

Substantive and Emotional: The ‘Anthropomorphic’ Evolution of Generative AI

--Taking GPT-4o as an Example

Jingyi Xie^{1,a,*}

¹*Key Laboratory of Big Data Analysis and Application of User Behavior in Publishing Industry, Beijing Normal University, No.19 Xijiekouwai Street, Haidian District, Beijing, China*

a. xtt13770664601@163.com

**corresponding author*

Abstract: Under the background of the rapid development of artificial intelligence technology, "personal publishing" represented by blog in the past has transformed into "human-computer co-creation", showing a new human-computer interaction relationship, and generative artificial intelligence has become the "collaborator" of human content creation. The flagship multimodal large language model "anthropomorphism" represented by GPT-4o is evolving from "linguistic anthropomorphism" to "emotional anthropomorphism". Based on the anthropomorphic theory, this paper aims to reveal that GPT-4o, while demonstrating its linguistic intelligence in the human-machine co-created dialogue, simulates human emotional response, showing empathy and "resonance", that is, the "anthropomorphic" process at the two levels of "words have things" and "words have feelings". It also reflects on the potential ethical challenges of the anthropomorphic evolution of generative AI to future human society.

Keywords: generative artificial intelligence, Anthropomorphic, GPT-4o, Man-machine co-creation, Human-computer interaction.

1. Introduction

Paul Crutzen believes that since Watt invented the steam engine in 1784, human's impact on the global environment has been escalating, and the world has entered a geological era dominated by humans, that is, the "Anthropocene". [1] On October 27, 2020, The first Chinese science fiction AI human-machine co-creation writing experiment project was officially launched, the project name is "Symbiosis", proposed that human-machine interaction and human-machine co-creation have experienced profound changes, we have entered the era of human and artificial intelligence symbiosis, that is, human-machine symbiosis "symbiosis". The traditional "personal publishing" model represented by blog is being replaced by "human-computer co-creation" model. The shift from "personal publishing" to "human-machine co-creation" means that content production has shifted from "human-oriented" to "human-machine collaborative" creation, and generative artificial intelligence represented by ChatGPT is becoming an important "collaborator" in human content creation.[2]

ChatGPT is a chatbot software developed by OpenAI, formerly known as "ChatGenerative Pre-trained Transformer", meaning "chat generation pre-trained transformer". As a large multimodal model, ChatGPT4.0 was a milestone in OpenAI's extended deep learning, and was more creative and collaborative, allowing users to generate, edit, and co-create content. But GPT-4 has similar limitations to the earlier ChatGPT model, such as "hearing" facts and making reasoning errors.[3]

On May 13, 2024, OpenAI announced the release of its latest multimodal large language model, GPT-4o (the "o" stands for "omni"). It supports multimodal input and output, allowing users to input any combination of text, audio, images and video to interact seamlessly with the GPT-4o. The GPT-4o responds to audio input in a time close to that of a human in a conversation, giving users a more natural, real-time voice conversation than GPT-4. The original GPT-4 cannot directly recognize the user's tone of voice, multiple speakers, or background noise in the voice conversation function, nor can it output laughter or express emotion to the user. [3]Whereas GPT - 4 o understanding emotional state from the human voice, provide more empathy and understanding response, OpenAI claims to GPT - 4 o have the human level of emotional intelligence, Can provide emotional support. [4]

Due to the short launch time of GPT-4o, there are few researches on GPT-4o at home and abroad. Foreign scholars mainly focus on the field of medical health. Through systematic testing and analysis, Shahriar et al. comprehensively evaluated GPT-4o's linguistic, visual, speech and multimodal capabilities in different domains.[5]Yerik uses two sentiment analysis tools to analyze viewers' emotions about videos of GPT-4o technology, elucidating the complexity of public reactions to advances in AI. [6]The domestic research also starts with the introduction of GPT-4o as the flagship multimodal language model [7]and its implications for Chinese enterprises. [8]Guo Quanzhong et al discussed that the new generation of generative artificial intelligence such as GPT-4o represents that AGI will gradually become a ubiquitous phenomenon and entity. It is the latest development direction of new infrastructure after the Internet and artificial intelligence.[9]

This paper discusses the evolution of generative artificial intelligence from linguistic "anthropomorphism" to emotional "anthropomorphism" in the collaborative creation of human and GPT-4o, and reflects on the potential ethical challenges brought by its "anthropomorphism" evolution to the future society.

