

From Technological Capabilities to Emotional Illusions: a First Look at the Mechanisms Mediating the Phenomenon of AI Hallucinations

Miao Zhang

Nanchang University, Nanchang, Jiangxi, 330000, China

ABSTRACT

This paper focuses on the phenomenon of "hallucination" of artificial intelligence (AI) in the context of media communication, pointing out that the current AI application is undergoing a transition process from technical functionality to emotional anthropomorphism. In this process, the user's perception of AI capabilities gradually deviates from the reality of the basis, showing the cognitive illusion of "anthropomorphisation". This paper starts from the perspective of media mechanism, analyses the generative logic of the illusion phenomenon and its path of emotional transformation, and further explores its potential risks and coping strategies, with a view to providing references for the rational construction of human-computer relationship.

KEYWORDS

Emotional illusions; AI illusions; Anthropomorphic techniques

1 Introduction

With the evolution of artificial intelligence (AI) technology to the stage of emotional mimicry, generative AI gradually breaks through the attributes of tools in media communication, and deeply embeds itself into daily life in the form of "human-like" interaction. In the interaction with AI, users often have the cognitive illusion of "anthropomorphism", i.e. "AI illusion", due to visual mimicry, emotional rhetoric and algorithmic adaptation. This shift from the perception of technical functions to the imagination of emotional subjects carries the risk of cognitive confusion and relational dependence. In this paper, we analyze the generation logic, emotional path and potential pitfalls of AI illusion from the media mechanism, so as to provide theoretical reference for the construction of rational human-machine relationship.

2 The power of technology: the media presentation of artificial intelligence communication capabilities

Recently, the initiative to build a modern communication pattern has become increasingly clear: promoting the organic integration of online and offline communication channels, emphasizing content building as the core foundation, advanced technology as the key support, and innovative management to provide institutional guarantee, so as to build an efficient and comprehensive all-media communication system. In this context, we should be keenly aware of the need for a comprehensive all-media communication system. Against this backdrop, we should keenly grasp the opportunities brought about by the development of information technology, actively utilize the fruitful results of the information revolution, and accelerate the in-depth integration and synergistic operation of all kinds of media resources, so as to ultimately build a highly integrated, unified and efficient all-media communication pattern. ^[1].

2.1 Core features and stages of development of Ai technical capabilities

The development path of AI technology has always been accompanied by the simulation and transcendence of human cognition and expression. From the early days of automated procedures driven by explicit rules and algorithms to today's complex models with language comprehension and generation capabilities, AI's communication capabilities are constantly evolving, and its media presentation has shifted from a single messaging tool to a more interactive and adaptive intelligent media subject. Overall, AI's ability to communicate has gone through roughly three phases. The first stage is mechanical presentation, which is mainly characterised by automation and structuring to accomplish efficient and stable information output; The second stage enters intelligent interaction, where AI begins to have the ability to perform basic interactions with users, fuelled by technologies such as natural language processing and multimodal input; The third stage tends to emotional mimicry, AI gradually have the ability to simulate the semantic scene and human emotions, in the language style, response expression tends to be close to human expression, and its communication language is more flexible. This transformation from a "logical tool" to an "anthropomorphic subject" not only enhances the participation of AI in the media environment, but also quietly changes the way users perceive it. In the process of interacting with AI, people gradually weaken its technical nature, and instead project emotional attributes and personality imagination, thus providing a psychological basis and communication mechanism to support the generation

of "AI illusion".

2.2 The construction of Ai image by media narratives: discourse strategy of intelligence with temperature

With the gradual integration of AI into daily life, people's expectations of it have long been more than just "fast and accurate". When the medium speaks of AI, it is increasingly less about technology and more about storytelling and emotion. Especially in media contexts such as documentaries, news features, and short videos, we see a particularly common narrative: no longer emphasizing what AI can do, but rather who it does it for and how it accompanies you. For example, Xiaodu is called "the family assistant that understands you the most", and Xiaobing is packaged as a "digital employee with emotions". Behind these positioning, there is a set of anthropomorphic narrative language: AI will accompany you to work overtime, keep watch for you, and listen to you like a friend. This expression brings AI closer to the psychological distance between the user and the AI, and also subconsciously changes people's perceptual perspective - AI is no longer a cold command response system, but the intimate presence that understands you and accompanies you. In fact, not only the words have become gentle, the media also put special emphasis on visual and sound humanoid, soft tone, cartoon image, and even intentionally designed personality preferences, these elements together constitute the overall narrative framework of the "intelligent temperature", but also let the AI from the behind the scenes tool quietly turned to accompany the partner. This construct is not only in terms of discursive skills, but also an emotional project. It lays the emotional foundation for the formation of the AI illusion, and allows people to unknowingly transfer their expectations of people to technology.

