ChatGPT

ChatGPT is a computer program designed to conduct natural and compelling conversations with humans. It employs techniques of artificial intelligence and machine learning to comprehend and respond to text-based input. This enables it to engage in conversations that more closely resemble those between two humans. This makes it an effective tool for customer service, marketing, and content creation, among others. ChatGPT is analogous to a virtual assistant in that it can converse with you and answer your queries just like a human would.

OpenAI has created a language model for chatbots. This variant of the GPT-3 language model was developed specifically for generating conversational discourse. ChatGPT is a language model created by OpenAI for natural language processing (NLP) duties including text generation and query responding. It is based on the GPT (Generative Pre-trained Transformer) architecture, which employs deep learning techniques to generate responses to text input that are like those of humans.

ChatGPT is taught natural language patterns and structures using a large corpus of text data, such as novels, articles, and websites. It can generate responses with high accuracy and fluency to a variety of prompts, including chatbot conversations, email messages, coding in different programming languages, and social media posts.

To summarize, ChatGPT is an advanced language model developed by OpenAI based on the GPT-3 (Generative Pre-trained Transformer 3) architecture, representing a significant milestone in NLP. It has garnered widespread attention due to its remarkable ability to generate human-like text and engage in coherent conversations. ChatGPT brought GPT into the limelight made the process of interacting with an AI text generator simple and free to everyone.

GPT is presently the most popular Large Language Model (LLM), while it is likely to be more competition over the next years. For instance, Google has Bard which is powered by its own language engine Pathways Language Model (PaLM 2).

# Working

## ChatGPT works by attempting to understand the prompt and then spitting out strings of words that it predicts will best answers the question, based on the data it was trained on.

## GPT-3 was trained on roughly 500 billion "[tokens](https://beta.openai.com/tokenizer)," which allow its language models to assign meaning and predict plausible follow-on text more easily. Many words map to single tokens, though longer or more complex words often break down into multiple tokens. On average, tokens are roughly four characters long. OpenAI has stayed quiet about the inner workings of GPT-4, but we can safely assume it was trained on much the same dataset since it's even more powerful.

All the tokens came from a massive corpus of data written by humans. That includes books, articles, and other documents across all different topics, styles, and genres—and [an unbelievable amount of content scraped from the open internet](https://commoncrawl.org/). Basically, it was allowed to crunch through the sum of human knowledge.

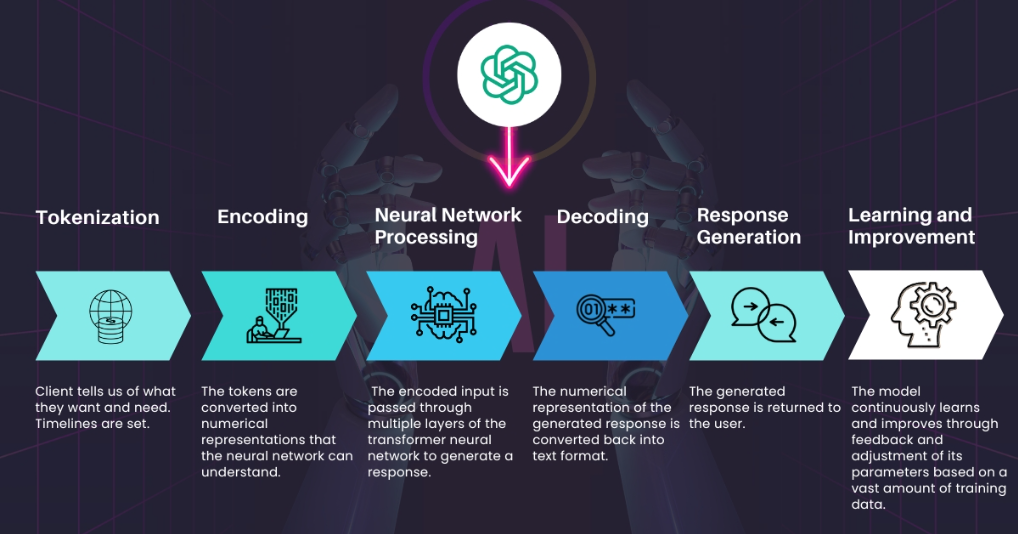
This humongous dataset was used to form a deep learning neural network—a complex, many-layered, weighted algorithm modelled after the human brain—which allowed ChatGPT to learn patterns and relationships in the text data and tap into the ability to create human-like responses by predicting what text should come next in any given sentence.

