How to implement ChatGPT: a sober look at the opportunities and risks of a new tool with near limitless potential

Generative AI is an exciting new technology with seemingly endless applications, undoubtedly kicking off a race for forward-looking businesses to start using it. Dr Caroline Chibelushi, Artificial Intelligence expert at Innovate UK KTN, discusses how we should be approaching ChatGPT, exploring the opportunities and pitfalls of this AI chatbot so that businesses can use it to their benefit and with minimal risk.

We are experiencing the early days of a new industrial revolution. From the printing press to the internet, we're becoming used to big step-changes in technology that have almost unforeseeable consequences long into the future. AI is the next step in that evolution of technology, but its potential for both good and bad is far greater than anything that's come before.

Either way, it's going to change the way we work and interact with each other permanently. That is abundantly evident with the release and mega-popularity of ChatGPT. The technology is not strictly speaking "new", but it is the most advanced natural-language processing tool publicly available. It's trained on massive amounts of data and it is spooky in its ability to pass the Turing Test – a test of machine sentience imagined by of the greatest minds of the 20th Century, Alan Turing – that basically states that if an AI can mimic human speech so well as to be convincing, it is sentient.

Obviously, the Turing Test is not a true test of intelligence, and ChatGPT is not sentient at all. It is, however, convincing. And that makes it capable of supporting businesses in all sectors, with tasks that would have previously taken a decent amount of mental manpower.

Whether assisting teachers with creating lesson plans, writing customised job descriptions for HR departments, writing and debugging code, there is a lot of potential for ChatGPT to augment business for the better.

But being convincing can also be a risk. Implementing any new system in a business comes with risk, this is normally self-evident, but ChatGPT seems to be an exception. If we get ahead of ourselves without truly understanding the tool we are using, we are exposed.

The first thing to be clear about is that there should be a clear distinction between testing a new tool, and adopting it wholesale. Testing should be done in a sandbox, so there is no risk to your day-to-day operations. Only once you've tested a tool to its limits can you safely implement it.

This is how OpenAI themselves have approached the development of ChatGPT – the tool itself is very clearly still in BETA, it's free and so there is limited liability. That's why the launch of the "ChatGPT Plus" paid-for version came after the free version.

OpenAI has been commendably open about ChatGPTs limitations so far, but the testing phase is not yet over.

We know, for example, that ChatGPT can generate an answer which is convincing, human-like and confidently stated: but not all of these answers are correct. Unless we're fact-checking, there's a real risk of spreading misinformation. And it's not just a general muddying of the waters at stake here, there are demos online of ChatGPT's ability to generate seemingly sound legal agreements. If anyone was foolhardy enough to trust the result without

fact-checking it, they could be in for costly legal troubles.

Recently, we've seen Google's AI 'Bard' cost its parent company Alphabet over $100B after making a factual error in its first demo. In response, Microsoft, which backs ChatGPT, has added a disclaimer to its AI-powered search engine, stating that " […] Surprises and mistakes are possible. Make sure to check the facts and share feedback so we can learn and improve."

There are other, more subtle limitations too. ChatGPT is trained on data up to 2021, so it may generate reems of text that include out-of-date information, and may attempt to "guess" information that occurred after 2021. Likewise, AI is susceptible to bias, because it is created by and trained on data generated by humans. Racism, ageism, misogyny, politics: these are all very real and very human biases implicit in society, AI can easily reflect that back at us, and there are many more types of bias that we can't even necessarily account for until they crop up. When ChatGPT was asked to generate code to predict how senior someone would be in their career it, unfortunately, considered age, race and gender in its calculations.

So what can we learn from this? Is it even worth testing, let alone adopting, with all these risks?

It's important to be cautious of our own biases and naivety. AI can seem like something from science fiction, it can seem authoritative, *it can generate things that an individual user could never personally dream of creatin*g. But it is not infallible.

With that firmly in mind, ChatGPT becomes a tool of immense potential. As long as you are prepared to look and account for its limitations, there's no end to what generative AI can be used for.

By immense potential, I do mean immense. AI is not just a copywriter. It won't be long until we see AI tools that will support healthcare professionals and even the public in diagnosing their own non-surgical conditions, for example.

In fact, large-scale genomic and metabolomic data is already available that could be used to train AI tools to support self-diagnosis and treatment. An AI tool being fed with the results from PoC (point of care) devices for blood, urine, and stool test, and other results, can make diagnoses. The UK is extremely well placed to take advantage of research assets that combine genomic and other-omics data at scale.

UK Biobank has sequenced the exomes and whole genomes of its 500,000

participants which represents the largest collection of genome sequences anywhere in the world, all of which are linked to participants' detailed NHS health records.

This massively benefited the world by identifying the COVID-19 Alpha variant, which helped to explain epidemiological changes in data and transmission, as well as helping to track the emergence and spread of the Delta and Omicron variants as they replaced Alpha in turn. This knowledge, combined with the capability to develop PoC devices such as the COVID-19 lateral flow test – which proved to be fairly accurate and easy to use – allowed governments to tackle the pandemic more efficiently, and thus massively reduce deaths.

All businesses and sectors, and humanity itself, can benefit from AI, as long as we embrace it with a healthy dose of caution. To find out about Innovate UK KTN's role in supporting the UK AI industry, visit, *www.ktn-uk.org*.

Dr Caroline Chibelushi, Artificial Intelligence expert at *Innovate UK KTN*,

---

Article by DR CAROLINE CHIBELUSHI