

A Brief Analysis of the Architecture, Limitations, and Impacts of ChatGPT

Noah M. Kenney

March 23, 2023

Abstract

First, we begin with a description of the technical architecture of ChatGPT and how it differs from other large scale artificial intelligence language models. From there, we define the current use cases of ChatGPT, followed by an analysis of some of the current limitations of ChatGPT. In particular, we look at the inability of ChatGPT to be creative, issues of perpetuating biases, and possibility of identification. Finally, we look at some of the likely key impacts of ChatGPT, including copyright considerations and economic ramifications. As much as possible, this analysis is done from a non-technical standpoint, and seeks to show how AI is beginning to connect with daily life.

1 ChatGPT Technical Architecture

1.1 Parameters

Like most artificial intelligence models, ChatGPT utilizes parameters, which are "variables in an AI system whose values are adjusted during training to establish how input data gets transformed into the desired output" [1]. ChatGPT-3 uses 175 billion parameters [2], a massive upgrade from the 117 million parameters used in the first version Open AI released [3]. The sheer quantity of parameters is one reason why ChatGPT has been so effective. To understand why so many parameters were necessary, it is necessary to understand the intended use of ChatGPT.

ChatGPT is designed to fall into the category of Artificial General Intelligence (AGI), meaning that is designed to mimic human cognitive abilities [4]. Unlike specialized artificial intelligence (AI), which is designed to accomplish one relatively predictable task repeatedly, AGI is designed to either replace or augment human cognitive processing and output. This can only be accomplished if the AI is reasonably capable of answering questions and generating content related to a variety of topics with a high degree of accuracy. Thus, the AI relies on the use of a large parameter count.

1.2 Reinforcement Learning

ChatGPT utilizes reinforcement learning, which is designed to improve the accuracy of the artificial intelligence application over time as the application is used by more people. It accomplishes this by providing feedback to the model, either positive or negative. For example, if a user asks a question and then asks the same question reworded slightly upon receiving an answer from the AI, we can infer that perhaps the AI did not correctly answer the question the first time. In this case, we would provide feedback to the AI that it incorrectly responded to the user's question, with the goal of increasing the model accuracy if asked in the future. Theoretically, this allows an AI tool to learn through trial and error, given a long enough duration of time with enough usage by humans [5].

1.3 Supervised Learning

The use of supervised learning allows for answers that are not only accurate, but also precise. Accuracy and precision are two metrics that are commonly charted on either some form of regression model or circular graphical representation. AI models are generally either extremely accurate and not precise or extremely precise and not accurate. Accuracy in artificial intelligence could be defined as producing factually true or correct statements, while precision could be defined as closely answering a user's input. Building an accurate (and not precise) AI model is much easier than developing a precise AI model. In this case, answers to specific questions can be hardcoded in the database used by the AI. Thus, when a user asks a question, the AI model attempts to match the user input to a known question in the database, and displays the corresponding answer with minimal to no deviation. The result is a fairly generic answer that may not closely match the user's question. This form of AI is also very slow computationally, given that the database record count is so high when dealing with AGI. Thus, this form of AI architecture is beneficial primarily in cases of specialized AI.

In the case of AGI (such as ChatGPT), precision and accuracy are equally important. Additionally, massive datasets are not an option, given that the AI would be too computationally slow to be usable for everyday consumers. This is why supervised learning is used by ChatGPT. Supervised learning uses weighted input data, cross validation processes, and labeled datasets for algorithmic training, classifying, and prediction of data outcomes [6].

The primary purpose of supervised learning in AGI is classification, or the breaking down of data into categories. There are several reasons why this might be necessary, or at least beneficial. In the case of AGI, faster computational processing can be achieved in part by utilizing segmented AI, which requires scanning the user input twice. The first time, the AI scans the user input for keywords that help it determine which segment of the database

to match the user input against. The second time, the AI scans the input against the segmented portion of the database to match the user input to a programmed question, similar to the process utilized by a specialized AI tool. Supervised learning assists in the first of these two scans, ensuring that the user input is properly matched to a category or database segment.

One additional use of supervised learning in AGI is regression, which can be used to make predictions based on the relationship between dependent and independent variables [6]. In these cases, the independent variables may be the data fields collected about the user of a particular application.