Though really, that massively undersells things. ChatGPT doesn't work on a sentence level—instead, it's generating text of what words, sentences, and even paragraphs or stanzas could follow. It's not the predictive text on your phone bluntly guessing the next word; it's attempting to create fully coherent responses to any prompt.

To further refine ChatGPT's ability to respond to a variety of different prompts, it was optimized for dialogue with a technique called reinforcement learning with human feedback (RLHF). Essentially, humans created a reward model with comparison data (where two or more model responses were ranked by AI trainers), so the AI could learn which was the best response.

Back to the neural network it formed. Based on all that training, GPT-3's neural network has 175 billion parameters or variables that allow it to take an input—your prompt—and then, based on the values and weightings it gives to the different parameters (and a small amount of randomness), outputs whatever it thinks best matches your request. OpenAI hasn't said how many parameters GPT-4 has, but it's a safe guess that it's more than 175 billion and less than the [once-rumoured](https://www.theverge.com/23560328/openai-gpt-4-rumor-release-date-sam-altman-interview) 100 trillion parameters. Regardless of the exact number, [more parameters doesn't automatically mean better](https://www.theatlantic.com/technology/archive/2023/03/openai-gpt-4-parameters-power-debate/673290/). Some of GPT-4's increased power probably comes from having more parameters than GPT-3, but a lot is probably down to improvements in how it was trained.

# Phases



**Tokenization:**  The input text is broken down into smaller pieces called tokens. It uses a tokenizer to segment the input text into individual words, phrases, or symbols and then creates tokens that represent each of these segments.

**Encoding:** The tokens are converted into numerical representations that the neural network can understand. It uses an encoder to convert each token into a vector of numbers that represents its meaning and context.

**Neural Network Processing:** The encoded input is passed through multiple layers of the transformer neural network to generate a response. The transformer network uses attention mechanisms to analyse the relationships between the input tokens and generate a response that is relevant and appropriate based on its vast knowledge base.

**Decoding:** The numerical representation of the generated response is converted back into text format. It uses a decoder to convert the numerical representations back into human-readable text.

**Response Generation:** The generated response is returned to the user. It generates a response based on its vast knowledge and analysis of the input query. And returns a response that is intended to be relevant and helpful to the user.

**Learning and Improvement:** The model continuously learns and improves over time through feedback and adjustments to its parameters based on a vast amount of training data. It is trained on a massive dataset of text from a variety of sources and is constantly updated with new data to improve its accuracy and relevance.

# Evolution

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **2018** | **2018** | **2019** | **2020** | **2022** | **2023** | **2023** |

InstructGPT

GPT-1

GPT-4

ChatGPT

GPT-3

Transformer

GPT-2

# ChatGPT vs Search engine

Google Search Engine and ChatGPT both have their place and purpose in the modern digital world, and they do those things very well. When using Google's potent search engine, users enjoy a seamless user experience characterized by relevance, accuracy, and search speed. Google actively crawls the totality of the web, retrieving content using sophisticated algorithms that provide superior search quality. ChatGPT, meanwhile, is meticulously designed for conversational assistance, employing cutting-edge AI and natural language processing to understand and respond to questions posed in commonplace language. Google extends a conventional search box, whereas ChatGPT facilitates dynamic, human-like communication.

Both systems use personalization features to generate customized results, but they manage user data differently. Google collects information from numerous data sources in order to remain current, whereas ChatGPT's knowledge base is vast but only contains data through 2021.

Although they share a commitment to user input and an obsession with improving the always shifting environment of user experience and information distribution, these two platforms couldn't be more dissimilar:

1. Search functionalities

Google's search engine is highly adept at scouring the vastness of the internet, using complex algorithms and ranking methodologies to catalogue and index its various search categories, such as web pages, images, videos, and other digital content. This ensures that users receive results that are relevant to their queries. ChatGPT employs cutting-edge AI and natural language processing techniques to interpret queries and generate conversationally appropriate responses. Its primary purpose and outcomes are to provide conversational support and natural language responses to questions.

