The Ideals, Achievements, and Controversies of OpenAI: A Discussion on the Backdrop of ChatGPT

Jiacheng Xie

Abstract—This paper analyzes OpenAI and its accomplishments and controversies, with a focus on the development of the highly acclaimed chatbot ChatGPT. OpenAI was founded in 2015 as a non-profit research company with the objective of developing strong artificial intelligence that is safe and beneficial to humanity. OpenAI's founders included renowned investors such as Elon Musk, Peter Thiel, and Reid Hoffman. ChatGPT has been used to create a variety of applications, including writing articles, generating code, and translating languages. However, OpenAI has faced significant controversies, including accusations of being controlled by Microsoft and criticisms for becoming a closed system. The paper provides an in-depth analysis of ChatGPT and the company behind it to evaluate the credibility of its developments and understand the basis of the doubts it has faced.

Index Terms—OpenAI, Artificial General Intelligence, Controversies, Technology industry.

1 INTRODUCTION

The field of technology has witnessed the rise of big tech giants who have achieved remarkable success in terms of profitability. Science and technology graduates are often underestimated. They are seen as being more interested in theory than in practice, and as being less concerned with making money than with making a difference in the world. However, this perception is often inaccurate. Science and technology graduates are just as capable of making money as anyone else, and they are often very good at it, particularly in the realm of open-source technology and financial gain, as this has been controversial.

One notable example of this is the company OpenAI, a non-profit research company founded in 2015 by Elon Musk, Sam Altman, and others [5], which has developed the highly acclaimed chatbot ChatGPT. The technology behind the chatbot has received widespread recognition and minimal criticism. ChatGPT is still under development, but it has already been used to create a variety of impressive applications, such as writing articles, generating code, and translating languages. However, OpenAI has faced significant controversies over the past four years despite its remarkable achievements. Some have accused the company of being controlled by Microsoft, especially when OpenAI was acquired by Microsoft for $1 billion in 2019 [4], while others have criticized it for devolving into a closed system.

Therefore, it is essential to consider the accomplishments and controversies surrounding OpenAI to evaluate the credibility of its developments and understand the basis of the doubts it has faced. In this paper, I will provide an in-depth analysis of ChatGPT and the company behind it. I will explore the achievements of OpenAI, as well as the controversies that have surrounded it, in order to gain a better understanding of this highly influential company.

• *Jiacheng Xie is an undergraduate student of Duke Kunshan University, Jiangsu, China 215316.*
E-mail: jiacheng.xie052@dukekunshan.edu.cn

2 AGI

In December 2015, the establishment of OpenAI marked the advent of an idealistic venture. Its objective was AGI or Artificial General Intelligence [5], which was coined by Mark Gubrud as early as 1997 [10], a technology that transcends the limitations of the AI applications that are available today. Unlike AI, which is specialized to perform narrowly defined tasks such as face recognition or locating text, AGI is an all-encompassing intelligence capable of surpassing the average human's cognitive abilities. While the idea of AGI may sound like a fantasy, Greg Brockman, one of OpenAI's co-founders, in a podcast with Lex Fridman in 2019 [8], talked about his thoughts on the AGI. He believes that the recent AI wave that began in 2012 has three key features that could make AGI a reality. The first feature is the generality, meaning that AI models can solve problems in a variety of categories using neural network models plus backpropagation solutions. The second feature is the power of this generalized approach, which can disrupt traditional methods and revolutionize fields such as computer vision. Finally, the third feature is scalability, which suggests that the larger the model, the better its performance. The trend of larger models continues even now, as computing power becomes increasingly affordable. Thus, as the concept of AGI shifts from science fiction to technology, we can envision an intelligent entity that can excel in multiple domains and solve complex problems, including those that plague humankind. It is a tantalizing prospect, one that raises the question of what academic contributions could arise if geniuses like Einstein were endowed with the poetic talents of Chopin. Nonetheless, the development of such supreme intelligence also raises ethical questions that must be addressed.

