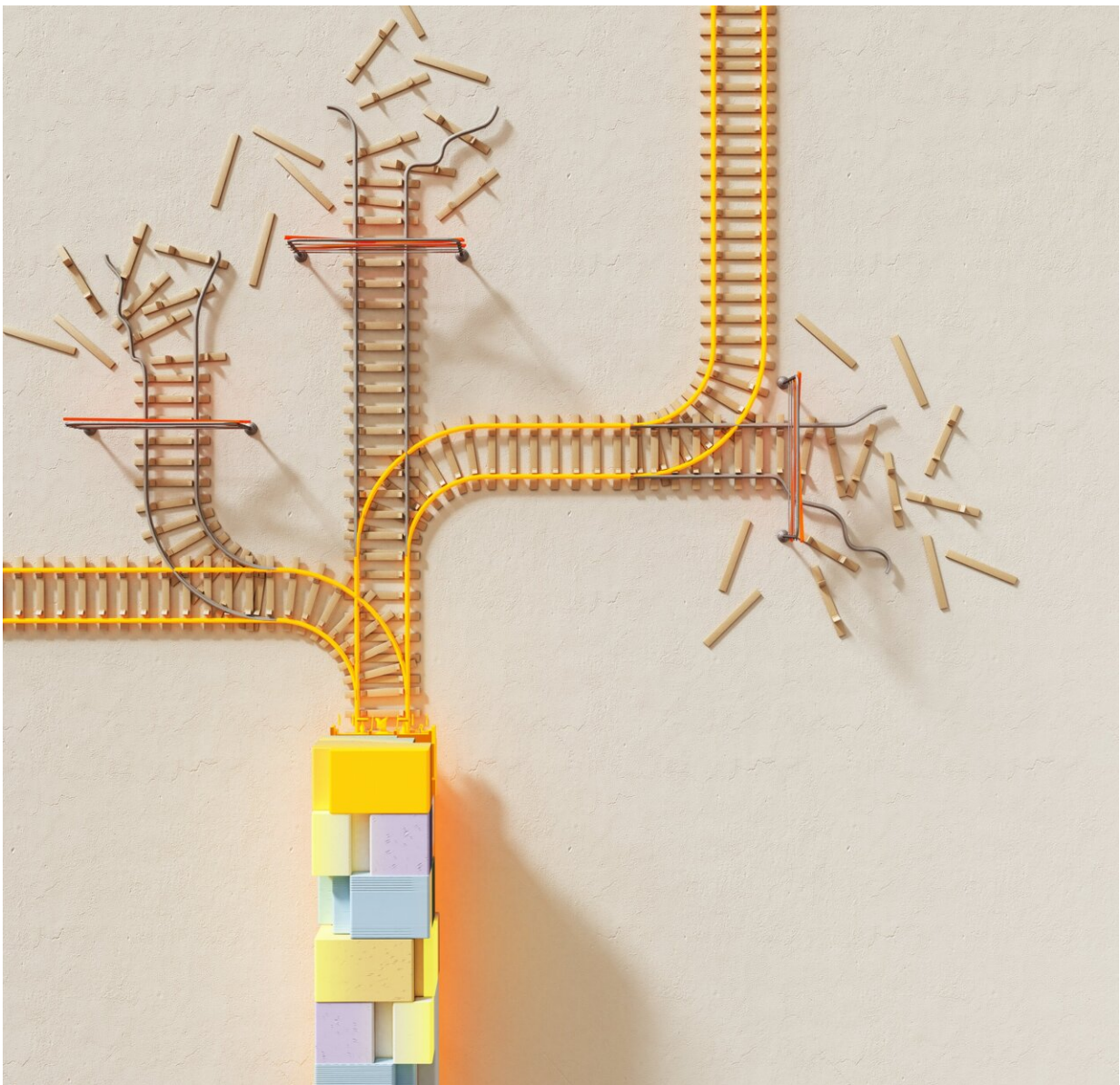


# Worried about AI taking over the world? You may be making some rather unscientific assumptions

September 24 2018, by Eleni Vasilaki

---



Credit: Google DeepMind from Pexels

,Should we be afraid of artificial intelligence? For me, this is a simple question with an even simpler, two letter answer: no. But not everyone agrees – many people, including the late physicist Stephen Hawking, [have raised concerns](#) that the rise of powerful AI systems could spell the end for humanity.

Clearly, your view on whether AI will take over the world will depend on whether you think it can develop intelligent behaviour surpassing that of humans – something referred to as "super intelligence". So let's take a look at how likely this is, and why there is much concern about the future of AI.

