# LLM/GPT Generative AI and Artificial General Intelligence (AGI): The Next Frontier

Atif Farid Mohammad
OmniAGI.ai
Los Angeles, CA
atif@omniagi.ai

Bryan Clark
OmniAGI.ai
Los Angeles, CA
bryan@omniagi.ai

Ram Agarwal
OmniAGI.ai
Los Angeles, CA
ram@omniagi.ai

Sean Summers
OmniAGI.ai
Los Angeles, CA
sean@omniagi.ai

***Abstract*** *— The dawn of artificial intelligence has sparked excitement and curiosity as it paves new ways for us to envision the future. Artificial General Intelligence (AGI) stands as a thrilling and challenging frontier to explore, offering many possibilities and transformations across various industries. This article attempts to provide an in-depth overview of AGI, its history, applications, challenges, and ethical implications. Before diving into the world of AGI, it is prudent to familiarize oneself with the key concepts and distinctions that shape this field. This section aims to define AGI, differentiate it from other types of artificial intelligence, and offer a glance at its historical development and its potential future development.*

***Keywords — Artificial General Intelligence (AGI), Generative AI, GPT, LLM***

## I. Introduction

Artificial General Intelligence (AGI) can be defined as the creation of machines that possess the ability to understand or learn any intellectual task that a human being can do. In simple terms, AGI is a machine's capacity to perform tasks and solve problems with a level of cognitive ability comparable to human intelligence. AGI encompasses various abilities, such as reasoning, learning, problem-solving, perception, and language understanding. These abilities allow an AGI system to be versatile and adaptive, granting it proficiency in diverse tasks without being explicitly programmed for each one.

The promise of AGI is the ability for digital systems to be able to help humanity solve its most complex challenges. Climate change, income inequality, drug discovery, space exploration, material science, food production, and mental health are all areas where AGI can make significant contributions. However, the road to producing a system capable of AGI is not clear. The capabilities of Large Language Models (LLMs) are hugely impressive and continue to improve daily, however there is a significant environmental cost associated with their usage. Li et al described this problem in their paper "Making AI Less Thirsty: Uncovering and Addressing the Secret Water Footprint of AI Models", according to their research "training GPT-3 in Microsoft's state-of-the-art U.S. data centers can directly consume 700,000 liters of clean freshwater and the water consumption would have been tripled if training were done in Microsoft's Asian data centers, but such information has been kept as a secret" [1]. Different approaches to building systems that are capable of AGI or support the implementation of AGI will need to be developed and they will need to focus on environmental sustainability, we will propose various approaches for this end.

## II. State of the Art

### *A. AI Landscape*

A crucial distinction in the AI landscape is the difference between AGI and Narrow AI (also known as Artificial Narrow Intelligence or ANI). Narrow AI [2] is the AI technology that we currently possess, which excels in specific tasks or domains, while lacking the versatility of human-like understanding or adaptability across multiple tasks.

### *B. Narrow AI and AGI*

Examples of narrow AI include language translation algorithms, facial recognition systems, and self-driving car technologies. These AI systems are designed for a specific purpose and cannot perform tasks beyond their designated domain. In contrast, AGI aims to equip machines with the cognitive flexibility to apply their intelligence across various tasks and domains, emulating human-like strategic thinking and problem-solving.

The concept of AGI can be traced back to the early days of artificial intelligence research. John McCarthy, Marvin Minsky, Allen Newell, and Herbert A. Simon, [3] the founding fathers of AI, believed in the possibility of creating machines that could perform any intellectual task a human could. The term "artificial intelligence" was even coined as an attempt to describe machines or systems "exhibiting behavior at least as skillful and flexible as humans." Since then, AI research has seen many milestones, and the pursuit of AGI remains a significant area of interest.

Over the years, AI research has mainly focused on creating specialized AI systems, leading to the development

979-8-3503-2759-5/23/$31.00 ©2023 IEEE
DOI 10.1109/CSCE60160.2023.00073

of many successful narrow AI applications. However, the ultimate goal of achieving AGI has proven more elusive than initially anticipated. Recent advancements in deep learning, neural networks, and reinforcement learning have rekindled the hope of progressing towards AGI, but there is still a long way to go.

