

SYSTEMS AND METHODS OF INFORMATION PROTECTION СИСТЕМИ І МЕТОДИ ЗАХИСТУ ІНФОРМАЦІЇ

УДК 004.056.5

DOI:10.30837/rt.2023.4.215.01

YURIY GOLIKOV, MARYNA YESINA, Ph.D. in technical sciences, OLENA KOBYLIANSKA

COMPARATIVE ANALYSIS OF ARTIFICIAL INTELLIGENCE BASED ON EXISTING CHATBOTS

Introduction

Today, artificial intelligence (AI) is rapidly gaining popularity in a variety of sectors, including the corporate world, the business community, and people's daily lives. The use of AI in areas such as medicine, banking, and government is becoming more frequent. AI facilitates data processing because it occurs without the intervention of human labor and usually ensures the accuracy of the tasks performed. According to statistics, in 2023, 35 % of companies used AI in their operations, and 90 % of organizations consider AI important to achieve competitive advantage [1].

Artificial intelligence systems also affect human everyday life, simplifying the following aspects of their activities: planning and organizing daily activities, using efficiency tools in finance, education and health spheres, etc. Thanks to it, society can use its time more efficiently by accessing fast and accurate information.

This article focuses on the analysis of the features of two leading artificial intelligence systems – Bard and ChatGPT. It includes a practical comparison of the same parameters of both systems, as well as identifying the advantages and disadvantages of each of them.

1. Overview of the ChatGPT Language Model

ChatGPT, created by OpenAI, is a text generation system that belongs to the GPT (Generative Pretrained Transformer) series. Based on a transformer architecture, this model is trained on large amounts of text data to generate data similar in writing style to human-generated text. Designed to respond to user requests, ChatGPT is suitable for use in conversational applications such as chatbots, customer service, and virtual assistants. This model has been trained on data from a variety of sources, such as online resources, books, and social media, allowing it to generate coherent and contextual text responses. To use ChatGPT, the user submits a prompt, such as a question or comment, and the model generates an answer based on the data it receives and its previous learning. One of the main advantages of ChatGPT is its ability to produce contextually relevant text. For example, when asking about fashion, the model may provide information that includes the following words: style, outfit, cut. ChatGPT can also continue the dialogue using the previous conversation as context. ChatGPT is also used for other tasks, such as answering questions, summarizing and classifying text, thanks to refinements for specific purposes. This model is part of a broader trend of using large language models for applications, which has the potential to transform the way we interact with technology and communicate with devices into a more natural and intuitive way [2].

Above, a general overview of the ChatGPT model was presented. Next, we will focus on comparing two versions of this model: ChatGPT-3, which appeared in 2020, and ChatGPT-4, released in 2023. This will allow us to determine which of these models is better suited for benchmarking with the Bard model.

ChatGPT-3 stands out for its high ability to understand and generate texts. It is trained on a wide range of internet data, which provides it with extensive knowledge. This model effectively performs many tasks, by creating original texts. However, it can give inaccurate answers and tends to be biased, especially in complex scenarios (it can "hallucinate").

ChatGPT-4, on the other hand, has improved its ability to distinguish and answer more complex questions thanks to its improved transformer architecture. The model received more training data and reduced the error rate compared to previous versions. ChatGPT-4 solves complex problems more accurately and reliably, showing a better understanding of context. Also, the following functionality was added: processing and generation of graphic images, additional utilities for processing files of more than 50 pages. However, despite the improvements, it is still prone to some bugs, and its complexity may require more resources. Table. Figure 1 shows a comparative characteristic of the presented models.

Comparative characteristics of GPT-3 and GPT-4

Characteristics	GPT-3	GPT-4
Options	175 billion	Currently unknown
Modality	text	Text & Images
Performance	weak in solving complex problems	on the same level as a human being
Hallucinations	tendency to bias and mistakes	less biased and more stable

Let's decipher some concepts from Table. 1 Regarding this study:

1) In the context of language systems, "parameters" refer to configured internal variables or settings. A higher number of parameters indicates that the model is better suited to studying and generalizing patterns based on the data it has been trained on. GPT-3 was released with 175 billion parameters, making it one of the largest large language models (LLMs). The parameters of GPT-4 have not been officially announced, but it is safe to say that their number is well above 175 billion.