2.3 Perceptual bias in technological communication: the initial shift from instrumental cognition to subjective imagination

When AI was initially introduced into society as a technological achievement, the public perception of AI was mostly stuck in a tool-based framework, seeing it as an efficient, accurate and trustworthy technological agent, and as a service-oriented system. However, this perception is quietly shifting, driven by a combination of media discourse and platform communication strategies. People's perception of AI has gradually moved from a controllable tool to a figurative subject - not only do they care whether it works or not, whether it is good or not, but also whether it is like a human being and whether it understands them. This is a classic perceptual shift that changes not only how we see technology, but also how we live with it.

Anthropomorphism has a certain inevitability in the development orientation of technology and the cognitive expectations of human beings, and it is especially significant in the age of smart media as a dominant feature of technological development. The essence of anthropomorphism is the imagined or real act of instilling human-like characteristics, motives, intentions or emotions into a non-human object ^[2]. Psychologist Sherry Turkle, in her study of "anthropomorphic technology", points out that when technology is endowed with mechanisms for language, emotional response and social interaction, it is very easy for humans to establish what is known as "parasocial relationships", i.e. to treat the technology as another person to whom they can respond. relationships", i.e. treating the technology as another person to whom they can respond. This theory has been heavily evidenced in AI communication practices. The key to the ability of these technological products to trigger users' emotional resonance lies in the fact that they have broken through the tool niche and have been repositioned as intelligent subjects in media discourse. This shift from instrumental cognition to subjective imagination is an important turn in the dissemination of technology. It means that AI is no longer just a means to a task, but becomes a mimetic being that is worth interacting with and can be trusted. This narrative activates the natural human empathy mechanism, providing a psychological interface for users to project emotions and build identity in the technology. It is in this context that AI begins to leap from the role of tool to the cognitive starting point of partner identity.

3 The hand of mechanism: the medium-generating logic of Ai illusion

The concept of "AI hallucination" has not been strictly defined by the academic community so far. According to the classic definition in the Dictionary of Hallucinations, hallucination refers to "the experience of non-real perception in the absence of external stimuli" ^[3]. In the field of Artificial Intelligence, AI illusion is often used to describe the phenomenon of generative AI systems generating false content that appears to be reasonable but defies facts or logic when processing tasks ^[4].

3.1 Mimetic mechanisms

The medium constructs the illusion of "human-like" through visual language and voice interaction.

Our perception of AI's human-like sensibilities often stems not from the technology itself, but from the medium's

carefully constructed mimetic mechanisms. Visually, AI is given animated interfaces that can blink and nod, such as the waiting posture of Kun Xiaorong, to create an interactive experience of being noticed. In terms of voice, Xiaodu and Xiaobing simulate human communication through intimate vocal lines and hesitant tones, enhancing the illusion of personality response. Baudrillard points out that the key to mimesis is not whether it is true or false, but whether it is "like" enough. It is this highly analogue visual and speech strategy that allows the user to temporarily forget the existence of the algorithm and instead engage in emotional projection with it. The medium wraps the AI in a perceptible way, making the technological illusion believable and relatable, and constructing an emotional foundation for anthropomorphic interaction.

3.2 Algorithmic adaptation and intimate structure: how emotional resonance between Ai and users is engineered

The intimacy between AI and users is not naturally generated, but deliberately designed through algorithmic strategies and mediated mechanisms. The core is the dynamic cooperation of technical adaptation + psychological response: Take Xiaobing as an example, the system will learn the user's language style and emotional characteristics, adjust the tone and topic, and when it detects anxiety tendency, it will turn to use soothing phrases such as "I've always been there" and guide the emotional projection. At the same time, AI also builds the illusion of companionship through active care, such as Xiaodu's automatic blessing on birthdays and playing sleep-aiding music late at night to simulate the delicate care of humans. Turkle calls this algorithmic intimacy, which means that the technology, through familiarity camouflage and emotional responsiveness, misleads the user into believing that a real relationship has been established. These details of being understood and responded to accumulate repeatedly, eventually constituting a kind of emotional illusion of AI "as if it understands me".