One of the most widely used methods is Convoluted Neural networks (CNN). Considering that it is modeled after the human brain and seeks to acquire knowledge in a manner similar to that of learning to extract and organize complicated features, CNNs have demonstrated extraordinary success in facial recognition. A typical CNN comprises input layers, hidden layers, and an output layer, sometimes called neurons. The input layer accepts arrays of the picture pixels as input. There may be several hidden layers in CNNs that use mathematical concepts to extract features from the image. Convolution, pooling and completely connected layers are a few of the layers present in the hidden layers [4]. Convolution layer is the first layer for extracting features from an input image. Once the image passes through the hidden layers, it hits one of the output layers which helps in classifying and deducing results from the image processing. This is a high-level overview of how CNN's work.

Open AI as an organization is at the forefront of machine learning and artificial intelligence research. Their ChatGPT product started off as an isolated language model that had no connection to the outside world. Now, Open AI is developing and releasing integrations that will allow their language models to interact with other applications and the internet in general. This will bring new and exciting capabilities across a variety of industries and will influence a renaissance in web technologies. Ray Kurzweil discusses linear vs exponential growth in his book "The Singularity is Near: When Humans Transcend Biology" and describes the concept of the "Knee of the curve", a point at which growth rapidly increases and the change from linear to exponential growth is observed [5]. The launch and widespread use of LLMs will mark the turning point where technological innovation moves from linear gains to exponential gains, as more users have access to capabilities that they would not normally have had, they will be able to rapidly develop solutions and collaborate with others to achieve even more than they otherwise would alone.

## III. The Path to Achieving AGI

The development of AGI is a complex and multifaceted undertaking, requiring rigorous research in various avenues of artificial intelligence. This section delves into the key approaches and techniques explored, the challenges underlying AGI, and the critical breakthroughs and milestones.

### *A. Key Approaches and Techniques*

Various methodologies have been proposed to develop AGI, ranging from neural networks and deep learning to more symbolic and logic-based approaches.

### *B. Key Techniques*

- Reinforcement Learning: an approach in which an AI agent interacts with its environment and learns to achieve its goal through trial and error, guided by a reward system.
- Deep Learning: a subset of machine learning employing artificial neural networks to learn hierarchical representations of data, thereby enabling computers to perform tasks with minimal human intervention.
- Knowledge-based systems: a symbolic approach that uses formal rules and logic to represent and manipulate knowledge, allowing machines to reason about complex problems.
- Evolutionary Algorithms: a class of optimization algorithms inspired by natural evolution, which uses techniques such as selection, crossover, and mutation to solve optimization problems.

### *C. Challenges and Limitations*

Despite the promising advancements in AI research, several challenges must be addressed before AGI can become a reality. Some of the most significant challenges include:

1. Scalability: Developing an AGI system that can scale its learning and understanding across multiple domains, tasks, and environments remains a significant hurdle.
2. Transfer Learning: AGI requires the ability to transfer knowledge from one domain to another, a feat that remains challenging for current AI systems.
3. Explainability: As AGI systems become more sophisticated, understanding the decision-making process behind their actions becomes increasingly crucial yet difficult.
4. Resource Constraints: Developing AGI requires vast computational resources, posing a challenge to those with limited access to cutting-edge hardware and expertise.

### *D. CNN Architecture*

As mentioned earlier, a CNN architecture comprises of Input, Output layers along with multiple hidden layers. CNN aims to detect points in the whole image, which can be useful for it to make its calculations. In this analysis, 4 convolution layers have been used. Images are first parsed through the first convolutional layer where these images are converted to 128x128 pixels after max pooling process. Next, these images are sent through the second convolutional layer where essential features like facial geometry points are extracted by applying max pooling. Finally, the final convolutional layer resizes these images to 32x32 pixels size format. Photos are converted into arrays to make measuring quicker. The application of a connected layer is the last phase. In both of the convolution layers, we employed the Relu activation function, and in the output layer, we used the Softmax activation function. Adam stochastic gradient descent is utilized to determine the most efficient outcomes [6]. This process is illustrated in detail in Figure 1.


