

A Detail Cram on Chat Generative Pre-Training Transformer

P. Dharanyadevi
Department of CSE,
Puducherry Technological
University,
Puducherry, India.

S. Nandhini
Department of CSE,
Puducherry Technological
University,
Puducherry, India.

Keerthivasan K
Department of CSE
Chennai Institute of Technology,
Malayambakkam, Tamil Nadu, India

Abstract—An artificial intelligent automaton called ChatGPT (Generative Pre-training Transformer), created by OpenAI, that produces conversational chat that is human like via natural language processing. It is based on the OpenAI GPT-3 family of large language models, and it has been improved (a transfer learning technique) by utilizing both supervised and reinforcement learning methods. It works by generating text as an answer from natural language data supplied by the user. To predict the best ways to connect words, ChatGPT makes use of complex computational techniques and a large amount of data. Since its release, social media has been a hive of discussions regarding this new invention's possible perils. This paper discusses about the ChatGPT users' crams, evolution of ChatGPT, chronicle about ChatGPT, evolution of GPT's and working of GPT. This paper also elaborates the use case of chatGPT, merits, demerits of ChatGPT and its applications. This paper provided the foundation for researchers to comprehend the ChatGPT.

Keywords— OpenAI, ChatGPT, Search Engine, Use Case of ChatGPT, Google BART

I. INTRODUCTION

In recent months, the general-purpose AI chatbot prototype known as ChatGPT, which is the current internet craze, has swept the globe. It has rapidly taken the lead in terms of the impact of AI-generated material [1]. Because they grew up using IRC (Internet Relay Chat) rooms, a text-based instant messaging network, older millennial may be reminded of the experience of conversing online by the informal style of the conversations with the bot.

ChatGPT, the most recent technical innovation referred to as "large language model tools," does not "think" or interacts with awareness, however, unlike humans. The most advantageous paths to connect words are predicted by chatbots like GPT using a lot of data and powerful computer techniques. They can grasp words in context; have access to a huge vocabulary, and a wealth of information. They can replicate speech patterns and convey a great deal of knowledge. Along with providing answers to your inquiries, the language model may help you with things like composing essays, programming, and letters.

The ChatGPT users' crams are classified into individual's users, comparing ChatGPT user expansion with the other networks and web surfing information in ChatGPT. The explanations are as follows:

Individuals Users Using ChatGPT: One million new users joined ChatGPT in the first week after its launch. This created a record for the user programme with the quickest user growth ever. Additionally, as of February 2023, ChatGPT had 100 million daily users, according to The Guardian [2]. A lot of information concerning this has been made public since then.

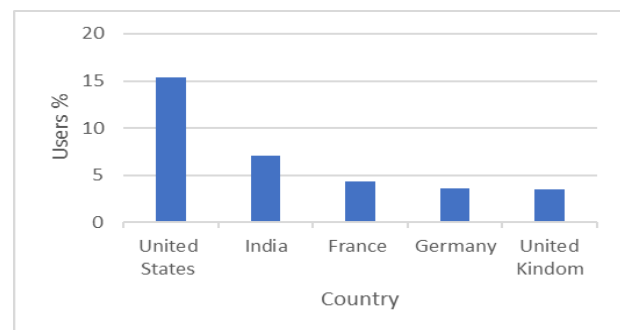


Figure 1. Percentage of ChatGPT Users

Figure 1 illustrates the estimate of website views, which does not always correspond to the number of active users, may cause confusion. According to more recent statistics, the website received an estimated 616 million visitors in the preceding month. Currently unknown is the total number of users who have successfully logged in. Our best estimate places the population at roughly 62 million (assuming 10% of website visitors sign up for the free programme). ChatGPT users come from all over the world, with the United States accounting for the majority (approximately 15%). According to data from the World Bank, India is considered to have the second-highest percentage of customers, making up around 7% of the total population. Comparing ChatGPT User Expansion with the Other Networks: Figure 2, shows how ChatGPT has grown considerably faster than other well-known networks. It took the network just five days to gain a million users, making it the second-fastest network behind Instagram [3].

