

The World of Artificial General Intelligence

If one day, we can simply tell an AI to create an app like TikTok or YouTube with a single command, and it provides us with a complete project containing millions of lines of code, that would be truly amazing.

When ChatGPT was released in late November 2022, it could perform tasks like fixing a snippet of code and answering a wide range of questions. It excelled at explaining concepts such as neural networks and providing “Hello, World!” examples in various programming languages. It was like having a very powerful search engine that could give direct answers.

Then, it had search functionality. I could instruct ChatGPT to retrieve a list of links from a website and create a document about them.

Then, versions like ChatGPT 3.5, 4, 4o, o1-mini, and o1 were released.

Now, it can handle requests like adding a dark mode to a website. It can provide the necessary code and instructions to update HTML, CSS, or scripts, and even suggest adding a dark mode toggle. Implementing dark mode involves changing CSS, and if the site uses Markdown, that needs to be updated as well.

It’s as if the AI can implement entire features or functionalities, not just snippets of code.

By combining these features, we can create an application. So, one day, if we tell an AI tool to build a terminal, a browser, a to-do list, a task app, a calendar, a code collaboration tool, or a meeting app, it could provide the entire project code.

We can then make the task more complex. For example, we could ask the AI to integrate all existing YouTube code and use APIs from OpenAI, Claude, or Deepseek to add AI functionalities to YouTube. This could include adding a smart assistant, replacing current translations with AI-powered ones, enhancing search capabilities with AI, and even curating dedicated short videos, like requesting YouTube to provide 100 funny short videos about Japanese life.

So, this is an app. But what about more advanced tasks, like creating an operating system? We could tell the AI to design a new, fully open-source operating system with a fresh design, basic apps, a terminal, a command line, and a scheduler, similar to Oberon, using data structures to interact between processes instead of strings.

What’s next? We could ask the AI to design the latest Mac and update its operating system.

And then, what’s next? We could tell the AI to design and update an entire home, customizing all electrical products based on our activities, the latest knowledge, and our needs to create a better living environment.

And then, what's next? We could ask the AI to design an entire city, tailored to its citizens' behaviors and the latest knowledge, to improve their lives.

And finally, what's next? We could tell the AI to improve the Earth, using all available knowledge and information to enhance everyone's life.

I am struggling to come up with a title for this essay. Let's call it "The World of Artificial General Intelligence"

According to Deepseek, "Artificial General Intelligence (AGI) refers to a type of artificial intelligence that possesses the ability to understand, learn, and apply knowledge across a wide range of tasks at a level comparable to human intelligence." Unlike Narrow AI, which is designed for specific tasks such as facial recognition, language translation, or playing chess, AGI can perform any intellectual task that a human can do.

When considering the future of AI, there are two basic points to grasp: algorithms and compute. AI algorithms are mainly involved in calculus, backpropagation, transformers, GPT, and multihead latent attention.

In the digital world, there will be mappings from X to Y, where X can be anything from text, images, video, audio, code, to any byte data. Y can be any of these as well.

Computers don't inherently understand AGI; it is merely a definition created by humans and doesn't matter much to machines.

The application of AI in the physical world will include areas like autonomous driving and robotics. If the digital world can map X to Y, the physical world will follow suit. For example, a robot can turn ingredients into dishes, construct Legos, decorate a house, tile floors, install air conditioners, and assemble IKEA furniture.

There are already industrial robots in use. Notable companies in Japan include FANUC, Kawasaki Heavy Industries, and Yaskawa Electric Corporation.

So why aren't there more robots in homes? Consumer robots need to be versatile and capable of performing multiple tasks. For instance, a cooking robot might only stir and fry ingredients, requiring users to prepare the ingredients and clean up afterward.

In the future, robots will be present in homes, stores, schools, offices, cinemas, and tourist attractions—essentially anywhere human workers are currently employed.

There will be a world model in the cloud, a very large model potentially around 100 petabytes in size. For reference, 1 petabyte is 1,024 terabytes, and 1 terabyte is 1,024 gigabytes. One version of the Llama 3 70B model has a file size of 21.1 GB.

Robots in the world will need to consult this world model in the cloud to take actions. A network delay of 100 milliseconds or even 1 second is acceptable as long as the robot can perform its tasks effectively.