1.4 Training Data

Although not directly related to the technical architecture of ChatGPT, it is worth noting the amount of data used in the training of the AI model. By volume, ChatGPT-3 was trained with approximately 570GB of datasets [7]. The amount of the training data has also gone up with each version of ChatGPT released. ChatGPT-1 utilized a 5GB dataset, while ChatGPT-2 utilized a 40GB dataset [7]. The data types used in the dataset include websites, books, and articles, among others. ChatGPT-3 was trained with 300 billion words in total [2].

1.5 Technology Summary

ChatGPT utilizes 175 billion parameters, reinforcement learning, supervised learning, and a training dataset of 300 billion words (570GB) to effectively answer user queries through the use of predictive text. While most large scale language learning models use input to database matching algorithms, while ChatGPT's approach relies instead on significant training data and human reinforcement to generate text in a predictive manner that (theoretically) allows it to answer questions quicker due to the reduction in computational processing power.

2 Current Use Cases

ChatGPT is a large scale language learning models and is used for a variety of purposes, primarily involving text based responses. While not an inclusive list, below are a few of the primary use cases.

2.1 Coding

ChatGPT is capable of writing code in multiple coding languages, including Python, JavaScript, C++, C#, Java, Ruby, PHP, Go, Swift, TypeScript, SQL, and Shell [8]. While

the code written by ChatGPT may not always be optimized fully, it is capable of writing reasonably good code very quickly and accurately [9].

2.2 Conversational AI

ChatGPT presents responses to user queries in a conversational manner, allowing it to answer questions, provide information, and summarize content. While ChatGPT is not classified as a conversational AI model (given that it is a language model instead), it is capable of providing human-like responses and understanding natural language [10].

2.3 Content Generation

ChatGPT can be used to generate both short and long form text, including blog posts, articles, and essays. ChatGPT-4 expands the length of text supported, allowing the AI to write full length books as well. In addition, ChatGPT is capable of writing advertisements, poetry, and even songs.

3 Analysis of Limitations

3.1 Creativity

ChatGPT lacks the ability to be creative, at least according to the traditional definition of creativity. Oxford defines creativity as "the tendency to generate or recognize ideas, alternatives, or possibilities that may be useful in solving problems, communicating with others, and entertaining ourselves and others" [11]. Creativity in artificial intelligence, particularly AGI, is a unique challenge. To understand why AI can't truly be creative, we have to first understand that AI only operates within the parameters that it's been programmed to operate within. Creativity relates to the production of artistic work with an emphasis on the originality which, by definition, means that it must work outside of fixed parameters. For something to be original, it has to be unique, and that cannot happen in the context of fixed parameters. Parameters are useful for generating and outputting factual answers to complex problems and questions. ChatGPT excels in providing factual answers, given the wealth of knowledge and information it is programmed with. It is then able to distill down and articulate in response to user inputs, which it is highly effective at doing. However, it's only effective at doing this because it is operating within the parameters that it was programmed with. If we ask ChatGPT to operate outside of these programmed parameters, it is not able to do so. Thus, we could say that ChatGPT is not capable of being truly creative.

3.2 Lack of Personal Experience

While this limitation is not unique to ChatGPT, it is worth noting. Today, a significant portion of media utilizes personal stories and experiences, as opposed to strictly factual information. For example, we can read facts about a tragic event, but that will be very different than reading the personal experiences of an individual who lived through them. The facts may provide useful information, but it cannot provide the personal stories and experiences that a human can. We can even see the impacts of this on marketing, where a significant portion of marketing is based on personal stories or testimonials.

3.3 Inherent Bias

There is not enough data to provide a thorough analysis of the specific biases of ChatGPT. However, there is preliminary information and data showing ChatGPT discriminating on multiple grounds [12]. To an extent, this is expected since the developers of AI technologies are not a representative sample of the overall population. Historically, AI engineers are predominately male (over 90%) and predominantly white [13]. AI engineers also generally need some level of formal education [14], and earn higher salaries than the majority of the population [13]. Ultimately, the uneven distribution in demographics among the AI engineers training AI models, such as ChatGPT, will result in inherent bias.

3.4 Possibility of Identification

Although the research conducted about the possibility of identification of content written by ChatGPT is preliminary, there are several possible methods of identifying this content. First, ChatGPT is capable of identifying a significant portion of the content it wrote itself, assuming it is of reasonable length. By feeding the content back into ChatGPT and asking it whether it wrote the content, it is able to identify this content in most cases. Additionally, OpenAI is working on cryptographic watermarking that could assist in the identification of content written by ChatGPT [15]. This possibility of reidentification shows that ChatGPT is not capable of writing content that is truly indistinguishable from content written by a human. However, in many practical cases, the content will be similar enough for this limitation to be negligible.