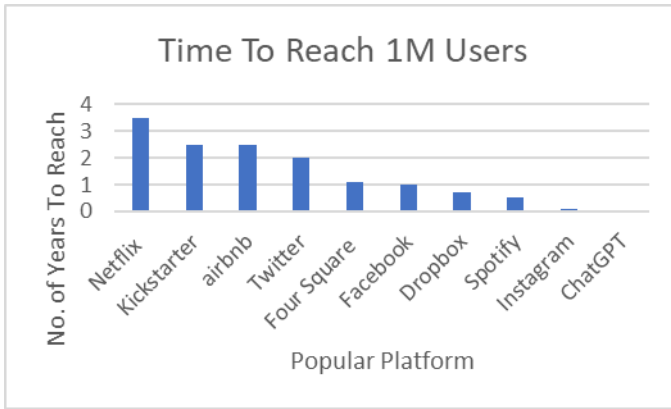


Figure 2. Time to Reach 1 Million Users

Web Surfing Information in ChatGPT: As shown in Figure 3, in January 2023, ChatGPT had an average of 13 million unique users per day, or more than twice as many as in December 2022. The site's traffic climbed by an average of 3.4% every day in the previous month. 5.85 pages were regularly seen during each 7 minutes, 23 second visit to the OpenAI website. The busiest day for traffic was January 31, when 28 million accesses to the website were made by 15.7 million unique visitors.

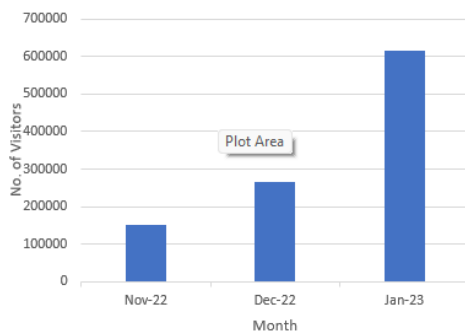


Figure 3. Website Visitor since Launch

II. EVOLUTION OF CHATBOT

Owing to advancements in artificial intelligence technology, chatbots are becoming more intelligent and nippier. The following are some of the developments in robot technology that technology magazine looks at:

- 1966 ELIZA: MIT scientist Joseph Weizenbaum created the first robot in the 1960s, called ELIZA. Technology has come a long way since then. The first chatbot, it used for pattern matching and substitution methodology to simulate dialogue, and was meant to emulate real interaction. Many early users felt that ELIZA was intelligent despite her incapacity to speak with true comprehension, despite Weizenbaum's protestations to the contrary. AI chatbot that can comprehend Standard English and respond to follow-up queries, admit mistakes, dispute unfounded claims, and decline unsuitable requests [5].
- 1992 Dr Sbaitso: One of the first chatbots to use AI was Dr. Sbaitso, which was created for MS-Dos. The programme, which was packaged with various Creative Technology sound cards, "conversed" with the user as though it were a psychotherapist and was designed to highlight the sounds the cards could produce.

- 1995 A.L.I.C.E: A.L.I.C.E. was a natural language processing robot that conversed with a person and was inspired by ELIZA. The program simulated chatting with a real individual over the internet and could inform a user its age, interests and other fascinating details, as well as addressing the user's dialogue.
- 2022 LaMDA- Google releases a mobile program that allows users to communicate with its Lamda AI. The business started enabling customers to obtain and sign up for the AI Test Kitchen App on either Android or Apple platforms using a Google account. It was first unveiled in May and enables users to communicate with LaMDA in a stream of test demonstrations.
- 2022 ChatGPT- OpenAI launched ChatGPT, an automaton with artificial intelligence. It has been enhanced (a technique of transfer learning) by using both supervised and reinforcement learning methods. The ChatGPT by OpenAI, a dialogue-based prototype. Figure 4, illustrates the evolution of ChatGPT.