3 VISIONS AND CHALLENGES

The initial statement of OpenAI highlights a lofty aspiration to develop strong artificial intelligence that is not only effective but also safe and beneficial to humanity. In pursuit of this aim, OpenAI

adopted an idealistic approach by forming a non-profit organization devoid of financial interests as written in their announcement in 2015 [5]. While financial gains are not inherently negative, OpenAI recognized that commercial competition could hinder its research objectives by prioritizing speed and profitability over security and safety. This misaligned focus could lead to disastrous consequences, as unchecked strong artificial intelligence could be exploited by malicious individuals or organizations. To avoid such outcomes, OpenAI chose to remain a non-profit organization, allowing it to prioritize research objectives over shareholder interests. This approach also enabled OpenAI to share its research results publicly without the pitfalls of commercial competition. However, this non-profit model is not without its limitations, as it may not always be effective in achieving its goals. OpenAI's charter includes a waiver of competition clause, which stipulates that the organization will shift from competition to collaboration with other entities striving to achieve strong artificial intelligence. Ultimately, OpenAI must maintain a balance between its idealistic aspirations and the need to remain competitive, ensuring that its research goals align with the interests of humanity.

4 NON-PROFIT DILEMMA

OpenAI's founding announcement showcased not only a distinguished lineup of top researchers but also an impressive array of investors. Noteworthy figures such as Elon Musk, renowned for his contributions to the automotive and aerospace industries, as well as Reid Hoffman, co-founder of LinkedIn, Peter Thiel, co-founder of PayPal, and Sam Altman, the then-president of Y Combinator, all lent their support to the company [5]. Y Combinator, a prominent name in Silicon Valley's tech investment sphere, had successfully nurtured companies with a combined market value exceeding $65 billion by 2015 [1], including well-known entities like Airbnb, Dropbox, and Docker.

Sam Altman, who had dropped out of Stanford's computer science program early in life and honed his skills in operations and finance at Y Combinator, assumed the role of OpenAI's CEO in 2019. However, prior to this pivotal moment, OpenAI had experienced a rather lackluster trajectory, characterized by incremental growth that fell short of its potential. Despite having a star-studded roster of investors and substantial funding amounting to billions of dollars, OpenAI's early results in the first three years appeared puzzlingly modest when compared to subsequent achievements like DALL-E and ChatGPT. This contrast was particularly noticeable when compared to the accomplishments of Google's DeepMind during the same period, with DeepMind's creation of AlphaGo and its victory over top Go player Lee Sedol in 2016 generating a global AI frenzy [3].

Before 2018, OpenAI had largely operated within the confines of the AI community, lacking widespread recognition. It wasn't until the second half of 2017 when OpenAI gained some prominence by defeating a highly ranked Dota 2 player in a one-on-one match, albeit limited to the mid lane and a specific hero selection [13]. However, this fleeting burst of attention stemmed solely from the Dota 2 International Invitation Tournament and quickly subsided. The lackluster performance of OpenAI in 2016 and 2017 can be attributed, in part, to the challenges posed by its non-profit model. Despite claiming to possess a substantial $1 billion, this figure was merely an aspirational target. The actual funding available to OpenAI was not as abundant as perceived.

Operating as a non-profit organization presented significant difficulties for OpenAI, as it struggled to raise necessary funds and provide competitive compensation packages to attract top talent. Unlike for-profit companies, OpenAI couldn't offer lucrative stock options or compete with the financial incentives provided by larger corporations. This predicament resulted in decreased income for renowned researchers and hampered the organization's ability to attract and retain talent. The financial demands of AI research are substantial, as exemplified by DeepMind's significant losses of $164 million in 2016 and $368 million in 2017 [23], even after accounting for generated revenue from services rendered to Google and health systems.

In contrast, OpenAI's funding fell far short of the same magnitude. As a non-profit organization, OpenAI is required to disclose its financial information annually. In 2016, its total expenses amounted to $11 million, with employee compensation accounting for $7 million [18]. When compared to the approximate $35 million spent solely on training AlphaGo by DeepMind [11], OpenAI's limited resources became evident. Insufficient funding hindered OpenAI's ability to realize its ambitious goals and impeded its researchers' pursuit of cutting-edge advancements. Consequently, several prominent individuals, including Ian Goodfellow, the author of the adversarial generation algorithm, departed OpenAI for other opportunities as he confirmed under a post on Reddit [9].

