Special report

GHOST IN THE MACHINE

Artificial intelligence fever is upon us again, with our increasingly wily ways of smartening up silicon prompting Stephen Hawking, Bill Gates and Elon Musk to issue warnings about super-intelligent robots. What happens when you weigh up the potential against the peril? *Director* asked the experts

It was one giant step for computer-kind. In May 1997, a 1.4-tonne block of silicon, christened 'Deep Blue' by its creators at IBM, beat the then world chess champion, Garry Kasparov. The Russian grandmaster was not only incredulous, but more than a little spooked. Only human intervention, he suggested, could have pulled off the decisive move – a sacrifice as part of a long-term strategy.

Deep Blue's victory was widely thought to usher in an age in which machines could genuinely think. Bar-room pundits talked of imminent doom, their fears fostered by a tide of fictional accounts of cybernetic revolt that began with Czech writer Karel Čapek's play *R. U. R.*, which introduced 'robot' to the global lexicon in 1921, and crested decades later with the release of films such as *2001: A Space Odyssey* and the *Terminator* franchise. (Interestingly, *The Matrix* was released two years after Deep Blue's coup.)

But the reality was actually more prosaic. Kasparov was not beaten by a machine that could think, but by some extremely powerful processors and a dizzyingly complex set of algorithms, which together could analyse an almost infinite number of variables and outcomes. While chess masters typically think 10 moves ahead, Deep Blue could plan its game 74 moves ahead. "Kasparov felt at key stages of the match that the computer was reading his mind," says Nigel Shadbolt, professor of artificial intelligence at the University of Southampton. "It wasn't, but he got unnerved by that."

In fact, Shadbolt points out that, for all its complexity and sophistication, Deep Blue was really a one-trick pony. "That machine couldn't have played a game of draughts, whereas a human could be taught a range of games very quickly," he says. "A machine cannot transfer expertise from one domain to another like we can. Humans don't really have one brain – it's several brains layered on top of one another, from amphibian to reptile to early mammal. We have an endocrine system – we're hormonally driven. We live in complex carbon-based bodies. And so this whole richness of being a human is something that it'll be extraordinarily hard to capture."

Charlotte Golunski, co-founder of Sense – an intelligent recognition platform for wearables and other smart devices – agrees, pointing out that machines haven't yet reached what she calls 'the fusion layer'. "Humans understand the world around us by fusing information from different sources simultaneously – for example, by combining visual cues with sounds plus information about our context and how we got into each situation," she says. "This combination of data helps us make faster and more accurate decisions. Such fusion is still very difficult for computers."

Golunski also points to our unique capacity for absorbing our surroundings. "Once someone understands the environment of what is going on around him or her, they can then form inferences that lead to hidden information and understanding,"

»

Words Nick Scott

she says. "This is often hinted at with contextual subtleties that currently prove incredibly hard for computers to understand. Computational neural networks only loosely mimic the way the brain works – which in itself remains a mystery to neuroscientists."

And yet there's been another spike in AI hysteria of late, prompted by notables such as Bill Gates, SpaceX founder Elon Musk (who referred to it as "summoning the demon") and Professor Stephen Hawking warning of the perils of creating anything which surpasses human intelligence. "It would take off on its own and redesign itself at an ever-increasing rate," Hawking told the BBC last December. "Humans, who are limited by slow biological evolution, couldn't compete, and would be superseded."

So is this man-made, rational, self-aware silicon entity – capable of emotional responses, inference and, ultimately, rebellion – a possibility in the future? Are we a few careless technological advances from creating a mechanical Bond villain – possibly stroking an equally malevolent animatronic moggy? Those in the field insist we're nowhere near that. "We haven't got the faintest clue as to how to build a self-aware, genuinely intelligent AI," says Shadbolt. And yet, the sober reality is actually not only less scary, but also much, much more exciting than the hype-fuelled, rabid speculation.

Smart thinking
Defining AI is a tricky business. "We've got used to the term now to denote any program that is kind of smart," says Shadbolt. While the pedantic observer might argue that the abacus, along with early automatons built by Greek, Egyptian and Chinese engineers, were all forms of it, the AI story really begins in 1956. That summer, at a two-month, 10-person conference at Dartmouth College in Hanover, New Hampshire, the term 'artificial intelligence' was coined. Attendees concluded that a machine as intelligent as a human would be created in no more than a generation. Funding was poured their way.