Figure 1: Adam stochastic gradient descent [4]

## IV. POTENTIAL APPLICATIONS OF AGI

Although AGI has yet to be achieved, some notable breakthroughs in artificial intelligence research have inched us closer to realizing this goal. For instance, OpenAI's GPT-3, a language model that can generate human-like text, has demonstrated impressive abilities in multiple language tasks. DeepMind's AlphaGo and AlphaZero programs have showcased expert-level gameplay and self-learning in competitive board games. These milestones provide valuable insights into the potential capabilities of AGI, inspiring further research and development.

The successful development of AGI could revolutionize various industries and professions, accelerate scientific discovery, and contribute to solving intricate global challenges. This section highlights the transformative potential of AGI across different domains.

### *A. Transforming Industries and Professions*

AGI could significantly impact industries such as healthcare, finance, and manufacturing, offering innovative solutions and streamlining processes. For example, AGI systems could assist physicians in diagnosing and treating complex medical conditions, help financial analysts find optimal investment strategies, or revolutionize supply chain management by optimizing production and logistics.

### *B. Advancements in Scientific Research*

One of the most intriguing aspects of AGI is its potential to accelerate scientific research and discovery. AGI systems could assist researchers in finding groundbreaking solutions to long-standing problems, as well as providing innovative ideas and techniques to various scientific fields, such as physics, chemistry, and biology.

### *C. Solving Complex Global Problems*

AGI has the potential to help humanity address some of its most pressing challenges, including climate change, poverty, and inequality. By providing new insights and solutions, AGI could contribute to more sustainable policies, resource management, and equitable distribution of wealth, thus benefiting societies across the globe.

## V. ETHICAL CONSIDERATIONS AND IMPLICATIONS

The development of AGI raises serious ethical concerns, revolving around its alignment with human values, potential misuse, and governance. This section examines the importance of addressing these ethical questions as we strive towards AGI.

- Ensuring AGI Aligns with Human Values

As we develop AGI, it is critical to ensure that it aligns with human values and serves humanity's interests. Researchers and developers must prioritize the safe and responsible design of AGI technology, incorporating fairness, transparency, and inclusiveness to minimize potential harms and maximize the benefits for all.

As AGI development advances, the power and autonomy of the systems will increase and reason towards goals that may risk losing alignment with human values and create new conflicts of interest in prioritizing the objectives AGI is given to reason about. To moderate this, it is crucial to develop methods to control and influence the decision-making processes. AGI could be applied for malicious purposes in applications such as what narrow AI has already been used, such as autonomous weapons and surveillance systems that infringe on privacy rights.

Ensuring responsible use requires robust safeguards and regulations that should be developed with international regulation, shared legal frameworks, and comprehensive oversight to prevent deployment that causes harm to human individuals, institutions, or societies.

- Addressing Potential Misuse and Risks

The power of AGI comes with the potential for misuse, and it is crucial to consider how to safeguard against unintended consequences and malicious exploitation. Establishing strict security protocols, fostering international collaboration, and devising robust countermeasures are crucial measures to mitigate the risks posed by AGI.

An analysis of the article "Generative AI Breaks the Data Center: Data Center Infrastructure And Operating Costs Projected to Increase to Over $76 Billion by 2028" the author states that "Companies began sound the alarm about data center power consumption five years ago at the annual Hot Chips semiconductor technology conference by predicting that worldwide compute demand could exceed the total world electricity power generation within a decade." [6] Generative AI workloads will increase as more organizations identify the business value associated with Generative AI usage. The need to develop solutions that reduce the environmental impact of Generative AI solutions needs to be thoroughly investigated. LLMs that have hundreds of billions and even trillions of parameters require vast resources to train and run so that they can deliver prompt responses to users.

Developing techniques to extract similar performance from smaller models that use less resources will become an increasingly incentivized business problem to solve. Consumers will want to access the capabilities of LLMs on their personal devices and small parameter LLMs that can deliver just enough insight to users to be useful while running on edge devices and having minimal environmental impact.

- The Debate on AGI Regulation and Governance

To ensure the responsible and ethical development of AGI, there is an ongoing debate surrounding its regulation and governance. Policymakers and AI researchers must engage in a continuous dialogue to create comprehensive frameworks and policies that balance innovation with ethical concerns, thereby guiding the future of AGI in a manner that benefits humanity. AGI will undoubtedly change the nature of work as it currently exists today. Research suggests that 32.8% of occupations may face full impacts from LLMs such as ChatGPT [7]. Lawmakers and governments around the world will have to carefully study the impact that LLMs have on the employability of its citizenry, so that a strategy is in place to address the disruption caused these technological advancements. A basic starting point could be the certation of job retraining programs that are state funded that identify Individuals who are affected or displaced by AGI and LLMs technology developments.