Despite these early challenges, OpenAI has since made significant strides in the field of AI research. In 2018, the organization announced the creation of GPT-2 [19], a language model capable of generating highly realistic and coherent text, which was deemed too dangerous to release due to concerns about potential misuse. Subsequently, OpenAI released a smaller version of the model, known as GPT-2 117M [20], which garnered widespread attention and was widely praised for its natural language processing capabilities.

In 2020, OpenAI made headlines again with the unveiling of GPT-3, an even larger and more powerful language model than its predecessor. With 175 billion parameters [7], GPT-3 represented a significant improvement over GPT-2 and was praised for its ability to generate high-quality and contextually appropriate responses to a wide range of prompts. The model was also notable for its zero-shot learning capabilities, which allowed it to perform certain tasks without any prior training.

In addition to language models, OpenAI has also made significant progress in other areas of AI research. The organization's robotic hand, known as Dactyl, demonstrated impressive dexterity and manipulation capabilities, while its reinforcement learning algorithms have been used to teach robots how to perform complex tasks [17]. OpenAI has also been involved in the development of generative models for images and videos, including the aforementioned DALL-E [21].

Overall, OpenAI's journey has been marked by ups and downs, but the organization's recent achievements have firmly established it as a leader in the field of AI research. While the non-profit model may present certain challenges, OpenAI has demonstrated that it is possible to achieve significant breakthroughs in AI without being driven solely by profit. As the field continues to evolve, it will be interesting to see what new developments emerge from OpenAI and other organizations working in this exciting area.

5 CAPPED-PROFIT

Immediately after his resignation as the president of Y Combinator, Sam Altman became OpenAI's first CEO, and under his leadership, the organization underwent a significant transformation. The company, which had previously been a non-profit entity, transitioned to a "semi-profit" business model and received a $10 billion investment from Microsoft [6]. As a result, OpenAI's valuation skyrocketed to $29 billion [22].

Altman's background as an entrepreneur and his experience at Y Combinator, where he coached startups on how to burn through money to achieve success, likely influenced the company's new direction. OpenAI's research efforts expanded, and the organization became involved in developing language models such as the GPT, which became a major focus of their work.

One significant achievement of OpenAI's research efforts was the development of AI technology capable of playing the complex game Dota 2 in 5v5 matches. This required significant investment in hardware, including 128,000 CPU cores and 256 GPU accelerators, but allowed for parallel AI training and achieved the equivalent of 180 years of gameplay in a single day [**?**]. Ultimately, OpenAI's bots were able to defeat the top team of Dota 2, OG [15].

Overall, Altman's leadership has transformed OpenAI from a small, fresh non-profit to a major player in the AI industry, and the company's valuation reflects this change.

6 GPT

In June 2018, OpenAI conducted a study on general-purpose language models that introduced a novel concept called Generative Pre-training (GPT) [19]. Although the term GPT was not explicitly used in the article, the model described is now known as GPT-1. GPT-1 is distinct from traditional task-oriented training methods in that it does not require manual labeling. Rather than creating datasets for specific tasks, GPT-1 takes existing human text and uses it to train the AI.

The process of Generative Pre-training involves asking the AI to generate the next sentence based on the previous one. For example, given the sentence "Mike studied hard every day and his teachers praised him as a good ____", the AI would complete the sentence with a word like "student." If the AI's response is incorrect, it is trained to move in the right direction. This process allows researchers to avoid the time-consuming and resource-intensive process of manually labeling training data.

Generative Pre-training can be applied to any textual information available to humans, including novels, texts, song lyrics, forum replies, and software code. By feeding more text into the neural network model and increasing the number of parameters, the pre-trained model's ability continues to grow. This process, combined with a model structure called Transformer, resulted in the Generative Pre-trained Transformer (GPT). GPT-2 and GPT-3 were subsequently developed, with the latter having a staggering 175 billion parameters [7].

OpenAI's GPT models have proven to be highly effective in specific tasks, such as translation and chatbot interactions. OpenAI's success has established it as a world-famous star AI company.