The enormity of their task soon became brutally apparent, and the only real breakthrough in the subsequent years was Eliza, a chatterbot built in 1966 by Dr Joseph Weizenbaum. Named after George Bernard Shaw's character Eliza Doolittle, Eliza was an early model for automated customer service systems. The problem was, any conversation with her was rudimentary, stilted, repetitive and often fruitless. (Whether we've come on leaps and bound by sticking with humans is, of course, another issue.)
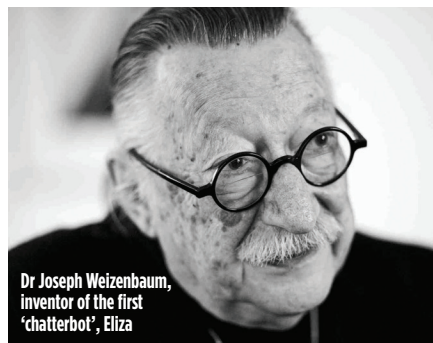
Dr Joseph Weizenbaum, inventor of the first 'chatterbot', Eliza
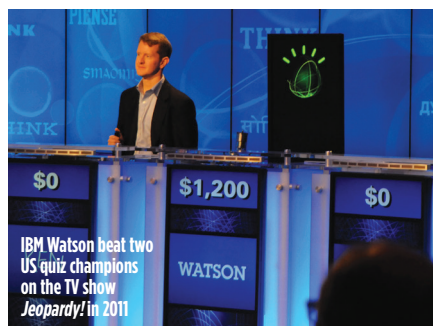
IBM Watson beat two US quiz champions on the TV show *Jeopardy!* in 2011

❝ **Stopping would hold back the progress driving smart cities, connected devices and wearable technology"**
Charlotte Golunski, Sense

The field has endured peaks and troughs in interest, and therefore government funding, ever since, and IBM's Deep Blue was the next major milestone. Now, though, AI is hitting a purple patch – hence the finger-wagging caution from Gates, Hawking and Musk. IBM Watson – the cognitive system that beat two grand champions at US quiz show *Jeopardy!* in 2011 – is these days helping oil and gas companies decide where to drill, lawyers to compile cases, and police to navigate seemingly unsolvable crimes. It's now being primed to go into medicine, an area in which its ability to hypothesise, based on the processing of vast reams of medical research papers and diagnostic images, may correct the status quo, in which preventable medical errors, resulting from poor decision-making, are the number-three killer in the US, behind only heart disease and cancer.

Meanwhile in Japan – perhaps unsurprisingly, a nation that embraces AI with gusto – it was recently announced that Nao, a robot developed by the French company Aldebaran Robotics (a subsidiary of Japanese telecoms and internet giant SoftBank), will be greeting customers at branches of Mitsubishi UFJ

Financial Group from April. Japan was also the origin of Paro – a robot seal used for dementia therapy as part of a trial project in Sheffield last summer.

British intelligence
The UK punches well above its weight in AI. Last October, Dark Blue Labs and Vision Factory, two Oxford University spin-off companies specialising in machine learning and computer vision, were bought out by Google (the US giant declined to comment for this article). Another impressive AI innovation, developed in the UK and set to make waves in the workplace, is Anomaly42 – a technology whose transparent and configurable algorithms eliminate one of AI's fundamental flaws: the dangers, ranging from inconvenience to Armageddon, of creating inanimate objects 'intelligent' enough to become autonomous. "The traditional problem is that algorithms can over-learn, which can result in a gradual deterioration in decision-making," says Freddie McMahon, director of strategy and innovation at Anomaly42. "This new form of AI intends for the 'IQ' of the 'machine' to be an aggregation of human intelligence which, over time, will typically outperform the capability of a human individual." Anti-money-laundering efforts and combatting the financing of terrorism are among Anomaly42's current applications. "We're also pioneering new capabilities in areas such as patents and healthcare – prevention at scale," says McMahon.

Also making the leap from artificial knowledge to artificial understanding, meanwhile, is 'Amelia', a new virtual customer service agent that can, manufacturers IPsoft claim, understand both what callers say and how they feel. "Other systems recognise words, but they don't grasp their underlying meaning, which limits their ability to solve many everyday business problems in a natural way," says IPsoft's UK CEO Richard Warley. "Because Amelia's neural ontology is modelled on the way humans process information, she's able to grasp concepts and meanings conveyed in dialogue, understand context, apply logic, infer implications and even sense emotions."

Amelia's ability to 'listen' and establish precisely what a customer wants by asking clarifying questions effectively introduces the power of reason into the AI arsenal, and could make *Little Britain*'s "Computer says no" sketch – which satirised the polarised positions of customer service professionals in the infancy of corporate technicalisation – look like a period piece just over a decade after it hit our TV screens.

Both Anomaly42 and Amelia are examples of a specific branch of AI focused on non-human entities having the capacity to learn. It's an area whose enormous potential for positive change is under threat, according to Golunski. "The damaging effect of negative publicity around AI is the restriction of development in machine learning," she says. "Stopping work within the AI field would hold back the progress which is driving smart cities, connected devices and wearable technology."

Other promising areas of AI include real-time translation and biomimetic engineering – observing the cunningly evolved systems that work in nature, and building those design cues into robots and programs, as Shadbolt describes it. "We've come up with some nifty ideas that nature hasn't got around to implementing," he adds. "There's a vast space of possibilities..."

The human touch
It's impossible to conceive exactly what possibilities AI offers for our future – something that will never stop people speculating, of course. Interest in the concept has been stirred by the techno thriller *Ex Machina*, directed by *The Beach* author Alex Garland and starring Colin Firth and Samuel L Jackson, which hit UK cinemas in January.