4 Impacts of ChatGPT

4.1 Copyright Considerations

First, it is important to make a distinction between two types of content: factual content and unique content. Factual content could be defined as indisputable information that can be agreed upon by virtually everyone (i.e., where someone was born, the current president, the date of a historical event). Unique content would include the creative slogans used

in marketing campaigns, an insightful blog post, or an opinion about a matter of public policy. For the purposes of this analysis, I will focus on unique content, given that we generally don't need to cite sources for factual content.

From a copyright standpoint, we may identify several ethical concerns. One concern would be how someone could be prevented from profiting exorbitantly from this technology, given that ChatGPT is able to produce information at a rate not possible without the technology. Although ChatGPT changed from a non-profit to for-profit in 2019, there is a profit cap (100x original investment) that is designed to address this ethical concern [16]. Some individuals would argue that it is fair the creator of the technology owns the content it creates and can profit from it as he/she wishes. However, if we view AI (and ChatGPT) as merely a tool, this argument may seem unreasonable. For example, we don't cite the use of a calculator that helps us answer a complex math question. The calculator is a tool and we view it as such. In regard to AI, this is slightly different given that the content may be unique. However, we are still faced with the question of whether we should need to cite ChatGPT if it is truly just a tool.

The issue of content ownership also raises a key question: If ChatGPT and another AI platform independently create content that is remarkably similar, who owns the content? Ordinarily, we would say that the first person to write the content is the rightful owner of the content. However, in the case of ChatGPT and other large scale language learning models, we have no method of tracking when a particular piece of content was originally written. Additionally, a legal question is raised: Can the owner of an AI tool competing with ChatGPT be penalized for copyright infringement if their AI produces an extremely similar answer? If this question is answered in the affirmative, we are faced with an issue resulting from the fact that content produced and published on the web will now be written largely by AI. When we go to train AI models, it will be challenging if we are not able to freely train the model using content available on the web, given that it may have been written by another AI. This is especially true for ChatGPT, since it is trained on information available to the public.

4.2 Higher Education Considerations

In some cases, ChatGPT may be a restricted technology because of an assumption that it diminishes either the quality of content produced or the educational value of an assignment. This is especially true in institutions of higher education. One challenge is whether we consider the goal of higher education to be the production of a valuable worker or the production of an educated person. These are two distinct goals. ChatGPT, and similar applications, will be available and prevalent for the rest of the foreseeable future. Yet, some institutes of higher education will go to great lengths to prevent students from using this powerful technology. This is because their view is that the goal is for students who gradu-

ate from the institution to be intelligent, learned, and capable of expressing, articulating, and even creating knowledge. Others argue that this technology is a tool that institutions of higher education should teach to students so they can become productive members of the workforce. This is a challenging dilemma and one that will be handled differently at every institution.

Additionally, ensuring that ChatGPT is used ethically and responsibly in academic settings poses challenges. ChatGPT is a tool, and it can be misused like any tool. The institution can develop policies against using ChatGPT, and if students are caught breaking the policy, the institution can take disciplinary action against them. However, determining how this technology is being used at a particular institution poses additional challenges in of itself, and the speed of technological innovation may make it difficult for institutions to stay ahead of a rapidly advancing technology.

Regardless of these challenges, ChatGPT can also be a powerful learning tool by accelerating the speed at which students learn and solve problems. It can provide an always available virtual second cognitive thought partner for students to use to detect flaws in their thinking and writing. For example, students may feed an essay they wrote into ChatGPT and ask it to critique the writing for the purpose of improvement. This form of instant feedback has not previously been available, and could be immensely valuable.

4.3 Economic Considerations

The economic considerations of ChatGPT are some of the most discussed, and the most controversial. It is highly likely that ChatGPT and other similar applications will begin to both augment and replace human labor. Cases where ChatGPT would augment human labor could include positions that require extensive copywriting or content creation. For example, a digital marketer may use ChatGPT to generate ideas of social media topic ideas or to edit blog posts before publication. In these cases, we would expect almost no economic ramifications, and no change to unemployment rates. In other cases, it is possible that ChatGPT could begin to replace human workers, particularly in cases where the tasks being performed by the worker are predictable and repeatable.

