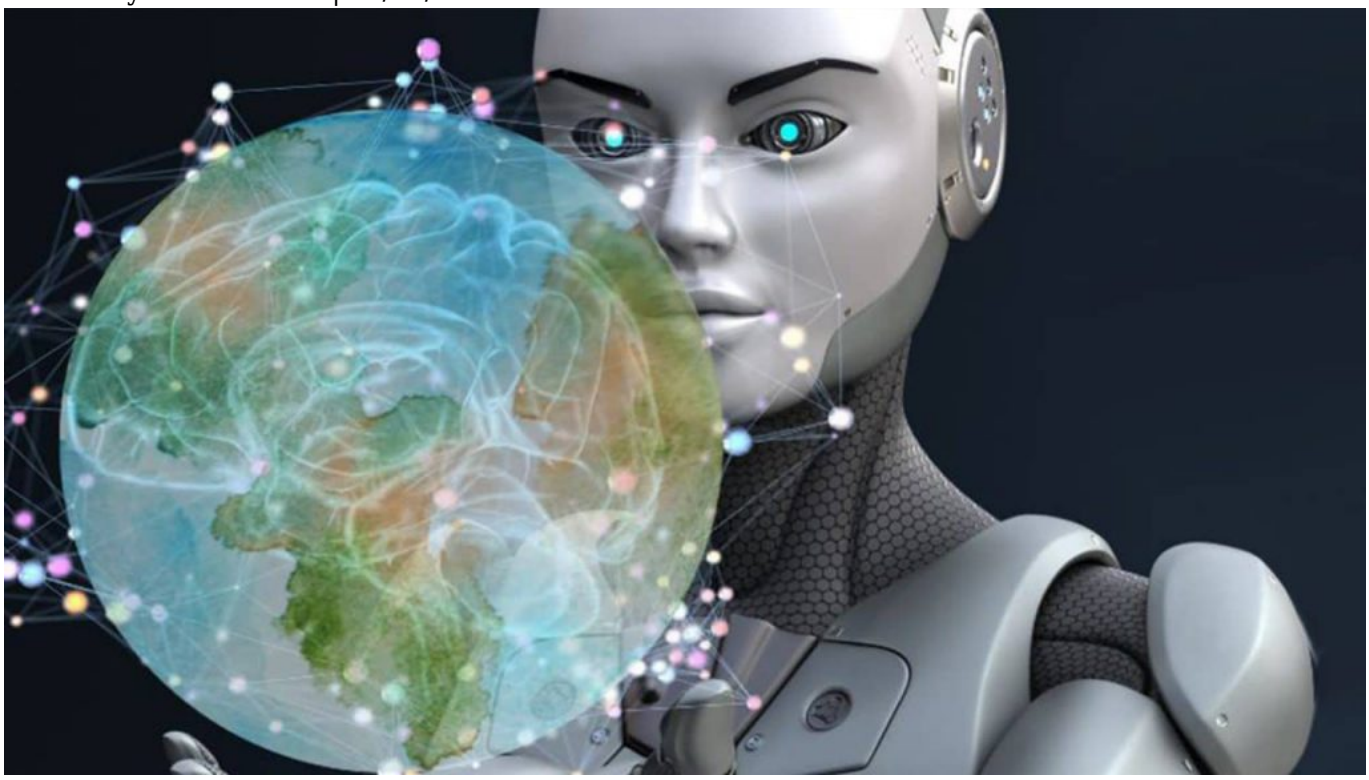
ChatGPT – Issues and Challenges

written by iasexam.com | 21/01/2023



Context – ChatGPT, an open AI platform which was released recently is taking online users by storm as they are marveling at its incredible power.

This coming of the age artificial intelligence tool is creating a buzz among computer scientists and programmers due to its creative capabilities.

What is ChatGPT?

- **ChatGPT** (Generative Pre-trained Transformer) is a chatbot launched by OpenAI in November 2022. It is built on top of OpenAI's GPT-3 family of large language models and is fine-tuned with both supervised and reinforcement learning techniques.

- **OpenAI –** ChatGPT has been developed by OpenAI, which is a research institute and company that focuses on developing **artificial intelligence** technology responsibly and safely. It was founded in 2015 by a group of entrepreneurs and researchers, including Elon Musk, Sam Altman, and Greg Brockman.

- **A human-like language model –** ChatGPT is built on a large-scale transformer-based language model that is trained on a diverse dataset of text and is capable of generating human-like responses to prompts. It is based on GPT-3.5, a language model that uses deep learning to produce human-like text.

- **More engaging with details –** However, while the older GPT-3 model only took text

prompts and tried to continue on that with its own generated text, ChatGPT is more engaging. It's much better at generating detailed text and can even come up with poems.

- **Keeps the Memory of the Conversations –** Another unique characteristic is memory. The bot can remember earlier comments in a conversation and recount them to the user.

