**Generative AI: A new weapon in cognitive warfare**

**Authors: Chen Dongheng and Xu Yan**   
**Date: 2023-04-03**

**introduction**   
Recently, OpenAI in the United States upgraded ChatGPT to a new version, GPT-4. Compared with the old version that can only receive text questions, GPT-4 can work in dozens of languages ​​and can also recognize image information. The rapid development of generative AI represented by ChatGPT has opened up a new path for humans to understand the world, and also provided new weapons for some countries to carry out cognitive penetration attacks against foreign countries.

**Cognitive Application Capabilities of Generative AI**

Generative AI refers to AI that uses various machine learning methods to learn and extract the core features of objects from data, and then generates new and completely original content. ChatGPT and its iterative products are essentially artificial intelligence language models based on human natural language. It uses large data as raw materials, and through intensive training and deep learning, it has a feedback and error correction mechanism similar to that of neural networks. It can create new content according to the specific needs of users, and automatically complete complex tasks such as writing texts, writing articles, coding, painting, editing videos, etc., giving wings to human cognition.

Improved cognitive speed. Generative AI is supported by AI as its core and uses high-performance algorithms as its weapon. It gathers large-scale data information and is the epitome of human wisdom and data information. The emergence of generative AI marks that the development of artificial intelligence is moving from quantitative change to qualitative change, and its cognitive speed far exceeds all previous AI cognitive machines. According to some data, the 45TB data and nearly 1 trillion word training model that support the more natural and smooth language synthesis results of ChatGPT are equivalent to 13.51 million Oxford dictionaries. Although GPT-4's cognitive algorithms, computing power and training data have not yet been made public, it is undoubtedly a more powerful super astronomical figure, which is the key support for its powerful cognitive speed and cognitive ability.

Cognitive trend prediction. The magic of generative AI is that it can predict the development trend of things based on strong data support, algorithms and the laws of things. Public information shows that the US military's "Advanced Algorithm Combat System" currently supported by GPT-like artificial intelligence can summarize thousands of enemy artillery bombardments, ground attacks and air strikes, and use AI algorithms to summarize the enemy's action patterns and combat habits, and even simulate the "thinking mode" of enemy commanders to propose targeted response plans.

Inducing cognitive bias. No matter how quickly generative AI iterates, it is still an AI created and controlled by humans. Behind the dazzling functions, its R&D operators use the design keys and training data sets they have mastered to manipulate the cognitive stance and attitude of the machine output. At the same time, major risks and challenges that may be caused by machine cognitive failure or loss of control are not ruled out. Data shows that OpenAI spends a lot of debugging energy on English data rather than Chinese content. Chinese data accounts for only 0.1% of ChatGPT's training corpus. The logic and accuracy of the English output results are much greater than that of Chinese, and its cultural value tendency can be seen. When generative AI and its products with clear political stances and value tendencies become more and more deeply rooted in people's lives, it is difficult for the audience's cognition not to be influenced by the tendency of the answers to the preset political stances they provide, and this system is therefore more likely to become a tool for cognitive attack and defense.

**Cognitive application characteristics of generative AI**

As a highly intelligent form of AI, generative AI has the general characteristics of AI cognitive warfare applications in cognitive warfare, while also having the special characteristics of highly intelligent cognitive confrontation. Understanding the characteristics and laws of generative AI cognitive applications will help us recognize its actual and potential value in cognitive offense and defense, and expand the methodological path of cognitive warfare.

**Autonomous operation**.   
This is a significant feature of generative AI cognitive applications. Generative AI has a certain degree of self-awareness and can autonomously "emerge" to generate new content according to the underlying logic of its R&D application, thereby producing a continuous and uncertain impact on human cognition. The value of its cognitive application lies in that the R&D and application users can rely on the mastery of system design and training language to continuously act on the thinking and cognition of the target object with higher quality and efficiency; similarly, the "audience" as the opponent is more likely to be confused and controlled under the seemingly objective continuous influence and intervention of AI. In future cognitive confrontations, with the blessing of generative AI, the organizers of the attack can hand over the cognitive attack and defense tasks to AI machines and programming. For the "audience", whether the information is true or false, whether the judgment is right or wrong, whether the decision is good or bad, and whether the action is good or bad, everything will become more confusing. Because you may not be facing a living person, but a program or machine with intelligent human-like cognitive ability.