1. User experience

Users must input a search query and then browse and select information based on Google's results to initiate a Google Search. ChatGPT engages users through natural language in a conversational format and provides instant interaction and responses. Google Search Engine provides customized search results based on a user's browsing history, geographic location, etc. Similarly, ChatGPT generates customized responses for users based on their previous interactions and preferences.

1. Accuracy and relevance

The Google Search Engine utilizes complex algorithms to determine the relevance of search results to a user's query. ChatGPT employs advanced AI and NLP techniques to comprehend user inquiries and provide pertinent responses. Utilizing advanced AI and NLP technologies, they both seek to provide users with accurate, up-to-date data. However, ChatGPT may not always have complete, current, or accurate information.

1. Response time

Having indexed millions of pages and generating search results in a matter of milliseconds, the Google Search Engine offers fast and effective search results. Similarly, ChatGPT provides real-time responses to user inquiries and processes questions quickly.

1. Personalization and user data

Google's search engine collects and stores user information, including search logs and geographic data, to provide customized search results and targeted advertisements. ChatGPT does the same thing by collecting and storing user information to provide more personalized responses and fine-tune its artificial intelligence and natural language processing.

1. Data source and accuracy

Information retrieval tools like Google Search and ChatGPT provide varying degrees of precision in response to user inquiries. ChatGPT's magic lies in its ability to grasp user context and provide appropriate replies, in contrast to Google Search's expertise in indexing billions of pages to display a wide collection of results. Google's search precision typically relies on well-established algorithms that consider credibility and prominence, whereas ChatGPT's accuracy is based on sophisticated semantic comprehension and extensive training data. While Google's index and the bots it employs to scan it are constantly being updated and expanded, the ChatGPT dataset may only include data up to the year 2021. While both have made progress in providing consumers with accurate results, it is important to recognize their distinct approaches, strengths, and sometimes shortcomings.



# Features

Developers that are interested in creating chatbot apps with natural language processing will find ChatGPT to be a very useful tool. Because of its capacity to grasp and react to complex and diverse linguistic inputs, as well as its ability to learn and improve over time, it is an excellent option for the creation of chatbots that can provide consumers something of value.

1. Size of database

Despite possessing a vast database, ChatGPT sometimes struggles to accurately capture the dynamic range of languages, experiences, and cultural diversities in real-time. Consequently, the nature of some replies occasionally results in their being unrelated or unsuitable in their context.

1. Pre-trained models

ChatGPT relies on machine learning algorithms that have been developed using data. Due to the potential presence of bias or inaccuracies in the training, it is not advisable to place full reliance on the accuracy of the replies. Systemic biases may also be present.

1. Resource intensive

The model derives its energy from multiple sources. This may be a limitation for mobile and low-power devices.

1. Timeline

ChatGPT does not yet encompass occurrences after 2021. Therefore, you will not discover answers to any questions after that time.

1. Cost

Currently, the model reportedly costs $3 million per month to operate. Long-term, this situation may become unmanageable. In addition, individuals must determine where the costs for users will be situated once the model review/research period concludes.

1. Complexity of facts

The model can manage a series of attributes about individuals, locations, and events, etc., only on occasion. Given its pre-trained structure, it could confound factors and provide random responses. Moreover, given that the information comes from the Internet, the responses may be biased. Even generative AI can be incorrect! It cannot always have multiple perspectives.

1. Consistency and accuracy

It delivers replies that are consistent and correct, so guaranteeing that consumers obtain information that is dependable and exact while they are conversing with the system.

1. Language flexibility

It can handle many languages, provide conversational help in a variety of languages, and break down barriers caused by language in order to provide a seamless experience for users.

1. Privacy and Security

ChatGPT offers a safe and private environment for conversations. It employs artificial intelligence to identify spam, censorship, and malicious content to provide a secure environment devoid of interference or manipulation. Moreover, ChatGPT neither stores nor transmits any personally identifiable information to third parties. To protect the privacy of users, all communications are encrypted and stored locally.

# Disadvantages

1. Lack of academic integrity

The fundamental problem in higher education about the use of ChatGPT is the preservation of academic integrity. Numerous instructors hold the viewpoint that the utilization of ChatGPT in writing tasks would just serve to foster instances of academic dishonesty, such as cheating and plagiarism. The rapid generation of replies using ChatGPT may potentially impede students' capacity to engage in brainstorming, critical thinking, and creative problem-solving while formulating their answers.