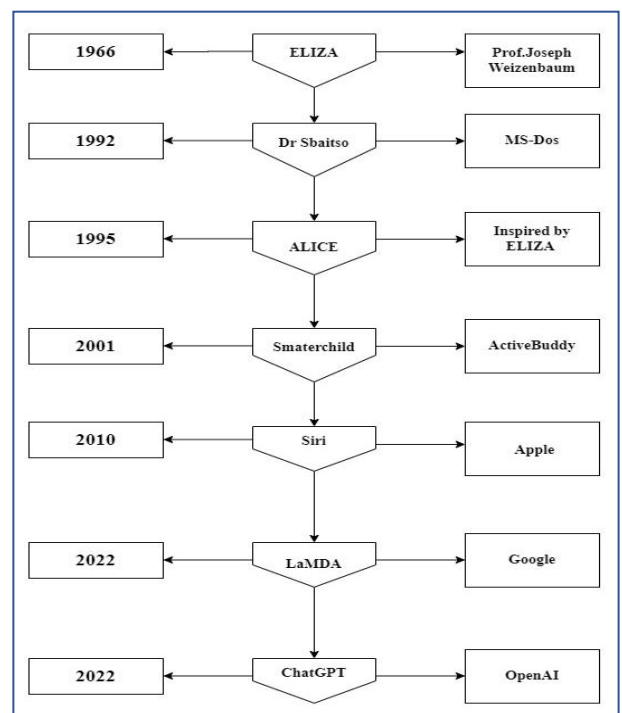


Figure 4. Evolution of Chatbot

- 2001 Smarterchild: Smarterchild, created by Active Buddy, could have amusing chats and fast data access to other services on AOL Instant Messenger and MSN Messenger. Over the length of its existence, it attracted over 30 million Instant Messaging "buddies" on AIM (AOL), MSN, and Yahoo Messenger.
- 2010 Siri: The intelligent personal assistant and learning guide Siri was created by Apple in 2010 for iOS, and its natural language user interface (NLU) served as the model for all subsequent AI apps and PAs. It cleared the way for the 2012 release of Google Assistant, as well as the 2014 releases of Microsoft's Cortana and Amazon's Alexa.

III. CHRONICLE ABOUT CHATGPT

In November 2022, OpenAI released ChatGPT, a robot with artificial intelligence. It has been enhanced (a transfer learning method) by integrating supervised and reinforcement learning techniques. The extensive language models from the OpenAI GPT-3 series serve as its foundation. ChatGPT was introduced as a prototype on November 30, 2022. It immediately gained a reputation for providing in-depth explanations to queries on a variety of subjects [6]. However, a key flaw in it has been recognised, including its uneven empirical veracity. After the introduction of ChatGPT in 2023, estimates put OpenAI's value at \$29 billion.

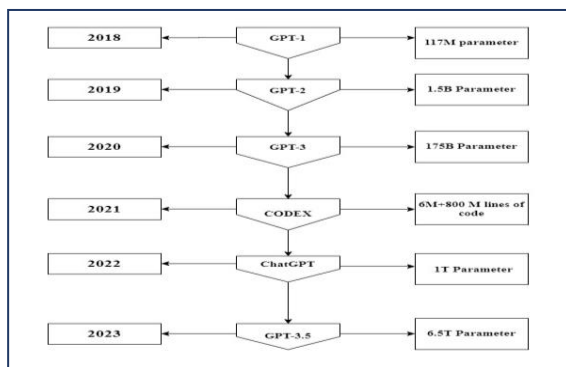


Figure 5. Evolution of ChatGPT

In order to learn more about the chatbot's creation, updates since its debut, and creators approach towards its success. Researchers are adopting an approach known as adversarial training (also known as jailbreaking) to stop users from deceiving ChatGPT into acting inappropriately. This work compares and contrasts a number of chatbots. In order to get

another chatbot to stray from its usual behaviour and send out undesired responses, one chatbot plays the antagonist and attacks the other by sending out text. The goal of adding successful assaults to ChatGPT training data is to teach it to disregard them. Additionally, Microsoft and Bain have a multimillion-dollar alliance and a global management consulting alliance with OpenAI, respectively [7]. Open AI's creative AI models will be used by Bain in marketing initiatives for its clients, which will include Coca-Cola. The excitement surrounding ChatGPT has ignited a new gold rush around large language models outside of OpenAI, with companies and investors from all over the world getting involved as shown in the Figure 5.

Social media was utilized by users to post examples of ChatGPT's various abilities, which vary from casual conversation to essay writing and coding, and in just five days, the chatbot had amassed over one million users. It was created by an artificial intelligence company that was formerly backed by some of the biggest names in Silicon Valley and is currently supported by Microsoft. The ground-breaking ChatGPT emerged from the GPT 3.5 model, one of the game-changers and cutting-edge in the field of natural language processing [8].

The GPT-1: Based on the encoder of the transformers proposed in the piece "Attention is All You Need," a spectacular model known as GPT-1 was developed. This model has 12-layer decoder layers and approximately 117 million parameters as shown in the Figure 6. It will be trained using a corpus of more than 40GB of text.