7 PROFITABILITY

The genesis of the "money power" phenomenon and its relationship to the departure of Elon Musk from OpenAI [14] can be traced back to the leadership of Sam Altman, the initial CEO of the organization. While Altman had achieved considerable success in developing a $100 billion business at Y Combinator, his expertise was primarily in the realm of Silicon Valley tech companies, rather than business. Prior to his departure from Stanford, Altman had studied computer science and possessed a technical rather than managerial background. It was under Altman's direction that OpenAI embarked upon a line of research into the demand for AI computing power [2], publishing a study on this topic in May 2018. This study concluded that the demand for AI computing power would soon outstrip Moore's Law and double every 3.4 months. This formed the basis for Altman's subsequent efforts to develop the technology required to meet this demand.

In March 2019, OpenAI established a limited partnership known as OpenAI LP [6], which was controlled by the board of directors of the former non-profit OpenAI. This entity used a unique "cap on returns" model, which enabled investors to earn a return on their investment, with a cap of 100 times their initial investment. Any returns above this cap would be directed to the non-profit OpenAI to fund inclusive research and education projects. Additionally, new investors were required to sign an agreement that placed OpenAI's charter first and vested decision-making power in the original non-profit OpenAI Board of Directors. While this corporate structure enabled OpenAI LP to raise capital and offer high equity incentives to employees, it also generated controversy and skepticism among some members of the tech community. Some criticized the 100x cap on returns, arguing that this was insufficient, while others expressed concern about the vagueness of the terms governing OpenAI LP. Despite these criticisms, OpenAI LP had the potential to serve as a vehicle for raising funds, supporting big projects, and enhancing employee morale, thereby transforming OpenAI from a defensive to an offensive force. Microsoft's entry into the fray, just three months after the establishment of OpenAI LP, added further intrigue to the unfolding story.

8 PARTNERSHIP WITH MICROSOFT

Since 2016, Microsoft has been providing cloud services for OpenAI's model training. The partnership was strengthened when Microsoft invested $1 billion in OpenAI LP shortly after its establishment, with a portion of the investment being in the form of vouchers for Azure cloud servers [4]. This collaboration has led to the rapid iteration and upgrading of OpenAI's GPT-2 and GPT-3 models, which have been transformed into services for widespread use, thanks to the support of Microsoft's cloud service capabilities.

One of the most significant outcomes of this partnership is the development of Codex [24], derived from GPT-3, which has been transformed into the GitHub Copilot service. The Copilot service is a code editor that automatically generates code based on user-provided text descriptions. The service has been popular among users, with the accuracy rate of generated code reaching almost 80%. Microsoft has been able to monetize the Copilot service by offering subscriptions to regular developers, while students and open-source project maintainers can use it for free.

In addition to the Copilot service, Microsoft has integrated OpenAI's DALL-E 2 image generation AI into its Bing Image Creator and Microsoft Designer software, which will generate direct revenue for the company [12].

Despite these developments, the explosion of ChatGPT, a conversational AI model based on GPT-3.5, has been a significant

development in the field [16]. ChatGPT's efficiency and accuracy have been remarkable, and the model has gained over 100 million users worldwide in just two months. Microsoft's Azure cloud service has played a significant role in the success of ChatGPT, and the model has been integrated into Microsoft's collaboration software Teams, which offers premium memberships for a fee.

Overall, Microsoft's investment in OpenAI has resulted in significant technological developments and direct revenue for the company. The collaboration has also put Microsoft in a favorable position against its competitors, including Google.

9 CONCLUSION

From a financial perspective, Microsoft's $10 billion investment in OpenAI does not result in a loss but raises justifiable doubts. OpenAI was initially founded with the goal of ensuring access to strong artificial intelligence for all of humanity. The intention was to prevent the monopolization of such powerful intelligence by individuals or organizations, which could lead to social polarization and catastrophic consequences. The original vision of OpenAI was that of a non-profit model, conducting research and sharing its results with society to prevent monopolization.