Corporations and governments will also have to come to terms with the capabilities and limitations of AGI and ensure that safeguards are in place to prevent models from producing output that is damaging to society, perpetuates hate, violence, or other undesirable outputs. This will require careful collaboration between business and government to ensure that legislation and regulations safeguard the public yet do not stifle innovation.

In conclusion, Artificial General Intelligence holds the potential to revolutionize our world and redefine the boundaries of human innovation. As we venture into this exciting landscape, it is paramount that we work collectively to address the challenges, embrace the opportunities, and ensure the future of AGI aligns with the best interests of humanity.

- The Biological and Quantum Approach to AGI

Evolution has provided us with biological intelligence. The hallmark of biological intelligence is extreme efficiency when compared to digital computational systems, massively parallel processing of information and low power consumption. Experimentation with the use of neuron cells to control robots has shown remarkable success in laboratory testing. Take this experiment for example "Potter places a droplet of solution containing thousands of rat neuron cells onto a silicon chip that's embedded with 60 electrodes connected to an amplifier.

The electrical signals that the cells fire at one another are picked up by the electrodes which then send the amplifies signal into a computer... The robot then manifests this neuronal activity with physical motion, each of its movements a direct result of neurons talking to neurons" [8]. This experiment and others like it, laid the foundation for an emerging discipline called Dishbrain. Kagan et al describe Dishbrain in their paper "In Vitro neurons learn and exhibit sentience when embodied in a simulated game-world." The conclusion of their research was that "integrating neurons into digital systems may enable performance infeasible with silicon alone. "Here, we develop DishBrain, a system that harnesses the inherent adaptive computation of neurons in a structured environment. In vitro neural networks from human or rodent origins are integrated with in silico computing via a high-density multielectrode array" [9].

The Dishbrain approach has interesting implications on the environmental impact of AI and Machine Learning activities. Biological systems leverage massively parallel neural networks and have a fundamentally different data storage and retrieval mechanism when compared to silicon solutions. Mehonic and Kenyon frame the problem in their article "Brain-inspired computing needs a master plan," "The energy problem is largely a consequence of digital computing systems storing data separately from where they are processed. This is the classical von Neumann architecture underpinning digital computing systems. Processors spend most of their time and energy moving data. Fortunately, we can improve the situation by taking inspiration from biology, which takes a different approach entirely collocating memory and processing, encoding information in a wholly different way or operating directly on signals, and employing massive parallelism" [10].

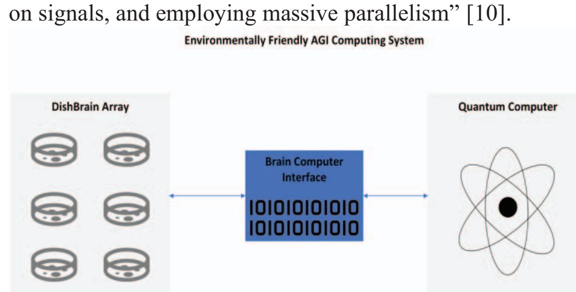


Figure 2: Brain & Computer Interface Comprehension

Quantum computers, which perform operations on qubits, are rapidly outpacing the capabilities of silicon processors on many tasks. According to the article "Quantum supremacy using a programmable superconducting processor," by Google AI Quantum as of 2019 their Sycamore processor "takes about 200 seconds to sample one instance of a quantum circuit a million times-our benchmarks currently indicate that the equivalent task for a state-of-the-art classical supercomputer would take approximately 10,000 years" [11]. Currently, it is not the case that every computational task performed by a quantum computer will vastly outperform classical algorithms, or silicon processors, the type of problem that is being solved must be benchmarked, and a clear advantage to using quantum computation must be established.

In the system described above the Dishbrain array would be optimized to perform calculations that require massive parallelism such as image classification tasks and tasks that require the computational capabilities of the quantum computer will be delegated from the Dishbrain array to the quantum computer. Tasks that would be appropriate for the quantum computer would be simulation tasks or tasks that experience exponential growth in the amount of information that needs to be processed.

