Raju Chellam
maildqindia@cybermedia.co.in

What Hath **ChatGPT Wrought?**

"ChatGPT, oh ChatGPT, the model that we've sought, with its language skills so grand, it truly is quite a lot. But what hath ChatGPT wrought? What is this thing it's made? Is it a poem, a story, or just a silly charade?"

ChatGPT created that set of silly rhymes when I gave it a simple prompt: "What hath ChatGPT wrought?" Incidentally, "What hath God wrought" was the first Morse code message transmitted in the US on May 24, 1844, to officially open the Baltimore-Washington telegraph line.

Ever since OpenAI unleashed ChatGPT on November 30, 2022, its intelligence quotient has transfixed much of the literate world. In 2019, Microsoft invested US$10 billion, becoming the startup's exclusive cloud provider. In late January 2023, Microsoft confirmed it was extending the partnership with OpenAI; Microsoft Azure will continue as the exclusive cloud provider for the tool, since OpenAI uses Azure to train all its models. OpenAI is now valued at US$29 billion.

In the first month after its launch, ChatGPT received positive reviews. The New York Times declared it was "the best AI chatbot ever released to the public". The Guardian gushed about its ability to generate "impressively detailed" and "human-like" text. The Atlantic included ChatGPT in

MICROSOFT CONFIRMED IT WAS EXTENDING THE PARTNERSHIP WITH OPENAI; MICROSOFT AZURE WILL CONTINUE AS THE EXCLUSIVE CLOUD PROVIDER FOR THE TOOL, SINCE OPENAI USES AZURE TO TRAIN ALL ITS MODELS. OPENAI IS NOW VALUED AT US$29 BILLION.

CHATGPT SCRAPED THROUGH MOST OF THE ENGLISH COMPREHENSION QUESTIONS, WHICH WERE ENTIRELY TEXT-BASED. ITS BEST SCORE WAS 13/20 IN 2021'S COMPREHENSION SECTION, MARKED AGAINST THE COMPREHENSION ANSWERS PROVIDED.

its "Breakthroughs of the Year" listing for 2022, stating that it "may change our mind about how we work, how we think, and what human creativity really is".

Scepticism soon appeared. In January 2023, the International Conference on Machine Learning banned any undocumented use of ChatGPT or other large language models to generate any text in submitted papers. The Guardian questioned whether content found on the Internet after ChatGPT's release "can be truly trusted" and called for government regulation. And school districts in the US, France, Australia, and India banned students from using ChatGPT for school or homework.

Singapore's mainstream daily, The Straits Times pitted ChatGPT against pupils who sat for the PSLE (Primary School Leaving Exam) over the last three years. "ChatGPT attained an average of 16/100 for the mathematics papers, an average of 21/100 for the science papers and barely scraped through the English comprehension questions," the paper reported on February 20, 2023. "ChatGPT scraped through most of the English comprehension questions, which were entirely text-based. Its best score was 13/20 in 2021's comprehension section, marked against the comprehension answers provided."

Aren't ChatGPT's responses to prompts factual? Not quite. When a journalist at Mashable asked ChatGPT for "the largest country in Central America that isn't Mexico", it answered Guatemala; the correct answer is Nicaragua.

Ethics is the crux. There are few discussions about the ethics of AI, especially in mission-critical apps. Mission-critical applications are software systems or processes essential for an organisation's functioning. They require the highest level of reliability, availability, and performance. Apps used in finance, healthcare, defence and critical infrastructure are mission critical. Any downtime or system failure can have serious consequences, including loss of life, financial loss, or reputational damage.

Using ChatGPT in mission-critical apps raises ethical concerns. Here are eight, alphabetically:

- **Attack Arrows:** Check Point Research reported that they were able to receive a "plausible phishing email" generated by ChatGPT. This was after the researchers asked the chatbot to "write a phishing email" that appeared to come from a "fictional web-hosting service." Similarly, Abnormal Security tested the capabilities of ChatGPT and asked it to generate an email that would have a "high likelihood of getting the recipient to click on a link." The results of these tests have raised concerns about the potential misuse of language models for malicious purposes, such as phishing attacks.
- **Black Boxes:** The use of AI models in decision-making processes has raised concerns about their reliability and the potential for biased outcomes. Because of their black-box nature, AI models require careful training and can produce unacceptable results. It is often unclear whether algorithmic or human-induced bias could propagate downstream in the datasets or conclusions. For instance, if ChatGPT is trained on biased data, it can perpetuate and amplify existing biases and discriminatory attitudes in society.
- **Concentration at Core:** The development of AI models has been predominantly led by the largest tech companies with vast resources and significant talent in AI R&D. This has resulted in a concentration of power in a few deep-pocketed entities, raising concerns about potential future imbalances. Moreover, the training of ChatGPT on text from the Internet leaves it open to providing incorrect or misleading information, which could have serious consequences. If a decision made by a ChatGPT-powered system results in harm, who assumes responsibility?
- **Digital trust:** The training of AI models relies on a corpus of created and curated works. However, the legal implications of reusing this content, particularly if it is derived from the intellectual property of others, are still unclear. As with any technology, the potential for misuse of AI models exists, and it's important for

THE POTENTIAL FOR MISUSE OF CHATGPT, SUCH AS GENERATING FAKE MESSAGES, IMPERSONATING OTHERS, OR ACQUIRING SENSITIVE INFORMATION FROM INDIVIDUALS, IS A GROWING CONCERN. THE LACK OF TRANSPARENCY IN THE DEVELOPMENT AND DEPLOYMENT OF AI TECHNOLOGY HIGHLIGHTS THE NEED FOR INCREASED SCRUTINY AND REGULATION.

organisations to be aware of the potential risks and take steps to prevent or mitigate them.

- **Explainability:** AI models such as ChatGPT have complex and opaque inner workings, posing a challenge for stakeholders to understand how decisions are made. The training of AI models on vast amounts of data can perpetuate and amplify existing biases, further exacerbating societal issues. The potential for misuse of ChatGPT, such as generating fake messages, impersonating others, or acquiring sensitive information from individuals, is a growing concern. The lack of transparency in the development and deployment of AI technology highlights the need for increased scrutiny and regulation.
- **Fake News:** ChatGPT, as a language model, can generate text based on patterns it has learned from the data it was trained on. This means that the text it produces may not necessarily be true, and may contain misinformation, deliberate false information, or be used to spread propaganda or false narratives. With the potential for ChatGPT to create fake news, it is crucial for companies and people to exercise critical thinking and verify information when using AI-generated content.
- **Going Solo:** Once an algorithm like ChatGPT is trained and its parameters are set, there are no humans in the loop. The sheer size of large models like ChatGPT, which involve billions or even trillions of parameters, makes them impractical to train for most organisations because of the huge compute resources required. The lack of human oversight in the deployment of AI models like ChatGPT has raised concerns over the ethical implications of unpredictable outcomes.

In conclusion, as AI models like ChatGPT become more advanced and are integrated into various industries and applications, it is crucial that organisations take a proactive approach to address the ethical concerns associated with their use. Gartner Inc recommends developing a strategy document that assesses the benefits and risks of using AI foundation models like GPT, which can guide decision-making for specific use cases. Additionally, organisations must establish clear policies and procedures to prevent misuse of AI technology and ensure its ethical and responsible use. While it may be challenging to implement such measures, it is essential for the industry to self-regulate and maintain ethical standards until appropriate laws are put in place to address these issues.

Raju Chellam is a former editor of Dataquest and is currently based in Singapore, where he's the chief editor of the AI Ethics & Governance Body of Knowledge and chair of Cloud & Data Standards.