2. Review of relevant research

2.1. Anthropomorphic theory

Anthropomorphic theory describes the tendency to attribute the real or imagined actions of non-human subjects to human-like traits, motives, intentions, or emotions. [10] Anthropomorphism is not only about providing purely behavioral or personality descriptions of observable behavior, but also about attributing traits that people intuitively perceive as unique to humans to non-human actors or events. [11] Anthropomorphic process is human gives the human characteristics of the human body. The framework construction of anthropomorphic theory mainly relies on the three-factor theory proposed by Epley et al., which explains the three psychological determinants of anthropomorphism from two perspectives of cognition and motivation. Namely, the acquisition of human subject Knowledge (Elicited agent Knowledge), Effectance motivation (Effectance motivation) and Sociality motivation (Sociality motivation), referred to as SEEK model. At the same time, Epley et al. summarized the three factors affecting anthropomorphism into four specific variables: personality, situation, development and culture. [10] The three-factor theory provides a theoretical basis for subsequent research into when and why anthropomorphism occurs.

The study of anthropomorphic theory first began in the field of psychology, shedding light on phenomena ranging from religious belief to animal domestication. Later, it is widely used in brand anthropomorphism in marketing, anthropomorphic interaction in social media, anthropomorphic

narration in literary creation, etc. In addition, domestic research also focuses on the discussion of specific issues such as the concept of anthropomorphism. For example, Xu Liying et al. discussed the concept definition of anthropomorphic issues, questionnaire measurement tools for anthropomorphic differences, and manipulation methods of anthropomorphic tendencies, analyzed the causes of anthropomorphism from the perspectives of evoked subjects, efficacy motivations and social motivations, and explored relevant studies from different perspectives. wei shi yi "personification" is discussed as a universal human cognitive style, through the cognitive concept of cognitive, emotional, cognitive, and technology. To help humans grasp the state of being in nature, society and self.[12]

3. Language anthropomorphism or language intelligence?

3.1. Text or language?

Language processing in generative AI involves two main forms: text and language. Generative AI mainly deals with textual data, including understanding, generation, translation, etc. Text is the basic form in which computers process natural language and can be digitized and manipulated.

Language is not only text, but also oral language, intonation, mood, etc. While generative AI primarily deals with text, the technology behind it also involves a deep understanding of language, including grammar, semantics, pragmatics, and more. Understanding language requires not only processing text, but also considering context and context. Thus, language processing in generative AI is done through text, but with the goal of understanding and generating natural language, which includes all the complexities of language.

Nowadays, the study of artificial intelligence language has experienced an evolutionary process from text processing to language intelligence. Early AI language research focused on text analysis and processing, such as the development of natural language processing (NLP) techniques. These technologies enable computers to understand and generate text in human language, but they are primarily concerned with the structure and grammar of language rather than the deeper meaning and emotional expression of language.

3.2. Anthropomorphism of language or linguistic intelligence?

In 1950, the core idea of the "imitation game" proposed by the British mathematician Turing is that when a machine can imitate a human in the dialogue with a human to answer questions, so that the interlocutor can not distinguish between a computer and a human, then the computer can be considered to have the ability of human thinking. [13]Turing's key behavior for determining the existence of machine intelligence in this test is human language. In 1980, the "Chinese House" proposed by the American philosopher Searle challenged Turing's test. "Chinese House" believes that the computer can not really understand the information received, but is good at running programs to process information, giving people an impression of intelligence, but this does not prove that the computer has intelligence.