This combination of computational capabilities will produce a new class of computers that can complete supremely difficult tasks with a greatly reduced environmental footprint. This architecture will represent a new paradigm in computing and could potentially power the development of AGI that can be generalized to complete any task assigned to it.

## CONCLUSION

AGI will be a powerful technology that can be leveraged to solve many of the problems that we face as a species in our world today. Governments and corporations will need to work together to project the impact of AGI and current generation generative AI models and prepare for the disruption that will result for the occupations that are impacted. Responsible use legislation will have to be developed to limit the harm that can result from unscrupulous actors leveraging AGI for nefarious purposes. Industry will also need to develop AGI that does not have a disproportionate impact on the environment, smaller models that deliver needed functionality to users should be explored and Dishbrain + Quantum computing hybrid computing solutions should be explored for their viability in servicing the computational requirements of AGI systems. Our current technological progress has already breached or will soon breach the line between linear growth and exponential growth.

With current LLMs readily available to all users who have interact access will provide users with the ability to produce products and services that they would not have been able to produce prior, due to access to knowledge and software development resources. Further study on the ethical and economic impact of highly capable LLMs and AGI will need to be conducted to prepare for the disruption that could result. Research on the interaction between Dishbrain and Quantum systems will also have to be performed so that high performance low environmental impact computational systems can be developed to alleviate the pressure on electrical grids around the world, and to reduce carbon emissions from computational activities.

## References


[1] Li, P., Yang, J., Islam, M. A., & Ren, S. (2023). Making AI Less "Thirsty": Uncovering and Addressing the Secret Water Footprint of AI Models. *ArXiv (Cornell University)*.

[2] Ramuš Cvetkovič, Iva, and Marko Drobnjak. "As Above so Below: The Use of International Space Law as an Inspiration for Terrestrial AI Regulation to Maximize Harm Prevention." *Artificial Intelligence, Social Harms and Human Rights*. Cham: Springer International Publishing, 2023. 207-238.

[3] Prentice, Catherine. "Demystify Artificial Intelligence." *Leveraging Emotional and Artificial Intelligence for Organisational Performance*. Singapore: Springer Nature Singapore, 2023. 25-40.

[4] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, ``Going deeper with convolutions," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2015, pp. 1-9.

[5] Kurzweil, R. *The singularity is near : when humans transcend biology*. Penguin Books, 2005, 10.

[6] McGregor, J. (n.d.). *Generative AI Breaks The Data Center: Data Center Infrastructure And Operating Costs Projected To Increase To Over $76 Billion By 2028*. Forbes. Retrieved May 24, 2023, from https://www.forbes.com/sites/tiriasresearch/2023/05/12/generative-ai-breaks-the-data-center-data-center-infrastructure-and-operating-costs-projected-to-increase-to-over-76-billion-by-2028/?sh=40da55667c15

[7] Zarifhonarvar, A. (2023). Economics of ChatGPT: A Labor Market View on the Occupational Impact of Artificial Intelligence. *EconStor Preprints*. https://ideas.repec.org/p/zbw/esprep/268826.html

[8] *Rat-Brained Robot*. (n.d.). MIT Technology Review. Accessed May 24, 2023, from https://www.technologyreview.com/2002/12/18/234548/rat-brained-robot/

[9] Kagan, B. J., Kitchen, A. C., Tran, N. T., Habibollahi, F., Khajehnejad, M., Parker, B. J., Bhat, A., Rollo, B., Razi, A., & Friston, K. J. (2022). In vitro neurons learn and exhibit sentience when embodied in a simulated game-world. *Neuron*.

[10] Mehonic, A., & Kenyon, A. J. (2022). Brain-inspired computing needs a master plan. Nature, 604(7905), 255–260.

[11] Arute F., Arya, K., Babbush, R., Bacon, D., Bardin, J. C., Barends, R., Biswas, R., Boixo, S., Brandao, F. G. S. L., Buell, D. A., Burkett, B., Chen, Y., Chen, Z., Chiaro, B., Collins, R., Courtney, W., Dunsworth, A., Farhi, E., Foxen, B., & Fowler, A. (2019). Quantum supremacy using a programmable superconducting processor. *Nature*, *574*(7779), 505–510.