One large fear of ChatGPT, and similar technologies, is that it will lead to mass unemployment and profound economic ramifications. However, this thought relies on a fundamental assumption that the quantity of money supplied will drastically decrease, which is not likely. Instead, it seems more reasonable to assume that, over time, all employees will work slightly fewer hours. This would ultimately lead to a higher wage per hour, keeping the supply of money the same and the impacts on Gross Domestic Product (GDP) minimal. Additionally, it is important to note that many new job positions would open as a result of innovation in AI.

5 References

1. “Number of parameters in notable Artificial Intelligence Systems,” Our World in Data. [Online]. Available: <https://ourworldindata.org/grapher/artificial-intelligence-parameter-count>. [Accessed: 22-Mar-2023].
2. A. Hughes, “CHATGPT: Everything you need to know about OpenAI’s GPT-4 tool,” ChatGPT: Everything you need to know about OpenAI’s GPT-4 upgrade — BBC Science Focus Magazine, 16-Mar-2023. [Online]. Available: <https://www.sciencefocus.com/future-technology/gpt-3/>. [Accessed: 22-Mar-2023].
3. “What is GPT-4.” [Online]. Available: <https://www.mlyearning.org/what-is-gpt-4/>. [Accessed: 23-Mar-2023].
4. B. Lutkevich, “What is Artificial General Intelligence?,” TechTarget Enterprise AI. [Online]. Available: <https://www.techtarget.com/searchenterpriseai/definition/artificial-general-intelligence-AGI>. [Accessed: 22-Mar-2023].
5. J. M. Carew, “What is reinforcement learning?,” TechTarget Enterprise AI. [Online]. Available: <https://www.techtarget.com/searchenterpriseai/definition/reinforcement-learning>. [Accessed: 22-Mar-2023].
6. “What is supervised learning?,” IBM. [Online]. Available: <https://www.ibm.com/topics/supervised-learning>. [Accessed: 22-Mar-2023].
7. B. Gratas, “50 CHATGPT statistics and Facts you need to know,” InvGate, 14-Feb-2023. [Online]. Available: <https://blog.invgate.com/chatgpt-statistics>. [Accessed: 22-Mar-2023].
8. A. Christensen, “How many languages does chatgpt support? the complete chatgpt language list,” SEO AI, 03-Feb-2023. [Online]. Available: <https://seo.ai/blog/how-many-languages-does-chatgpt-support>. [Accessed: 22-Mar-2023].
9. U. Abdullah, “Chat GPT can write code - here’s why that’s significant,” PC Guide, 03-Mar-2023. [Online]. Available: <https://www.pcguides.com/apps/chat-gpt-can-write-code/>. [Accessed: 22-Mar-2023].
10. U. Abdullah, “Is chat GPT a language model or a conversational ai?,” PC Guide, 07-Mar-2023. [Online]. Available: <https://www.pcguides.com/apps/chat-gpt-a-language-model-conversational-ai/>. [Accessed: 22-Mar-2023].
11. “What is creativity?,” California State University, Northridge. [Online]. Available:

<https://www.csun.edu/vcpsy00h/creativity/define.htm>. [Accessed: 23-Mar-2023].

12. H. Getahun, “CHATGPT could be used for good, but like many other AI models, it’s rife with racist and discriminatory bias,” Insider, 16-Jan-2023. [Online]. Available: <https://www.insider.com/chatgpt-is-like-many-other-ai-models-rife-with-bias-2023-1>. [Accessed: 23-Mar-2023].

13. “ARTIFICIAL INTELLIGENCE SPECIALIST DEMOGRAPHICS AND STATISTICS IN THE US,” Zippia. [Online]. Available: <https://www.zippia.com/artificial-intelligence-specialist-jobs/demographics/>. [Accessed: 23-Mar-2023].

14. M. Banoula, “How to Become an Artificial Intelligence Engineer: Roles, and More,” Simplilearn, 13-Feb-2023. [Online]. Available: <https://www.simplilearn.com/tutorials/artificial-intelligence-tutorial/how-to-become-an-ai-engineer>. [Accessed: 23-Mar-2023].

15. R. Montti, “ChatGPT for content and SE?,” Search Engine Journal, 12-Dec-2022. [Online]. Available: <https://www.searchenginejournal.com/chatgpt-for-content-and-seo/473823/>. [Accessed: 23-Mar-2023].

16. D. Ruby, “ChatGPT Statistics for 2023 (New Data + GPT-4 Facts),” Demand Sage, 18-Mar-2023. [Online]. Available: <https://www.demandsage.com/chatgpt-statistics/>. [Accessed: 23-Mar-2023].