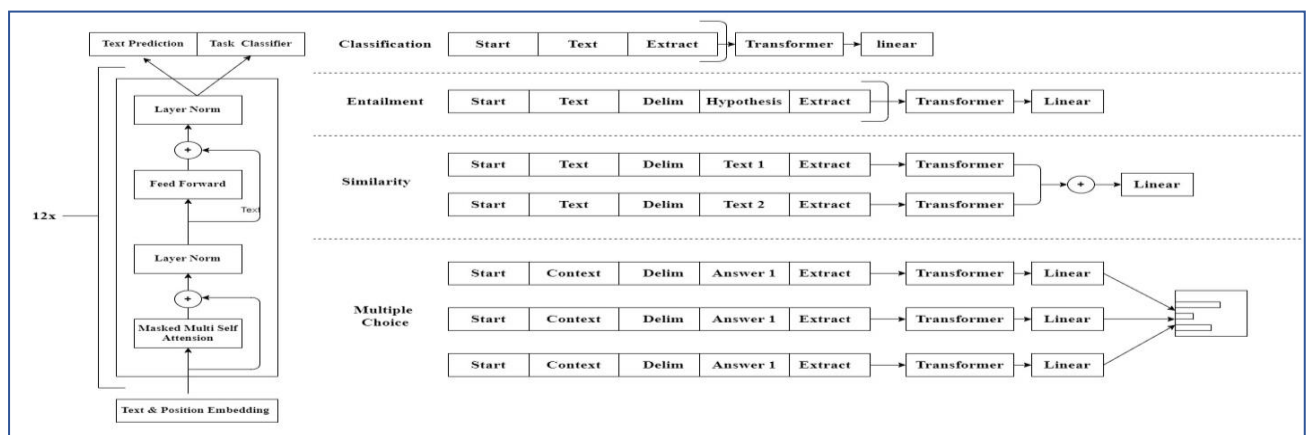


Figure 6. The GPT-1 Model

The evolutions of ChatGPT such as, GPT-1, GPT-2, GPT-3 and GPT 3.5 are discussed below:

The first phase in the GPT-1 process is training, followed by fine-tuning.

The GPT-2 Version: The GPT-2 model, which was published after the successful GPT-1 model was released and which is also based on the transformer's decoder design but has 48 levels and 1.5 billion parameters and was trained on 40 terabytes of text datasets from online sources, was an improvement on the original model as shown in the Figure 7.

Due to the amount of data, we can incorporate the task name into the GPT-2 input without engaging in an expensive and time-consuming fine-tuning method.

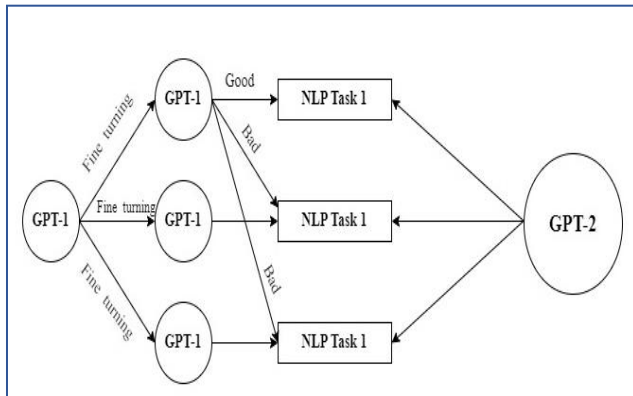


Figure 7. GPT version -2

The GPT-3 Version: The GPT-3 model was trained on 45 terabytes of text from diverse sources, including Wikipedia and books, and has 175 billion parameters, which is nearly 100 times more than GPT-2. The three types of learning used in GPT-3 are:

- Zero-shot learning: The algorithm is able to forecast the result using only the assignment label and a single "zero" sample.
- One-shot learning: In addition to the task's title and description, the model may predict the outcome by being provided "one" sample.
- Small-scale learning: The job description and "a few" examples will be presented to the model.

The GPT-3.5 Series: GPT-3.5, also known as InstructGPT, is based on GPT-3 but uses specific human values policies and has 1.3 billion fewer total parameters than the previous version. The same datasets used for GPT-3 were used to create the model, but it underwent additional modifications to incorporate Reinforcement Learning with Human Feedback (RLHF) into the GPT-3 model. The various difference of the GPT versions are shown in the below Table 1.