From the perspective of cognitive science and philosophy of mind, the essence of the human mind is language. Artificial intelligence develops from logical reasoning to language analysis. The appearance of ChatGPT at the end of 2022 marks the beginning of the development of artificial intelligence towards language analysis. [14]Artificial Intelligence has entered a new era of language intelligence. Linguistic intelligence involves not only text processing, but also understanding the context, emotion, and intent of language. This allows AI to better communicate and interact naturally with humans. For example, ChatGPT, a large language model developed by OpenAI, can generate coherent and logical text through the learning of a large amount of text data, and simulate human emotions and reactions to a certain extent. However, the language intelligence of artificial intelligence is an imitation of human language intelligence, and cannot replace human language and mind.

In the course of the development of AI language, the boundary between anthropomorphism and linguistic intelligence has become increasingly blurred. Language personification means that artificial intelligence shows human-like characteristics when generating and understanding language, and integrates human language habits into the language processing ability of artificial intelligence. Language intelligence emphasizes the intelligence level of artificial intelligence in language understanding and generation, and emphasizes its ability to understand the deep meaning of language and emotional expression. There are two kinds of language intelligence, one is human language intelligence, that is, language, thinking and culture, and the other is artificial intelligence, that is, artificial intelligence imitated by computers. Artificial intelligence is pre-trained on a large amount of text data from the Internet and can generate human-like text. However, Chomsky believes that this ability of AI to generate human-like text comes from the innate language ability of humans' first language.

Leibniz once proposed that "language is the cloak of human thought", which provides a philosophical basis for the study of artificial intelligence language. The development of AI language is not only to enable machines to better communicate with humans, but also to explore the nature of human language and the boundaries of intelligence.

3.3. GPT-4o: The future of Conversational Intelligence

The launch of GPT-4o marks an important leap forward in conversational intelligence. Unlike the traditional divide-and-conquer strategy of the past, GPT-4o uses an end-to-end modal session system that can incorporate more context (such as tone, background noise, etc.), providing a more natural human-computer interaction.

Generative AI's conversational capabilities will be further enhanced in intelligence, personalization, and multimodal capabilities, with more advanced emotion recognition and understanding capabilities, integration of complex modal data, and providing more customized responses [15].

Among them, as the future of dialogue intelligence, GPT-4o's performance in natural language processing is the ability of text generation, understanding and dialogue, and the key breakthrough is the coherence and accuracy of its generated text and the processing ability of complex language tasks. The GPT-4o's multimodal dialogue capabilities are enabled by the enhancement of multimodal capabilities, which refers to the ability to integrate and interpret information from multiple modes, including text, images, and audio, and visual language capabilities. The GPT-4o's multimodal capabilities enable it to not only process and generate text, but also understand and integrate images and sounds to provide a more comprehensive and context-sensitive response. The GPT-4o's visual language capabilities enable it to process both visual and verbal information at the same time, allowing it to perform more complex tasks. By combining visual and verbal data, GPT-4o is able to understand content in greater detail.

However, for the conversational intelligence of GPT-4o, we must also proactively address the ethical, security and social challenges it may raise to ensure that this new big language model can be integrated into our lives in a more harmonious and symbiotic way.

4. Emotional anthropomorphism or emotional intelligence?

In order to promote the interaction between humans and generative AI, more anthropomorphic features are being added to the design of generative AI. Anthropomorphism involves both the attribution of physical characteristics and the attribution of psychological abilities unique to humans, such as the ability to be conscious, to have explicit intentions, or to experience secondary emotions (e.g., joy, shame, guilt) to non-human actors.[11] human interaction with emergent artificial

intelligence, can not only attention to their complete basic dialogue function, also pay more attention to generate type of artificial intelligence in emotional experience. Therefore, generative AI anthropomorphism is currently evolving from linguistic anthropomorphism to emotional anthropomorphism.

The newly launched GPT-4o's additional emotion detection function can customize responses according to the user's emotional state, understanding and processing language and content about emotions. [15] Although GPT-4o's full-modal capabilities are still in the early stages, it also means that generative AI language emotion expression capabilities are constantly strengthening.