1. Provides inaccurate information

The information presented by ChatGPT has a semblance of plausibility and coherence in its composition, although it may be deficient in depth and perhaps lacks precision in its accuracy. Identifying the precise segments of the content that include factual inaccuracies might be a challenging task. The potential consequences of this situation include negative impacts on students' educational experiences and their ability to make informed decisions.

1. Biased responses

Artificial intelligence chatbots undergo training using an extensive dataset. If the dataset exhibits biases, it is likely that some outputs generated by Chat GPT will also reflect biases. These prejudices have detrimental effects and may exacerbate discrimination against certain demographic groups, leading to the establishment of an unfavourable social climate.

1. Limited knowledge

Despite the extensive training of ChatGPT, there exist some information that remains inaccessible to it. Because of this limitation, the system may exhibit limitations in generating accurate responses pertaining to specialized subject matters. It is possible that the system lacks awareness of current advancements across several disciplines.

1. Inability to multitask and understand context

The functionality of ChatGPT is limited to processing and addressing a singular job or enquiry at any one moment. When a student presents many inquiries simultaneously, ChatGPT may have challenges in effectively prioritizing and may experience a decrease in responsiveness, perhaps resulting in an incomplete response to all the posed queries.

In addition to this, ChatGPT may have challenges in comprehending the contextual intricacies and subtleties inherent in human language. For example, in situations when a student employs humour or sarcasm inside a query, ChatGPT may fail to discern such nuances, resulting in the provision of an incongruous answer.

1. Lack of EI

The importance of emotional intelligence (EI) is significant within educational environments. A human educator can comprehend the emotional states of pupils and then adapt their responses in a manner that is appropriate and effective. Educators could provide emotional support during times of difficulty. Virtual chatbots such as ChatGPT do not possess emotional intelligence (EI) and, as a result, are incapable of understanding human emotions. Despite seeming empathic, they have an inadequate response to intricate human emotions.

# SUMMARY

ChatGPT is an advanced artificial intelligence language model that has been created by OpenAI. The use of the GPT-3.5 architecture enables the provision of sophisticated capabilities in natural language interpretation and creation. Equipped with an extensive repository of information as of September 2021, ChatGPT can participate in text-based dialogues that resemble human discussions. It can respond to inquiries, provide explanations, extend support for creative writing endeavours, and perform more functions. The technology exhibits a wide range of applications in domains such as customer service, content development, education, and several other fields. As a result, it serves as a flexible tool for both organizations and people who need AI-driven language help. ChatGPT signifies a notable achievement in the advancement of conversational artificial intelligence, facilitating substantial and purposeful engagements with users across a wide range of scenarios.

# REFERENCES

* <https://www.youtube.com/watch?v=bSvTVREwSNw&pp=ygUHY2hhdGdwdA%3D%3D>
* <https://mlwhiz.com/blog/2023/07/07/how_chatgpt_works/>
* <https://zapier.com/blog/how-does-chatgpt-work/>
* <https://www.analyticsinsight.net/top-10-disadvantages-of-chatgpt-content/>
* <https://www.linkedin.com/pulse/5-advantages-disadvantages-chatgpt-frack-technologies/>
* <https://www.hurix.com/chat-gpt-pros-and-cons-of-using-chatgpt-in-higher-education/>
* <https://www.adamenfroy.com/chatgpt-vs-google>
* <https://www.syscraftonline.com/blog/how-does-chatgpt-work>
* <https://blog.digitalogy.co/top-useful-features-of-chatgpt/#:~:text=ChatGPT%20can%20be%20used%20for,and%20context%20of%20the%20user>.
* <https://www.profolus.com/topics/advantages-and-disadvantages-of-chatgpt/>
* <https://www.bostondigital.com/insights/google-search-versus-chatgpt-chatgpt-was-never-meant-be-search-engine#:~:text=The%20Google%20Search%20Engine%20employs,questions%20and%20deliver%20pertinent%20responses>.
* <https://thewebnoise.com/chatgpt/>
* <https://www.techtarget.com/whatis/definition/ChatGPT#:~:text=ChatGPT%20works%20through%20its%20Generative,of%20Generative%20Pre%2Dtrained%20Transformer>.
* <https://www.zdnet.com/article/how-does-chatgpt-work/>