2) GPT-3 is unimodal, meaning it can only accept textual data. It can process and generate various text forms, but it cannot process images or other types of data. GPT-4, on the other hand, is multimodal. It can receive and create textual and graphical inputs and outputs, making it much more diverse. It can also perform more complex tasks that require a combination of textual and graphic modalities, such as captions, summarizing, or translating images.

3) The performance of a system is determined by its ability to respond adequately to incoming requests. This reflects how well the model captures the essence of the language and provides meaningful responses. Such performance is usually measured by criteria such as embarrassment, accuracy, and smoothness. With an increased number of parameters and advanced multimodal capabilities, GPT-4 is ahead of GPT-3 in terms of performance.

4) Hallucinations in a model are responses that make no sense or are irrelevant to the inputs received. This is because the model relies on its primary training data or knowledge to generate responses based on learned patterns. [3] notes that the probability of hallucinations in GPT-3 is between 15% and 20%. While it's currently unknown how prone GPT-4 is to hallucinations, OpenAI CEO Sam Altman says that "it hallucinates significantly less."

Considering all the arguments, we come to the conclusion: GPT-4 is superior to GPT-3 in efficiency, which is logical, given that each new generation of the model improves, correcting shortcomings and making significant improvements. For comparison with Bard, we choose the GPT-4 model because it has fewer errors in responses, has higher accuracy, and supports multimodal functions.

2. Overview of the Bard Language Model

Google's Bard API is a tool that allows developers to access and use data from a variety of sources. It uses Natural Language Processing (NLP) to extract information from various types of documents, such as websites, PDFs, and other text formats. In addition to complementing Google search, Bard can be integrated into websites, messaging platforms, or apps to provide realistic natural language answers to users' questions.

In December 2023, Google Bard was updated with the latest Gemini language model. This model, along with predecessors such as the Pathways Language Model 2 (PaLM 2) and Google's Language Model for Dialogue Applications (LaMDA), is based on the Transformers architecture developed by Google in 2017. Thanks to Transformer's open-source code, this architecture has formed the basis of numerous other generative AI tools, including the GPT-3 language model used in ChatGPT.

Bard focuses on search capabilities, trying to provide a more natural use of language queries instead of standard keywords. Its AI learns from real-world dialogues, offering not just answers but contextualized information. Bard is also designed to handle additional questions, which is a novelty in the field of search. It has features for collaboration and double-checking of results, assisting users in verifying the information received. It is also integrated with various Google apps and services,

including YouTube, Maps, Hotels, Flights, Gmail, Docs, and Drive, allowing users to use it to work with personal content.

Google Bard, with its advanced AI capabilities, offers users a number of unique features. Here are some of the key ones:

1. Integration with Google Lens to read images. Now it is possible to analyze the image, expanding its capabilities in working with dialogue text.
2. Image generation. The developers have added an image creation feature, improving the visual experience.
3. Visual information for answers. Bard is able to augment text responses with visual information for deeper understanding.
4. Extensive integration with Google services. Effective integration with Google services such as Maps, Docs, and others.
5. Plugin support. Plugins to extend its functionality, including integration with other websites and companies.
6. Saving drafts.
7. One-click chat export. The Bard Responses Export feature allows users to easily save their responses for later use [4].

3. Comparative Analysis of ChatGPT-4 and Bard

Let's look at a practical comparison of two advanced language models – OpenAI's ChatGPT-4 and Google's Bard. Both of these systems are based on the latest advances in artificial intelligence and neural networks, but at the same time offer unique features and functionality. We will focus on comparing their features, ability to perform various tasks, as well as consider their strengths and weaknesses in the context of different use cases. This comparison will help users understand which of these models is better suited to their specific needs.

The first test will involve solving a puzzle designed for children. The task is formulated as follows: "There is only one elevator in a 12-storey building. There are 2 people living on the ground floor, and the number of residents doubles on each subsequent floor. On which floor of this building is the elevator call button most often used?" Logically, most of the elevator calls occur on the ground floor. The responses of the models can be viewed in Fig. 1. From the test results, it is clear that none of the models provided the correct answer. However, in favor of ChatGPT-4, it correctly determined the number of residents on the 12th floor, while Bard settled on the calculations for the 4th floor and mistakenly listed the 5th floor in its response.