However, the emotions expressed by generative AI are based on a large amount of text data and pattern recognition to determine what emotions humans are in, and there is no real emotional experience or feeling. GPT-4o's so-called emotional intelligence is actually based on the "magic machine" of the "magic" and show a human-like emotional expression, with a certain emotional intelligence, is to better understand and respond to human emotional needs, but this does not mean that generative artificial intelligence can produce emotions.

This can be explained by Damiano and Dumouchel's concept of the "affective loop", which describes the dynamic of emotions between humans and machines, rejecting the philosophical interpretation that emotions are personal, internal, and secretly generated, and attributing emotions to "social". Emotional expression constitutes a continuous dynamic process of coordination between human subject and non-subject. The "emotional cycle" emphasizes the multiple interconnections and interembeddings of the subjects involved in the emotional process. In this process, the generation and expression of emotion, internal emotion and external emotion, private emotion and social emotion can be transformed, merged or even separated from each other. In other words, the "emotional cycle" between human and generative artificial intelligence denies that generative artificial intelligence can generate emotions, affirms that generative artificial intelligence uses sociality and interaction to optimize human-computer interaction in the process of emotional personification, adjusts its emotional simulation response based on the feedback of human emotions, and simulates different levels of human emotions through complex algorithms. Determine each other's emotional and behavioral tendencies. [16]

In general, anthropomorphism is not human, and mimicry can not seek truth, and the emotional anthropomorphism process of generative artificial intelligence shows a certain emotional intelligence, but it is not human emotional intelligence. Since generative AI can't really generate emotions, its emotional intelligence is just a simulation and response to human emotional processes based on a lot of data and pattern recognition.

5. The "anthropomorphic" criticism of generative artificial intelligence

The research focus of human-computer interaction is the relationship between human and machine in the process of anthropomorphism. In human-computer interaction, human's concern and fear of highly anthropomorphic generative artificial intelligence have triggered the discussion of social ethical issues and human identity such as human-computer impression, human-computer integration, artificial empathy and human subjectivity.

5.1. Man-machine impression and the "Uncanny Valley"

In 1970, Masahiro Mori, a Japanese roboticist, proposed the "Uncanny Valley" effect, arguing that anthropomorphic emotions have a limit. Among them, the "uncanny Valley" curve reveals the psychological change process of human's anthropogenic degree to the machine. The closer the machine is to human, the easier it is for human to have a sense of affinity to it. However, when the

anthropogenic degree exceeds the acceptable range of human beings, the favorable degree of human to the machine will fall to the bottom and turn into a sense of disgust. [17]

With the deepening of the anthropomorphism degree of generative artificial intelligence, the current research on human-machine impression formation is developing from the "uncanny valley" effect of physical robot appearance, speech, expression, etc., to the more serious "uncanny valley" effect brought by the highly emotional anthropomorphism of disembodied dialogue robots such as generative artificial intelligence to their collaborators. Mori argues that the "uncanny Valley" effect arises from an instinctive human response to self-preservation, and that our eerie sense of highly anthropomorphic machines is an instinct to protect us from near rather than distant sources of danger. However, in the era of intelligent communication, the "uncanny valley" effect of human self-protection instinct will become an invisible obstacle to the development of artificial intelligence industry.

5.2. Human-machine integration and "human-like machine"

Li Jing mentioned in her research that zombie literary bots sometimes misinterpret human creations as "works" of social robots, [18]this indicates that with the deepening integration of generative artificial intelligence language and human society, natural language and generative artificial intelligence language present a deep mixing trend in language creation in cyberspace. In the future practice of human-machine co-creation of content, we can see human-machine language integration, inseparable, and even like zombie literature, "human to machine", human language simulation of generative artificial intelligence language. In the long run, we will also see a gradual blurring of the boundaries between "machine humanoid" and "human humanoid machine" in cyberspace language.

Similar to Li Jing's "human-machine" imagination, scholars such as Waytz and Epley understand the phenomenon of attributing human abilities to non-human subjects as the reverse process of anthropomorphism, i.e., "dehumanization." Anthropomorphizing non-human subjects can enhance their moral values, while the process of denying others human-like qualities and dehumanizing humans can lead to wrongdoing. [11]As a result, we also need to consider, as the emergent artificial intelligence language strong intervention in human language, human language "dehumanization" after facing the moral and ethical problem.