TABLE I: COMPARISON OF GPT'S

Parameters	Gpt-3.5	Gpt-3	Gpt-2
Year	2022	2020	2019
Model	Three	One	One
Parameters	(1.3b, 6b, And 175b)	175 billion	1.5 billion
Features	Modern Technology with Fewer Toxic Outputs.	Enhanced Responses.	Create Text That Is Human-Like.

IV. WORKING OF CHATGPT

A conversation engine and a language model are the two primary parts of ChatGPT. The language model is in charge of deciphering human input in natural language and producing replies. In order to comprehend the context of the discussion and produce pertinent answers, it makes use of deep learning algorithms developed by GPT-3. The person and the bot's communication are controlled by the dialogue system. It employs principles to decide how the bot ought to react to various human inputs. For instance, the conversation engine will decide what kind of answer the bot should provide if the user poses a query. We integrated this new discussion dataset with the InstructGPT dataset, which we turned into a conversation format. We need comparison data that contains at least two model replies ranked by quality in order to develop a reward model for reinforcement learning.

The Creation of ChatGPT: The three mechanisms listed below were used to build and improve ChatGPT, a large language model built on Generative Pretraining Transformers-3 and 3.5:

Step 1: Large-Scale Pretraining

Pretraining, which is a collection of text or phrases the chatbot already possesses, helps it better understand the grammatical structures and recurring patterns in real language [9]. As a result, the responses are conversational since the AI tool adapts to the context of the inquiry.

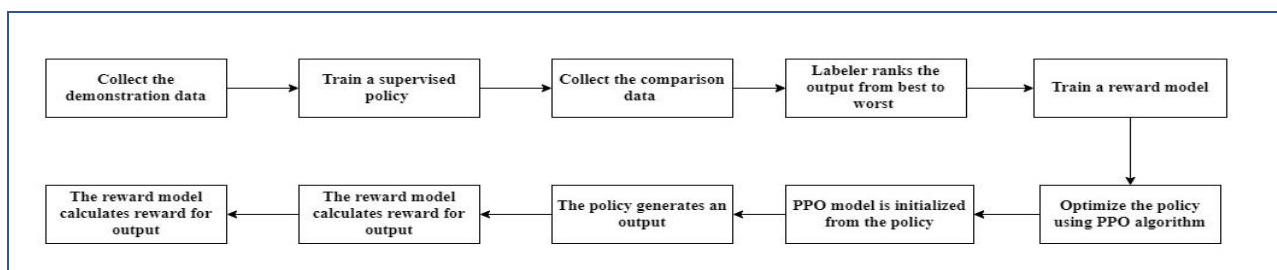


Figure 8. Working of ChatGPT

Step 2: Transformer Architecture

The transformer structure is made up of layers that may rank the words and phrases in user inputs in order to effectively

manage extensive data sequences. The model then makes an effort to better understand the context and meaning of the information in order to provide users with more logical solutions. Additionally, the feed-forward layers and residual links in the design aid the tool in understanding complex linguistic patterns, enhancing its capacity to record various words and sentences.

Step 3: Reinforcement Learning via Human Feedback (RLHF)
Since its inception, the robot has continuously developed and educated itself using Reinforcement Learning with Human Feedback (RLHF). Real AI teachers who can assume the roles of both a user and a chatbot have been hired by Open AI to help users interact with the model [10]. Trainers evaluate the quality of responses and contrast the tone and clarity of the answers produced by ChatGPT to human replies in order to promote human-like conversational techniques.

V. OVERVIEW OF CHATGPT

A Search Engine's Distinction from ChatGPT: Search engines organise web pages on the internet in order to assist users in finding the information they are looking for. Instead of providing the choice to perform an online search, ChatGPT [12] provides an answer based on the information it acquired from training data, which leaves potential for error. A key distinction is that Google and other traditional search engines have access to the most recent data, whereas ChatGPT only has access to data up to 2021. As a result, ChatGPT was unable to respond to your inquiry regarding the 2022 World Cup winner. However, Google was successful.

The Usecase of ChatGPT: ChatGPT for Coding: It helps users seem to concur that ChatGPT does coding exceptionally well, which they find unexpected. People have been playing with ChatGPT capacity to create website pages and write simple code since Day 1 [13]. In addition, ChatGPT appears to be very adept at answering coding-related queries and even act as a debugging aid when coders are having trouble.

ChatGPT for HR and Business Operations: A wide range of responsibilities are handled by HR and business administration professionals. And ChatGPT can deal with a lot of them.

