousness', 'thought', 'free will' and 'mind'. Philosopher Emmanuel Kant argued that the mind brings to experience certain qualities of its own that order it. These are the 12 *a priori* (deductive) categories of causality, unity, totality and the like, and the *a priori* intuitions of time and space. He also considered psychology to be an empirical inquiry into the laws of mental operations, where human intelligence is also visualised as a compendium of abilities.

Nick Bostrom, Professor of Philosophy at Oxford University and director of the Future of Humanity Institute, whilst categorising this intelligence in AI as general and super intelligence, also questions how we should act - mindful of the knowledge that we will eventually live alongside artificial minds exponentially more powerful than our own. The late Stephen Hawking also acknowledged this and was very vocal on the effect of this intelligence on human kind.



66, Old Bakery Street, Valletta, VLT 1454 - MALTA
T. +356 2124 2713 | +356 2722 6484
E. agtg@aflexion.com

 aflexion.com

AFILEXION ALLIANCE MEMBER FIRMS





Philosopher John Smart thinks that, given their processing capacity, AIs would be "vastly more responsible, regulated, and self-restrained than human beings" and that "if morality and immunity are developmental processes, if they arise inevitably in all intelligent collectives as a type of positive-sum game, they must also grow in force and extent as each civilisation's computational capacity grows".

To my mind, the correlation between the definitions of 'Artificial' and 'Intelligence', and the perceived outcomes, leave more questions than answers. What exactly would fall within this definition? Would this imply that AI can be self-sufficient and autonomous from the will of its creator? Can it thus act independently and be attributed rights, obligations and liabilities over such autonomous actions? Should it therefore also have moral status and/or legal personality?

In February of last year, the EU Parliament, with an unprecedented show of support, also took an initial step towards enacting the world's first Robot laws, where it was also suggested that sophisticated autonomous robots be given specific legal status as electronic persons. Saudi Arabia also granted citizenship to a robot, "Sophia". Estonia was also mulling over a decision to grant AI and robots a legal status that is somewhere between 'separate legal personality' and 'personal property' called 'robot-agent'. This could potentially allow AI to own their creations, as well as being liable for any damage, and introduce the concept of *lex machina criminalis*. However, this also highlights social, ethical and legal concerns that need a more profound analysis.

We might need some more time to understand the implications of such measures and maybe plan a more transparent and step-by-step transition to the reality that AI will bring about. The AI community, mindful of the dangers and uncertainties in this quadrant and the inefficacy of the science fiction norms created by Asimov, is building on the latter and working on the Asilomar AI Principles. These principles are aimed for the safe creation, use and existence of AI and will include, amongst others: Transparency (ascertaining the cause if an AI system causes harm); Value Alignment (aligning the AI system's goals with human values); and Recursive Self-Improvement (subjecting AI systems with abilities to self-replicate to strict safety and control measures). As George Puttenham's *The Arte of English Poesie* (1589) puts it:

'Ye haue another manner of disordered speech, when ye misplace your words or clauses and set that before which should be behind. We call it in English prouerbe, the cart before the horse, the Greeks call it Histeron proteron, we name it the Preposterous'.

Dr Ian Gauci is a partner at GTG Advocates and Afilexion Alliance, lectures in Legal Futures and Technology at the University of Malta and is a legal advisor on the National Blockchain Strategy Taskforce



66, Old Bakery Street, Valletta, VLT 1454 - MALTA
T. +356 2124 2713 | +356 2722 6484
E. agtg@afilexion.com

 afilexion.com

AFILEXION ALLIANCE MEMBER FIRMS