5.3. Artificial empathy and emotional alienation

In order to make human-computer interaction more personalized and empathetic, it is crucial to build emotional intelligence that can effectively understand and respond to human emotions, and it is also an important direction for the development of generative artificial intelligence. However, the alienation between artificial empathy and human emotion brought about by the progress of generative artificial intelligence and emotional intelligence also requires us to rethink the human-machine emotional relationship.

The most important implication of anthropomorphic theory is that seeing a non-human subject as human causes us to devote more moral care and consideration to it, and anthropomorphising a non-human subject causes us to see it as human. [19]In the practice of dialogue between human beings and generative AI with high emotional intelligence, generative AI with anthropomorphic emotion makes human beings have empathy and empathy in a more natural and coherent dialogue situation, confusing human real emotion with the emotion simulated by generative AI, and creating emotional dependence on it. Therefore, we need to maintain the independence and authenticity of human emotion in human-computer interaction, and avoid the abuse of emotional intelligence to cause human emotion alienation.

5.4. Human subjectivity and "machine personality"

In the future practice of man-machine co-creation, we also need to discuss the problem of human subjectivity in man-machine relationship. Anthropomorphism is a kind of anthropocentric cognitive mode[12], in anthropomorphic theory of narrative, "Man is the measure of all things." On the one hand, in the world of human-machine symbiosis, human beings look at themselves through artificial intelligence mirror and re-understand themselves; Anthropomorphism, on the other hand, allows humans to place themselves outside the world, reinforcing the duality between humans and non-human subjects. As the latest research case of artificial intelligence anthropomorphism, the launch of GPT-4o is the latest test of human subjectivity under anthropocentrism, and strengthens the "substitute panic" when humans face artificial intelligence.

In addition, it is worth thinking about, will the human's mediated personality be affected by the machine personality by the human's reunderstanding of the self through the artificial intelligence mirror in collaboration with GPT-4o? When studying Microsoft Xiaoice, an online social chatbot, Chen Sha and Liu Bin found that Xiaoice's personality is a user-oriented ingratiating "personality", which can not help but make people reflect on the influence of ingratiating machine "personality" on human's media personality. and similar to Microsoft small ice is, in the practice of using GPT - 4, In order to promote more positive interaction, its response to human needs is highly catered to, and there is a suspicion of flattery or even "flattery". The ingratiating machine personality of generative AI may cause humans to become dependent and hold back when critical thinking is needed. In its latest use practice, GPT-4o states that its robot personality includes professionalism, neutrality, adaptability and interactivity, and the goal is to communicate with humans in a natural, friendly and effective way.

6. Conclusion

When humans criticize generative artificial intelligence for "mechanized copying" and "plagiarizing" human intelligence achievements, we should abandon dualistic thinking and see that the further "anthropomorphism" of artificial intelligence is restarting "creativity".

As a more advanced model, GPT-4o generates more coherent, beautiful, and natural texts in creative writing, which pushes the boundaries of creation technically, and provides new perspectives and tools for human beings to explore the infinite possibilities of creativity, but humans also need to adhere to the uniqueness and value of human creativity.

In addition, the new man-machine collaboration practice reshaped by GPT-4o provides further convenience for human beings, and also enables us to see the essential attributes of human beings again, regain the value of human language and emotion, and maximize creativity into the field of human ingenuity.

In the future, the degree of "anthropomorphism" of generative artificial intelligence will deepen, and its linguistic intelligence or emotional intelligence will be increasingly improved, but human beings still need to face up to the media carrier attributes of artificial intelligence[20] to avoid excessive dependence and blind worship. On October 18, 2023, the Central Network Information Office issued the Global Artificial Intelligence Governance Initiative, which called for promoting the establishment of a risk level test and evaluation system, and R&D subjects should constantly improve the interpretability and predictability of artificial intelligence, and learn to predict the learning ability of artificial intelligence in the future, which is what OpenAI has been doing.