- **Human-like Resemblance –** A conversation with ChatGPT is like talking to a computer, a smart one, which appears to have some semblance of human-like intelligence.

Chatbots and AI

What is ChatBot? – A **Chatbot** is a computer program that simulates human conversation either by voice or text communication and is designed to help solve a problem. Various organizations use chatbots to engage with customers alongside classic customer service channels phone, email and social media.

What is Artificial Intelligence? – AI is a constellation of technologies that enables machines to act with higher levels of intelligence and emulate the capabilities of sense, comprehend and act. Thus, computer vision and audio processing can actively perceive the world around them by acquiring and processing images, sound and speech.

Working of ChatGPT

- ChatGPT is what is called a Language Model, rather than a chatbot. A language model is a software that prints out a sequence of words as output that are related to some words given as input with appropriate semantic relation;

- In practical terms, it means that it can perform tasks like answering questions and carrying on a conversation with humans. It is often used in natural language processing (NLP) applications, such as speech recognition, automatic translation, and text generation.

- It is also a neural network. A neural network can be thought of as a large network of computers that can fine-tune its output of words based on the feedback given to it during stages of training: this training process and the technology together are called Reinforcement Learning.

- With a further refining technique called "Transformer", a neural network can accurately "understand" the context of a sentence or a paragraph. This "comprehension" can be used for multiple purposes like answering a question, summarizing a paragraph or an article, translating documents and so on.

Applications

- **Virtual Studymate –** ChatGPT could be used to generate practice questions or prompts for students to use when studying or preparing for exams. Input a prompt into ChatGPT that relates to the topic you are studying. For example, you might write 'generate practice questions from Indian Polity on fundamental rights'. ChatGPT will then generate a series of practice questions that could range from multiple choice, true/false, or short answers.
- **Writing Assistant –** ChatGPT can be used to generate text in a variety of styles and formats, such as stories, news articles, poems, and more. This could be useful for creative writing, or for generating content for websites and social media.
- **Human-like Search Engine –** One of the most used applications of ChatGPT is to answer specific questions and get answers in human talking style-dialogue format. It can also simplify complex theories. Microsoft is planning to integrate OpenAI's ChatGPT chatbot into its Bing search engine in order to take advantage of its ability to respond to a wide range of questions while mimicking human conversation styles.
- **Language Translation –** ChatGPT could be used to automatically translate text from one language to another, making it easier for people to communicate with each other across language barriers.
- **Customizable Chatbot –** You could use the ChatGPT model to create a chatbot that can answer questions or have a conversation with a user. ChatGPT could answer customer questions or provide assistance in online chatbots or virtual assistants, allowing businesses to provide 24/7 support to their customers. It is trainable with specific data to create a virtual customer service bot that can converse and respond to queries.

Issues and Challenges

- **Implications for Cybersecurity –** Check Point Research and others noted that ChatGPT was capable of writing phishing emails and malware, especially when combined with OpenAI Codex.
- **Unethical Use –** ChatGPT attempts to reject prompts that may violate its content policy. However, some users managed to jailbreak ChatGPT by using various prompt engineering techniques to bypass these restrictions in early December 2022 and successfully tricked ChatGPT into giving instructions for how to create a Molotov cocktail or a nuclear bomb, or into generating arguments in the style of a Neo-Nazi.
- **ChatGPT is not entirely accurate –** It is not entirely accurate, something even OpenAI has admitted. It is also evident that some of the essays written by ChatGPT lack the depth that a real human expert might showcase when writing on the same subject.
- **ChatGPT lacks depth like the human mind –** It doesn't quite have the nuance that a human would often be able to provide. For example, when asked by ChatGPT how one should cope with a cancer diagnosis. The responses were kind but generic. The type of

responses you would find in any general self-help guide.

- **Lacks the same experiences as Humans –** AI has a long way to go. After all, it doesn't have the same experiences as a human.

- **ChatGPT does not excel in code –** ChatGPT is writing basic code. As several reports have shown, ChatGPT doesn't quite excel at this yet. But a future where basic code is written using AI doesn't seem so incredible right now.

- **Still prone to Misinformation –** Despite the abilities of the bot, there are some limitations. ChatGPT is still prone to misinformation and biases, which is something that plagued previous versions of GPT as well. The model can give incorrect answers to, say, algebraic problems.

- **ChatGPT can write incorrect answers –** OpenAI understands some flaws and has noted them down on its announcement blog that "ChatGPT sometimes writes plausible-sounding but incorrect or nonsensical answers.

Conclusion

OpenAI's ChatGPT turned that simple experience into something entirely different. ChatGPT is a path-breaking example of an AI chatbot and what the technology could achieve when applied at scale. Limitations aside, ChatGPT still makes for a fun little bot to interact with. However, there are some challenges that need to be addressed before it becomes an unavoidable part of human life.