However, while the development of strong AI is still a long way off, Microsoft is benefiting financially from OpenAI's GPT models. OpenAI has gradually reduced its model disclosure, citing security concerns as the reason. GPT-3 models are no longer released publicly, and OpenAI provides a paid API interface instead. OpenAI has cited three reasons for this approach. First, the need to generate revenue through commercial products; second, the cost-ineffectiveness of GPT-3 for ordinary developers due to its size; and third, security concerns that are better addressed through the API format.

The first two reasons are straightforward and understandable, as they are focused on financial viability. The third reason aligns with the security concerns raised by OpenAI, which are not exaggerated given the prevalence of bots on the internet. The proliferation of AI-generated content, including articles, comments, and even videos, raises the potential for misinformation and manipulation. OpenAI's concerns about the misuse of its models are justified, as a large number of fake accounts disseminating false information could have detrimental effects.

While OpenAI's concerns are reasonable, it is essential to acknowledge the commercialization of its research and the flow of benefits to Microsoft. OpenAI's approach raises questions about adherence to its original non-profit model, although it can be argued that a balance is being maintained between business interests and the public good. Despite commercialization, OpenAI continues to publish papers on its models and offers free versions of services like ChatGPT and Copilot.

In conclusion, OpenAI's partnership with Microsoft, facilitated by substantial investments, has enabled remarkable developments in AI. While OpenAI's image may have lost some of its initial perfection, it is still striving to serve humanity's interests in the realm of AI. The company plays the role of both a goalkeeper and a scorer, aiming to strike a balance between business and societal benefit. The presence of a company like OpenAI, although imperfect, is preferable to having none at all. It represents an attempt to navigate the challenges of AI while working towards the ideal of benefiting all of humanity.

REFERENCES

[1] S. Altman. Yc stats, Aug 2015.
[2] D. Amodei and D. Hernandez. Ai and compute, May 2018.
[3] BBC. Artificial intelligence: Go master lee se-dol wins against alphago program. *BBC News Online*, Mar 2016.
[4] G. Brockman. Microsoft invests in and partners with openai to support us building beneficial agi, Jul 2019.
[5] G. Brockman, OpenAI, and I. Sutskever. Introducing openai, Dec 2015.
[6] G. Brockman, I. Sutskever, and OpenAI, Mar 2019.
[7] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei. Language models are few-shot learners, 2020.
[8] L. Fridman and G. Brockman. Greg brockman: Openai and agi, Apr 2019.
[9] I. Goodfellow, 2017.
[10] M. Gubrud. Nanotechnology and international security. *Fifth Foresight Conference on Molecular Nanotechnology*, Nov 1997.
[11] G. Marcus. Deepmind's losses and the future of artificial intelligence, Aug 2019.
[12] Y. Mehdi. Create images with your words - bing image creator comes to the new bing, May 2023.
[13] OpenAI. More on dota 2, Aug 2017.
[14] OpenAI, Feb 2018.
[15] OpenAI. Openai five defeats dota 2 world champions, Apr 2019.
[16] OpenAI. Introducing chatgpt, Nov 2022.
[17] OpenAI, M. Andrychowicz, B. Baker, M. Chociej, R. Jozefowicz, B. McGrew, J. Pachocki, A. Petron, M. Plappert, G. Powell, A. Ray, J. Schneider, S. Sidor, J. Tobin, P. Welinder, L. Weng, and W. Zaremba. Learning dexterous in-hand manipulation, 2019.
[18] ProPublica, Dec 2017.
[19] A. Radford and K. Narasimhan. Improving language understanding by generative pre-training. 2018.
[20] A. Radford, J. Wu, D. Amodei, D. Amodei, J. Clark, M. Brundage, and I. Sutskever, Feb 2019.
[21] A. Ramesh, M. Pavlov, G. Goh, and S. Gray. Dall·e: Creating images from text, Jan 2021.
[22] P. Rosen. Chatgpt's creator openai has doubled in value since 2021 as the language bot goes viral and microsoft pours in $10 billion, Jan 2023.
[23] S. Shead. Deepmind losses grew to $368 million in 2017, Oct 2018.
[24] W. Zaremba, G. Brockman, and OpenAI, Aug 2021.