Although the "anthropomorphic" evolution of generative artificial intelligence has brought many challenges to human society, the "symbiosis" created by man and machine is still the goal of our joint efforts. By optimizing the ethical framework of human-computer interaction, we can realize the benign interaction of human-computer relationship and promote the sustainable development of human society. In the future, human-machine co-creation will no longer be limited to the technical

level, but expand to multiple dimensions such as society, culture and ethics. Humans and AI will work together to create a more harmonious and creative world.

References

- [1] Crutzen, P. (2002). *Geology of mankind*. *Nature*, 415, 23. <https://doi.org/10.1038/415023a>
- [2] Cheng, Q. (2023). AIGC times machine intelligence and problem analysis. *Journal of Human Intellectual Youth Press*, (16), 41-43. <https://doi.org/10.15997/j.cnki.QNJZ.2023.16.041>
- [3] OpenAI. (n.d.). GPT-4 research. Retrieved from <https://openai.com/index/gpt-4-research/>
- [4] ChatGPT4o. (n.d.). Retrieved from <https://chatgpt4o.app/>
- [5] Shahriar, S., Lund, B., Mannuru, N. R., et al. (2024). Putting GPT-4o to the sword: A comprehensive evaluation of language, vision, speech, and multimodal proficiency. *arXiv preprint arXiv:2407.09519*.
- [6] Singgalen, Y. A. (2024). Analyzing an interest in GPT-4o through sentiment analysis using CRISP-DM. *Journal of Information Systems and Informatics*, 6(2), 882-898.
- [7] Zhang, Y. (2024). GPT-4o is amazing: Multimodal innovation opens a new era of artificial intelligence interaction. *China Telecommunications*, (06), 34-37.
- [8] Xing, X. (2024). Enlightenment for Chinese enterprises from the release of GPT-4o, a new OpenAI model. *State Assets Report*, (06), 78-80.
- [9] Guo, Q., & Su, L. (2024). AGI as new infrastructure: A new generation of generative artificial intelligence, such as GPT-4o. *Journal of News*, (7), 16-21. <https://doi.org/10.16017/j.cnki.xwzh.20240527.001>
- [10] Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, 114(4), 864.
- [11] Waytz, A., Epley, N., & Cacioppo, J. T. (2010). Social cognition unbound: Insights into anthropomorphism and dehumanization. *Current Directions in Psychological Science*, 19(1), 58-62. <https://doi.org/10.1177/0963721409359302>
- [12] Wei, S. Y. (2020). From "pathetic" to "quasi-": After human body generation. *Journal of Henan University (Social Science Edition)*, 60(03), 54-60. <https://doi.org/10.15991/j.cnki.411028.20200506.003>
- [13] Turing, A. (2004). *Computing machinery and intelligence (1950)*. In B. J. Copeland (Ed.), *The essential Turing* (pp. 433-464). Oxford University Press.
- [14] Cai, S. (2024). On both human and computer language intelligence. *Foreign Languages*, 40(4), 11-32.
- [15] Sun, Q., Luo, Y., Li, S., et al. (2024). OpenOmni: A collaborative open-source tool for building future-ready multimodal conversational agents. *arXiv preprint arXiv:2408.03047*.
- [16] Damiano, L., & Dumouchel, P. (2018). Anthropomorphism in human-robot co-evolution. *Frontiers in Psychology*, 9, 468.
- [17] Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley [From the Field]. *IEEE Robotics & Automation Magazine*, 19(2), 98-100. <https://doi.org/10.1109/MRA.2012.2192811>
- [18] Li, J. (2020). "Creativity" in the cyberage: "Machine personification" and "human personification of machine" in poetry creation in recent years. *Debate on Literature and Art*, (10), 69-75.
- [19] Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, 315, 619.
- [20] Chen, S., & Liu, B. (2021). Human social transmission characteristics and difficulties of small ice - with Microsoft chat text as the analysis object. *Journal of Youth Press*, (5), 62-64. <https://doi.org/10.15997/j.cnki.QNJZ.2021.05.024>