A brief history of AI



With the current buzz around artificial intelligence (AI), it would be easy to assume that it is a recent innovation. In fact, AI has been around in one form or another for more than 70 years. To understand the current generation of AI tools and where they might lead, it is helpful to understand how we got here.

Each generation of AI tools can be seen as an improvement on those that went before, but none of the tools are headed toward consciousness.

The mathematician and computing pioneer Alan Turing published an article in 1950 with the opening sentence: “I propose to consider the question, ‘Can machines think?’”. He goes on to propose something called the imitation game, now commonly called the Turing test, in which a machine is considered intelligent if it cannot be distinguished from a human in a blind conversation.

Five years later, came the first published use of the phrase “artificial intelligence” in a proposal for the Dartmouth Summer Research Project on Artificial Intelligence.

From those early beginnings, a branch of AI that became known as expert systems was developed from the 1960s onward. Those systems were designed to capture human expertise in specialised domains. They used explicit representations of knowledge and are, therefore, an example of what’s called symbolic AI.

There were many well-publicised early successes, including systems for identifying organic molecules, diagnosing blood infections and prospecting for minerals. One of the most eye-catching examples was a system called R1 that, in 1982, was reportedly saving the Digital Equipment Corporation US$25m per annum by designing efficient configurations of its minicomputer systems.

The key benefit of expert systems was that a subject specialist without any coding expertise could, in principle, build and maintain the computer’s knowledge base. A software component known as the inference engine then applied that knowledge to solve new problems within the subject domain, with a trail of evidence providing a form of explanation.

These were all the rage in the 1980s, with organisations clamouring to build their own expert systems and they remain a useful part of AI today.

Enter machine learning

The human brain contains around 100 billion nerve cells, or neurons, interconnected by a dendritic (branching) structure. So, while expert systems aimed to model human knowledge, a separate field known as connectionism was also emerging that aimed to model the human brain in a more literal way. In 1943, two researchers called Warren McCulloch and Walter Pitts had produced a mathematical model for neurons, whereby each one would produce a binary output depending on its inputs.

One of the earliest computer implementations of connected neurons was developed by Bernard Widrow and Ted Hoff in 1960. Such developments were interesting, but they were of limited practical use until the development of a learning algorithm for a software model called the multi-layered perceptron (MLP) in 1986.

The MLP is an arrangement of typically three or four layers of simple simulated neurons, where each layer is fully interconnected with the next. The learning algorithm for the MLP was a breakthrough. It enabled the first practical tool that could learn from a set of examples (the training data) and then generalise so that it could classify previously unseen input data (the testing data).

It achieved this feat by attaching numerical weightings on the connections between neurons and adjusting them to get the best classification with the training data, before being deployed to classify previously unseen examples.

The MLP could handle a wide range of practical applications, provided the data was presented in a format that it could use. A classic example was the recognition of handwritten characters, but only if the images were pre-processed to pick out the key features.

Newer AI models

Following the success of the MLP, numerous alternative forms of neural network began to emerge. An important one was the convolutional neural network (CNN) in 1998, which was similar to an MLP apart from its additional layers of neurons for identifying the key features of an image, thereby removing the need for pre-processing.

Both the MLP and the CNN were discriminative models, meaning that they could make a decision, typically classifying their inputs to produce an interpretation, diagnosis, prediction or recommendation. Meanwhile, other neural network models were being developed that were generative, meaning that they could create something new, after being trained on large numbers of prior examples.

Generative neural networks could produce text, images or music, as well as generate new sequences to assist in scientific discoveries.

Two models of generative neural network have stood out: generative-adversarial networks (GANs) and transformer networks. GANs achieve good results because they are partly “adversarial”, which can be thought of as a built-in critic that demands improved quality from the “generative” component.

Transformer networks have come to prominence through models such as GPT4 (Generative Pre-trained Transformer 4) and its text-based version, ChatGPT. These large-language models (LLMs) have been trained on enormous datasets, drawn from the internet. Human feedback improves their performance further still through so-called reinforcement learning.

As well as producing an impressive generative capability, the vast training set has meant that such networks are no longer limited to specialised narrow domains like their predecessors, but they are now generalised to cover any topic.

Where is AI going?

The capabilities of LLMs have led to dire predictions of AI taking over the world. Such scaremongering is unjustified, in my view. Although current models are evidently more powerful than their predecessors, the trajectory remains firmly toward greater capacity, reliability and accuracy, rather than toward any form of consciousness.

As Professor Michael Wooldridge remarked in his evidence to the UK Parliament’s House of Lords in 2017, “the Hollywood dream of conscious machines is not imminent and indeed I see no path taking us there”. Seven years later, his assessment still holds true.

There are many positive and exciting potential applications for AI, but a look at the history shows that machine learning is

not the only tool. Symbolic AI still has a role, as it allows known facts, understanding and human perspectives to be incorporated.

A driverless car, for example, can be provided with the rules of the road rather than learning them by example. A medical diagnosis system can be checked against medical knowledge to provide verification and explanation of the outputs from a machine learning system.

Societal knowledge can be applied to filter out offensive or biased outputs. The future is bright and it will involve the use of a range of AI techniques, including some that have been around for many years.

Adrian Hopgood is an independent consultant and Emeritus Professor of Intelligent Systems at the University of Portsmouth in the United Kingdom. The opinions expressed in this article are that of the author and do not necessarily reflect any official policies or positions of the AEU or SSTUWA. This article was first published on The Conversation *website and has been reproduced here with permission.*

THE CONVERSATION

Western Teacher volume 53.6 September 2024

By Adrian Hopgood

Authorised by Sally Dennis, General Secretary, The State School Teachers' Union of W.A.
ABN 54 478 094 635 © 2026