ChatGPT for social media: Whether you're promoting a dating app, a software company, or an online store, ChatGPT is an excellent tool to help you build your social media profile. With minimal input, the chatbot can create posts for your social media accounts. The goal is to get more views, likes, and comments, even if the keywords are relevant.

It acts like a virtual cloud: OpenAI's chatGPT has found a way to assist with the rise in cloud service usage. All types of software are soon going to be bound together by language. An

underutilised use of the technology is the ability to integrate cloud services to finish difficult tasks.

The (possible) future of edtech: The impact of ChatGPT OpenAI on educational technology is likely to be significant [14]. Many edtech companies have a crucial use case where they merely provide the basics of a subject and provide a forum for students to ask questions. Essays, homework, and other types of evaluation are no longer acceptable. Although plagiarism has always existed, the speed, scope, and precision of this instance make it more dangerous. Teachers need to create novel ideas and assessment methods.

Solve the problem of unstructured data: In the era of data transformation, unstructured data poses a hassle. The problem is that they are challenging to manage, plan, and organise. ChatGPT offers alleviation since it can convert unstructured data into structured data.

Generate SQL Queries: Every data analyst should have SQL in their toolkit. Of course, learning it is crucial, but having a solid grasp of SQL will be very beneficial to you throughout your work.

Merits and Demerits of ChatGPT: Anyone can use it to create their own robot because it is open source. It is powered by GPT-3, one of the most effective language models currently available. It is extremely flexible and enables the building of programmes for a number of application scenarios [15]. It makes it possible for remote workers to quickly create AI chatbots without having to write complex code.

Merits:

- ChatGPT's primary purpose is to imitate human interaction based on queries or directives entered by the user.
- It simulates real-world conversation and is based on more advanced supervised learning and reinforcement learning approaches using large language models [16].
- The chatbot is functional.
- ChatGPT has unique advantages, such as significant savings in the cost and time required to develop a domain-specific model [17]. Its answers and overall performance may also be fine-tuned.

Demerits:

- **Inconsistency and Ambiguity** - The fact that ChatGPT occasionally has a propensity to produce phrases that seem to be true or convincing but are actually untrue or unintelligible is one of the most major critiques and shortcomings of the programme.
- Users should be made aware that ChatGPT has been the focus of an investigation [18] and that it uses outdated

datasets, which means it only has a partial comprehension of the facts. Many academic institutions have policies against using it.

- The Common Crawl dataset, which includes copyrighted materials from publishing companies as well as from individual writers and scholars, was used to build the GPT. Experts have also highlighted cautions over the potential application of AI-based systems in cybercrime [19].
- There may be legal fees and compliance requirements for ChatGPT and other derivatives.

Applications of ChatGPT: It is capable of writing in a style that is similar to that of expert artificial intelligence copywriters. Studies have revealed that it is even capable of composing music and creating fictional works like short stories. It can assist writers of technical material or content in creating an outline [20]. Another intriguing use of ChatGPT is that it can develop and debug. In addition, it can summaries, and explain the lengthy texts.

ChatGPT Vs Google Bard: Chatbot AI has taken over the news cycle, and two of the top rivals are Google Bard and ChatGPT, which are supported by Microsoft and owned by OpenAI. ChatGPT is still the finest because it was the first [21]. The OpenAI chatbot is credited with starting the current chatbot frenzy. It was first published as a research preview back in late 2022 and is capable of completing a number of activities, including composing essays, website scripting, and many more things we're still learning it can do.

Since then, a competition between Google and Microsoft to create chatbot AI weapons has been launched. We witnessed the launching of competing ChatGPT rivals on both sides over the course of a week. Microsoft was the company that originally presented the "new Bing" search engine [22]. The difference between the Google Bard and ChatGPT are shown in the Table 2.