Shadbolt has an explanation for why AI is riding the crest of the zeitgeist. "Partly because of the audacity of what it's trying to do, AI has attracted very bright people," he says. "What they've been trying to build – voice recognition systems, robots, remote navigation devices – is a result of looking at humans and thinking, 'How have we learned to do this so fantastically well?'"

Shadbolt says the technology we take for granted – voice-input command, thumbprint recognition, predictive text – has all come from algorithms and coding developed in AI research labs. "When you're trying to understand how to build intelligence systems, you divide and conquer," he says, "and along the way better programs have been built and computers have got more powerful – from when I started working in AI until now, there's been a million-fold increase in the power of the computers. Nothing else on the planet has

AI IN NUMBERS
200–300 million
The number of chess positions per second evaluated by Deep Blue, the super-computer that beat the then reigning world chess champion Garry Kasparov in 1997.

76
The percentage score achieved by Siri in a test comparing the abilities of three virtual assistants by asking 18 questions in English, French, Italian, Mandarin and German. Google Now came second, Cortana last.

13.6 million
The amount in US dollars raised in five months by Olcan Sercinoglu to launch behaviour prediction firm Scaled Inference Inc. "Many AI companies have raised significant money without any product plans," he later said.

33
The percentage of judges, in a 2014 experiment, who believed a super computer was a real boy. It was the first machine to pass the Turing test – when a computer's behaviour is indistinguishable from that of a human.

developed that fast." So what does this mean for company directors wanting to future-proof their businesses today? "Decision-making support will increasingly be provided by intelligent programs – whether you work in trading, manufacturing, retail or logistics," says Shadbolt.

And what of the more distant future? Ray Kurzweil, AI expert at Google, predicts that by 2029 robots will reach human levels of general intelligence – an event he refers to, with characteristic gravitas, as 'the singularity'. Is our status as planetary Big Cheese at risk? Not according to Shadbolt – though he concedes there are dangers. "We have to take real care when it comes to the powers we give to our machines – we've seen that with financial trading systems," he says. "When do you pull the plug if things get into deadlock? How much automatic control do you give when shutting down a nuclear reactor?
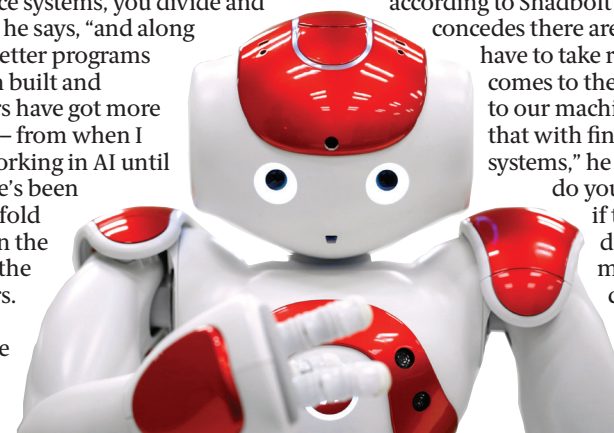
There are always issues when we apply our technology on the boundaries of critical decision-making. The threat [will only be realised] if we don't think hard about those boundaries – it's not that the machine might do the thinking for us and decide we're getting in the way."

Cyber Christians
For Shadbolt, the next stage is augmented intelligence – people with machine capability added on. "As we start putting engineered implants into ourselves to enhance our own abilities, there won't be a hard distinction between the stuff which rusts and the stuff which rots," he says, adding that if something which has a stream of consciousness is ever created, the ethical implications will be enormous. "At what stage in the evolution of our machines do we start to worry about cruelty, rights and so on?" he asks. The Reverend Christopher Benek of Providence Presbyterian Church in Florida, meanwhile, is more concerned with advanced AI forms' souls: last month he announced that, should they come into being during his lifetime, he intends to convert them to Christianity.

For William Higham, consumer futurist and founder of trends consultancy Next Big Thing, the immediate threats are akin to those posed by the Industrial Revolution. "It's got similar implications in terms of potential impact on jobs," he says. "Only this time, it's not going to affect only manual labours. Think of people like clerks in legal chambers, or anyone who works with data entry or analysis. But as with the Industrial Revolution, there are huge benefits too. You don't stop something like this in its tracks because of fear. But we need to ensure people are reskilled."

Higham is confident of a continued breaking down of boundaries between people and machines. "Their learning characteristics are going to deepen," he says. "They're going to be able to remember more and more about us, and recognise us better, and thus have more human functions and intuition. And our demands will grow – 'press one for yes, press two for no' isn't going to suffice. People will want things more humanised. The word 'human' will be a major buzzword over the next few years."

In short, we need to treat AI with the same rational, measured humanity and wisdom that we've not – yet – managed to replicate in non-human form. ⓓ

Watch a trailer for Alex Garland's *Ex Machina* at
exmachinamovie.co.uk

🐦 @NextBigThingCo

© MARK FILM