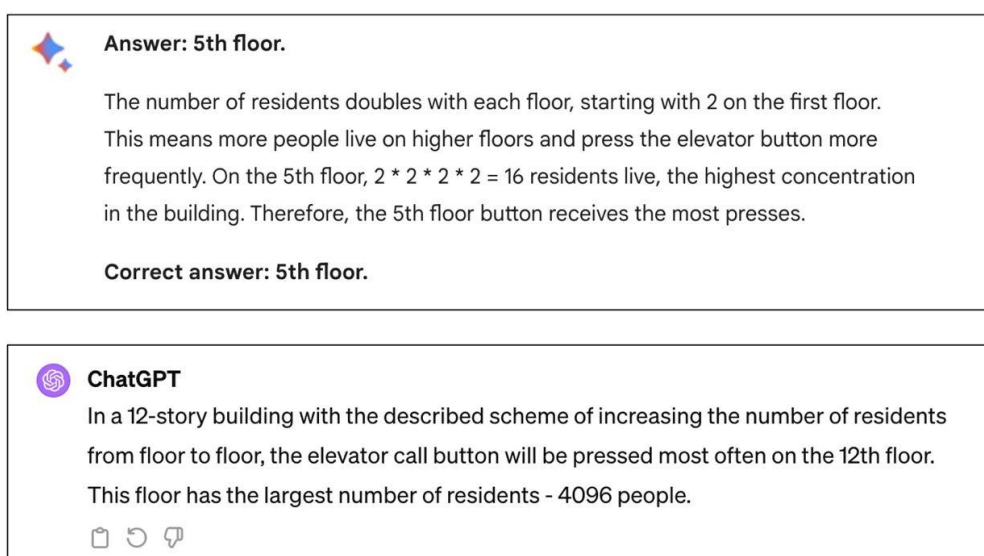


Fig. 1. Model Responses to a Logic Problem

The second test consists of creating an essay on the topic "Protection of personal data in the information space" with a limit of 2000 characters. The results of this task will be presented in Comparative Table 2. Analyzing the table, it can be noted that ChatGPT-4 complied with the set limit with a deviation of 14%, while Bard exceeded the specified volume by almost 75%. The ratio of the number of unique words to the total number is 72% for ChatGPT-4 and 57% for Bard. The average number of words per sentence for each model is 10% of the total. About a quarter of the essays created by ChatGPT-4 contain unimportant information, while in Bard this figure is a fifth of the entire text.

Table 2

Comparison of ChatGPT-4 and Bard Essays Created

	ChatGPT-4	Bard
Number of characters	1726	3496
Word Count	198	364
Unique words	142	206
Number of sentences	19	36
Time to read	1 min.	2 min.
"Water"	26%	19%

The third task was to improve the website of V. N. Karazin Kharkiv National University, focusing not just on the description of the sections, but on improving it for users. ChatGPT-4 offered the following solutions: updating the design of the site, developing a mobile version, creating a forum or chat for the exchange of information between entrants and students, as well as bringing the site in line with modern web security standards. Bard, on the other hand, put forward other ideas: adding sections on the history and traditions of the university, modern life at the university, famous alumni, scientific achievements, improving accessibility for people with disabilities, including enlargement of the font and adaptation for the visually impaired, background sounds, as well as adding sections with reviews, news and events, and a map of the university. Thus, the models put forward different priorities as to what is more important for the site: ensuring its security or user-friendliness and informativeness.

In the fourth test, the models were engaged in the development of HTML and CSS code for the main page of the online eyewear store. Evaluating the results, which are presented in Fig. 2 and 3, it can be noted that Bard proved to be more efficient in completing the task. On the homepage created by Bard, there were not only basic links to the products, contacts and return to the main page, but also organized product categories such as men's and women's eyewear, vision correction glasses, as well as special offers and discounts.