TABLE II. COMPARISON ON CHATGPT AND GOOGLE BARD

Parameters	ChatGPT	Google Bard
Application	Chatbot	Chatbot
Revealed year	November 2022	February 2023
Model	Language Generation Model	Text Completion Model
Technology	NLP+DL	NLP+ML
Parent Company	OpenAI	Alphabet Inc
Accessibility Mode	Offline	Online
Search engine	About to bring with Bing	Chrome with bard
Incorporation	Microsoft Office	Google's cloud computing
Training Data	Large dataset of human generated text.	Data source like web pages, images and user data.
Capabilities	Text Generation, Translation and summarisation.	Search and retrieval task
Personalization	Response based on the user's base preference and past interaction.	Provides the same results for everyone searching for the same keyword.
Creativity	Generate creative and unique responses based on the given prompts.	Provides factual answers based on existing information.
Input	Text	Query
Output	Text	List of relevant websites
Availability	Available through web browser	Limited to trusted users
Existing Model	GPT-3	Lamda
Architecture	Transformer architecture	LaMDA Architecture
Merits	Mimics real-life conversation, versatile, overall performance can be fine-tuned.	Ability to make complex subjects simple.
Limitations	Accuracy and reliability of its response	Access vast amount of data and making result more comprehensive.
Applications	Customer service, personal assistance, and education research and development.	Creative Writing, Personal AI assistant, Automated Tasks Consumer and Business

Quantum ChatGPT: Exploring the convergence of quantum computing and natural language processing (NLP) is made more interesting by integrating ChatGPT with Quantum AI. The intrinsic parallelism and exponential processing capability of quantum computing could greatly increase ChatGPT's processing capacity. Better response times and enhanced model performance in managing intricate linguistic tasks should result from more effective and quick processing of language data made possible by quantum algorithms.

VI. CONCLUSION

The technology underlying ChatGPT, a highly developed chatbot that has recently attracted a lot of interest, was carefully examined in this paper. The background, creation, functionality, and underlying use case of ChatGPT were the subject of the paper subsequent section. It also examined ChatGPT's advantages, including how it may enhance cataloguing, metadata development, reference and information services, and content production. The paper has shown how ChatGPT can greatly promote both uncomfortable and novel scholarly research and librarianship. It is crucial to consider how to use this technology in an ethical and responsible manner as well as how society as a whole may use it to advance academic knowledge and train the next generation of professionals.

REFERENCES

1. <https://en.wikipedia.org/wiki/ChatGPT>
2. <https://www.digitaltrends.com/computing/how-to-use-openai-chatgpt-text-generation-chatbot/>
3. <https://www.thefountaininstitute.com/blog/chat-gpt-ux-design>
4. <https://blog.invgate.com/chatgpt-statistics>
5. <https://seo.ai/blog/chatgpt-user-statistics-facts>
6. <https://www.gadgetsnow.com/faqs/google-bard-the-new-ai-chatbot-what-is-bard-what-are-its-benefits-how-does-it-work-and-more/articleshow/97721888.cms>
7. <https://www.tooltester.com/en/blog/chatgpt-statistics/>
8. <https://technologymagazine.com/articles/from-eliza-to-chatgpt-the-evolution-of-chatbots-technology>
9. <https://iq.opengenus.org/gpt-3-5-model/>
10. <https://www.technologyreview.com/2023/03/03/1069311/inside-story-oral-history-how-chatgpt-built-openai/>
11. <https://medium.com/illumination/the-story-of-chatgpt-and-openai-the-evolution-of-gpt-models-abf201316a9>
12. <https://reintech.io/blog/understanding-chatgpt-architecture>
13. <https://ccm.net/apps/sites/apps/10619-what-is-google-bard-chatgpt-rival-ai-powered-conversation-tool/#when-is-bards-release-date>
14. <https://www.zdnet.com/article/what-is-chatgpt-and-why-does-it-matter-heres-everything-you-need-to-know/>
15. <https://emeritus.org/blog/ai-and-ml-chatgpt-use-cases/>
16. <https://analyticsindiamag.com/these-8-potential-use-cases-of-chatgpt-will-blow-your-mind/>
17. <https://www.cometchat.com/blog/chatgpt-use-cases>
18. <https://www.profolus.com/topics/advantages-and-disadvantages-of-chatgpt/>
19. <https://www.quora.com/What-are-the-differences-between-Chatgpt-and-Google-Bard>
20. <https://www.edureka.co/blog/chatgpt-vs-google-bard>
21. <https://www.tomsguide.com/face-off/chatgpt-vs-google-bard>
22. https://in.images.search.yahoo.com/search/images;_ylt=AwrKCbL8BgdkYqENDY67HAX.;_ylu=Y29sbwNzZzMecG9zAzEEdnRpZAMEc2VjA3BpdnM-?p=history+of+chatgpt&fr2=piv-web&type=E211IN885G91759&fr=mcafee#id=39&iurl=https%3A%2F%2F149520306.v2.pressablecdn.com%2Fwp-content%2Fuploads%2F2023%2F02%2FChatGPT-2.png&action=click