**Interactive game.**   
This is the logical feature of generative AI cognitive application. Generative AI is actually the cognition of the interaction between human ontology and machine cognition. Humans empower programs and machines with their own wisdom. At the same time, cognitive machines are also answering audience questions, creating new content according to the specific needs of the audience, and then forming new capabilities, which are used to enrich and improve the system. Therefore, the cognition possessed by ChatGPT-like is the superposition of the interaction between humans and machines within and outside the program. In this process, based on the generative AI authorized by the developer, it can mobilize and use all data reserves, computing power, and algorithms to play games with people as "opponents" or the same ChatGPT-like on issues of concern, thereby subtly influencing, inducing or suppressing the other party.

**Hidden disturbance**.   
This is the deep characteristic of generative AI cognitive application. Generative AI is essentially a highly intelligent program or machine controlled by humans in the underlying loop. Its generation is based on mixed generation under human dominance, rather than a frivolous generation that is completely separated from humans. Like all technologies, AI that is independent of humans does not exist. It is the dazzling cognitive behavior supported by its high intelligence that provides a hidden space for the real cognitive actors to play their roles. It is based on this that the source, method, content and direction of generative AI generation have a distinct humanistic color, including political stance, value orientation, moral standards, cultural attributes and even personality traits. In turn, we can also recognize its origin and nature through these humanistic colors. When you are faced with a machine or program with considerable cognitive response capabilities and a sense of technology, it is difficult for you to judge whether there is another dominant force behind it from the response of the interaction with it. It is this a priori worship and trust in technology that often makes people lose their vigilance against the political tendencies, value criteria, moral standards and humanistic attributes it carries.

**Cognitive Applications of Generative AI**

**Cognitive habit shaping.**  
 It is to enhance the convenience of using superior technology products, and to form inertial thinking about the use of the product through long-term immersion, so as to achieve a credible, approachable, and unguarded cognitive convergence. This is the traditional practice of the United States and some other Western countries. In the early years when the United States and the West developed the international Internet and operating systems, they shaped the market through low prices or free use, so that opponents could become "addicted" to it unknowingly. ChatGPT appears as a "caring person" and "little assistant" at work, and also provides a free version. These seemingly harmless but actually artificially controlled generative AIs, once they become "close friends" that people trust and talk about everything, their influence on people will be subtle. This reminds us that there are traps in "being close to the people" and we must not be fooled by the appearance of warmth and affection.

**Cognitive information penetration.**  
 Exporting values ​​through media platforms and cultural and creative products to influence the value orientation of people in other countries is a common practice in the United States and the West. The United States has used its technological advantages to promote American values ​​by launching a large number of databases and information products through the Internet. The United States' first-mover advantage in generative AI undoubtedly gives it new tools to infiltrate information into its opponents. This reminds us that some seemingly objective and fair data and information are actually "colored" or even "deteriorated." We must effectively strengthen national security, especially the protection of strategic resources such as data and information, and prevent the replacement of value rationality with technical rationality.

**Cognitive thinking trap.**   
It is to take advantage of human weaknesses, stimulate primitive desires such as the pursuit of comfort, novelty, and stimulation, and influence people's thinking and cognition in the form of "curiosity entertainment + consciousness penetration". Generative AI such as ChatGPT targets the social and emotional dilemmas of modern people and appears as a pleasant chat robot, which meets the emotional appeals of many people. This kind of technology application that focuses on catering to and utilizing human characteristics can easily lead people's thinking and cognition into narrow, one-sided, and superficial traps. This warns us that we must remain rational enough about new technologies, and while using their advantages and strengths, prevent ourselves from falling into possible artificial traps.

**Cognitive standard hostage.**It is to bundle value logic with technical logic, and use technological advantages to forcibly infuse corresponding values ​​into products. This is an emerging practice in the United States, the West and other countries. For example, the Internet is an indispensable and important tool in modern society, and the United States and Western countries that have absolute control over the Internet use it as a weapon for value penetration. The United States has suppressed voices and practices that differ from those of the United States and the West on certain issues by controlling the use of the Internet and the use of international payment systems. Generative AI represented by ChatGPT can still replicate this "bundling" hostage trick, and use the opportunity of generating answers to forcibly provide "standard answers" that integrate American values. This reminds us that to strengthen cognitive confrontation, we must strengthen strategic information system technology innovation and product research and development, and use strong science and technology to support strong discourse power in cognitive confrontation.

(Author’s unit: Military Political Work Research Institute of the Academy of Military Sciences)