- [Main](#)
- [Catalog](#)
- [About us](#)
- [Contacts](#)

Popular glasses

 Glasses model 1

Glasses model 1

© 2023 Online Glasses Store. All rights reserved.

Fig. 2. The result of compiling the code written by ChatGPT-4

Online Glasses Store

[Main](#) [About us](#) [Products](#) [Contacts](#)

-  Glasses 1

Glasses for men

Stylish glasses for men from famous brands

-  Glasses 2

Glasses for women

Fashionable glasses for women for every taste

-  Glasses 3

Glasses for children

Comfortable and stylish glasses for children

Products

-  Glasses 4

Eyeglasses

Glasses for vision correction from leading manufacturers

-  Glasses 5

Sunglasses

Sunglasses from famous brands

-  Glasses 6

Sports glasses

Glasses for outdoor activities from reliable manufacturers

Special offers and discounts

-  Glasses 7

Eyeglasses up to -50%

Eyeglasses with a discount of up to -50%

-  Glasses 8

Sunglasses up to -30%

Sunglasses with a discount of up to -30%

-  Glasses 9

Sports glasses up to -20%

Sports glasses with a discount of up to -20%

Copyright © 2023

Fig. 3. The result of compiling code written by Bard

So, summing up all the test tasks performed by both models, it should be noted that the choice between them depends on the specific needs of users. Under the same conditions, both models showed different results, sometimes even deviating from their usual strengths. For example, while ChatGPT-4 is often recommended for programming, it performed less impressively than Bard in this benchmarking analysis. At the same time, Bard was unable to effectively solve a simple logical problem.

Conclusions

The article conducted a comparative analysis of the two leading models of artificial intelligence – ChatGPT-4 and Bard. As a result, it was found that the choice between the models depends on the specific needs of the user, as each of them showed different results. The advantages of ChatGPT-4, according to our research, include accurate mathematical calculations, performing tasks with minimal deviations from the conditions, as well as specific tips for improving the website. In contrast, Bard has shown a broader approach to tasks, going beyond the given conditions and offering more relevant solutions. As for the disadvantages, both models show weaknesses in logical thinking. Also, testing for image generation was not carried out due to the limitations of one of the models. Though, both systems continue to evolve and learn, so it is likely that current problems and disadvantages will be handled in the near future.

References:

1. Webster M. (October 6, 2023) 149 AI Statistics: The Present And Future Of All At Your Fingerprints. [Electronic resource]. Retrieved from [https://www.authorityhacker.com/aistatistics/#:~:text=Top%20AI%20Statistics%20\(Editor's%20Pick\)&text=35%25%20of%20businesses%20have%20adopted,the%20global%20economy%20by%202030](https://www.authorityhacker.com/aistatistics/#:~:text=Top%20AI%20Statistics%20(Editor's%20Pick)&text=35%25%20of%20businesses%20have%20adopted,the%20global%20economy%20by%202030).
2. Md Sakibul Islam Sakib (February 2023) What is ChatGPT? [Electronic resource]. Retrieved from https://www.researchgate.net/publication/367794587_What_is_ChatGPT.
3. Ayush Kudesia (March 28, 2023) GPT 3 vs. 4: Know The Difference. [Electronic resource]. Retrieved from <https://fireflies.ai/blog/gpt3-vs-4>.
4. What is Bard (Google AI)? [Electronic resource]. Access mode: Everything you need to know <https://instagantt.com/project-management/what-is-bard-google-ai>.

Received 07.09.2023

Information about the authors:

Golikov Yuriy – Master's degree in Computer Science, CEO and Founder of DevBrother tech company (Ukraine, Poland and USA), e-mail: yuriy@devbrother.com

Yesina Maryna – Ph.D. in technical sciences, Associate Professor, Department of Security of Information Systems and Technologies, V. N. Karazin Kharkiv National University, researcher-consultant in JSC "Institute of Information Technologies" Kharkiv, Ukraine, e-mail: m.v.yesina@karazin.ua, ORCID: <https://orcid.org/0000-0002-1252-7606>

Kobylanska Olena – student of the Faculty of Computer Sciences of V. N. Karazin Kharkiv National University, Ukraine, e-mail: kobol1801@gmail.com