Humans tend to be afraid of what they don't understand. Fear is often blamed for racism, homophobia and other sources of discrimination. So it's no wonder it also applies to new technologies – they are often surrounded with a certain mystery. Some technological achievements seem almost unrealistic, clearly surpassing expectations and in some cases human performance.

## **No ghost in the machine**

But let us demystify the most popular AI techniques, known collectively as "[machine learning](#)". These allow a machine to learn a task without being programmed with explicit instructions. This may sound spooky but the truth is it is all down to some rather mundane statistics.

The machine, which is a program, or rather an algorithm, is designed with the ability to discover relationships within provided data. There are

many different methods that allow us to achieve this. For example, we can present to the machine images of handwritten letters (a-z), one by one, and ask it to tell us which letter we show each time in sequence. We have already provided the possible answers – it can only be one of (a-z). The machine at the beginning says a letter at random and we correct it, by providing the right answer. We have also programmed the machine to reconfigure itself so that next time, if presented with the same letter, it is more likely to give us the correct answer for the next one. As a consequence, the machine over time improves its performance and "learns" to recognise the alphabet.

In essence, we have programmed the machine to exploit common relationships in the data in order to achieve the specific task. For instance, all versions of "a" look structurally similar, but different to "b", and the algorithm can exploit this. Interestingly, after the training phase, the machine can apply the obtained knowledge on new letter samples, for example written by a person whose handwriting the machine has never seen before.

Humans, however, are good at reading. Perhaps a more interesting example is Google Deepmind's artificial Go player, which has surpassed every human player in their performance of the game. It clearly learns in a way different to humans – playing a number of games with itself that no human could play in their lifetime. It has been specifically instructed to win and told that the actions it takes determine whether it wins or not. It has also been told the rules of the game. By playing the game again and again it can discover in each situation what is the best action – inventing moves that no human has played before.

## **Toddlers versus robots**

Now does that make the AI Go player smarter than a human? Certainly not. AI is very specialised to particular type of tasks and it doesn't

display the versatility that humans do. Humans develop an understanding of the world over years that no AI has achieved or seem likely to achieve anytime soon.

The fact that AI is dubbed "intelligent" is ultimately down to the fact that it can learn. But even when it comes to learning, [it is no match for humans](#). In fact, toddlers can learn by just watching somebody solving a problem once. An AI, on the other hand, needs tonnes of data and loads of tries to succeed. But while humans develop breathtaking intelligence rapidly in the first few years of life, the key concepts behind [machine learning](#) are not so different from what they were one or two decades ago.

The success of modern AI is less due to a breakthrough in new techniques and more due to the vast amount of data and computational power available. Importantly, though, even an infinite amount of data won't give AI human-like intelligence – we need to make a significant progress on developing [artificial "general intelligence"](#) techniques first. Some approaches to doing this involve building a computer model of the human brain – which we're not even close to achieving.

Ultimately, just because an AI can learn, it doesn't really follow that it will suddenly learn all aspects of human intelligence and outsmart us. There is no simple definition of what human intelligence even is and we certainly have little idea how exactly it is represented in the brain. But even if we could work it out and then create an AI that could learn to become more intelligent, that doesn't necessarily mean that it would be more successful.

Personally, I am more concerned by how humans use AI. Machine learning algorithms are often thought of as black boxes, and less effort is made in pinpointing the specifics of the solution our algorithms have found. This is an important and frequently neglected aspect as we are

often obsessed with performance and less with understanding. Understanding the solutions that these systems have discovered is important, because we can also evaluate if they are correct or desirable solutions.

If, for instance, we train our system in a wrong way, we can also end up with a machine that has learned relationships that do not hold in general. Say for instance that we want to design a machine to evaluate the ability of potential students in engineering. Probably a terrible idea, but let us follow it through for the shake of the argument. Traditionally, this is a male dominated discipline, which means that training samples are likely to be from previous male students. If we don't make sure, for instance, that the training data are balanced, the machine might end up with the conclusion that engineering students are male, and incorrectly apply it to future decisions.

Machine learning and artificial [intelligence](#) are tools. They can be used in a right or a wrong way, like everything else. It is the way that they are used that should concerns us, not the methods themselves. Human greed and [human](#) unintelligence scare me far more than [artificial intelligence](#).

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation

Citation: Worried about AI taking over the world? You may be making some rather unscientific assumptions (2018, September 24) retrieved 19 April 2024 from <https://phys.org/news/2018-09-ai-world-unscientific-assumptions.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.