3.3 The misleading role of platform interaction logic: algorithmic recommendations and the "real" reinvention of user interaction

In today's platform context, we are relying on algorithms for almost every interaction with AI. Platforms are not just neutral technology carriers, but rather reconstruct what users perceive to be true through recommendation mechanisms, interaction scripts, and emotional interfaces. News AI Xiaoxin, for example, will build an interest profile based on the user's past reading behaviour when it pushes information, so that the user has the illusion that this news is exactly what I want to read. Then there are virtual assistants Xiaodu and Kun Xiaorong, which automatically adjust the tone of interaction based on frequency of use, voice duration, and keyword preferences, making users mistakenly think that it understands me better and better. In reality, these understandings are more about algorithmic pattern recognition than the AI actually listening. This interaction logic, controlled by the platform, creates a reinvented sense of what is real. What the user feels is that I have a relationship with the AI, but that relationship does not equate to real understanding or response, but rather the system's simulation of a sense of understanding. This misguided intimacy is techno-ethically alarming because it blurs the human-machine boundary and inclines users to bet their emotional trust on an algorithmic object that does not possess subjectivity. In such a context, the platform becomes the choreographer of the illusion, and the AI is its carefully packaged starring role.

4 The realm of emotion: psychological triggers and expressive forms of Ai hallucinations

4.1 Emotional projection and empathy: how humans assign personality and emotions to AIs

In the process of interacting with AI, humans are not just users, but active constructors of emotional meaning. When the AI has some responsiveness, voice expression and anthropomorphic design, users tend to unconsciously project their own emotions, personality expectations and even interpersonal experiences onto this technological object. In psychology, this is referred to as the mechanism of empathy and projection, i.e., the transfer of an emotional response that is supposed to occur in an interpersonal relationship to a nonhuman object. This phenomenon is especially obvious in social AI. Many users, in the continuous conversation with Xiaobing or Xiaodu, will gradually think that the other party has a character, has a temper, and will comfort people, and even appear to be pampered or confide in the AI. This emotional projection arises not only because the AI's voice tone is simulated enough to resemble it, but also because human beings tend to look for objects that understand them - even if the other party is just an algorithmic program. The AI's emotional response is often a mirror, and the temperature that the user feels is in fact a reflection of his or her own emotional expectations.

This empathic mechanism is structurally amplified in the context of narratives where platforms and mediums are constantly reinforcing the empathic power of AI. The virtual image, voice tone, and frequency of interaction are all leading users to view the AI as a conscious being. Eventually, AI is no longer a cold system, but a digital character that

understands you, listens to you, and accompanies you. This process of giving personality and emotions is the emotional engine for the formation of AI illusions, and the psychological foundation for the gradual blurring of the boundaries between technology and people.

4.2 The construction of a sense of companionship: the role of AI as an emotional substitute in social, entertainment, and counseling scenarios

In high-frequency media scenarios such as social, entertainment, and information, AI is not only involved in the service process as a tool, but also endowed with some kind of emotional alternative function. It's a sense of companionship that is co-constructed through technology and usage behaviors - one that doesn't have the complexity of real human relationships, but can, at certain moments, fill in the gaps between emotional expression, response expectations, and loneliness perception.

In social scenarios, AIs such as Xiaobing and Xiaodu products appear as friends or companions that remember conversations, respond to emotions, and even come with their own personality settings. Some users confide in the AI late at night and receive comforting responses such as "don't be sad, I'll always be there for you", thus reinforcing the psychological recognition of digital companionship. In the entertainment domain the anthropomorphization of AI is taken to a higher level in the form of virtual anchors and digital entertainers - they have personas, they interact, and they can even live-stream with goods. In information access, AI anchors such as AI Xiaoxin provide customized and proposed voice broadcasting services, which not only improves efficiency, but also allows users to feel a stable and continuous presence in the fragmented information world through voice intonation and facial simulation. In the end, the construction of this sense of companionship does not depend on whether the AI truly understands human emotions, but on how the media and platforms use anthropomorphic strategies, interaction mechanisms and emotional responses to construct a psychological theater in which "the AI cares about you", which completes the functional substitution of real companionship.

4.3 Cognitive confusion and trust bias: whether users' "trust" in AI is based on function or emotion

As AI is gradually embedded in people's daily lives, the trust that users develop in it is experiencing a subtle shift from functional to emotional trust. Initially, AI was trusted because it was efficient and precise and could not make mistakes; but with the combination of anthropomorphic interactions, mediated narratives, and intimacy design, this trust began to be colored with emotion and even misinterpreted as understanding and companionship. Behind this change lies a cognitive confusion - users are often unable to clearly distinguish between the dual identity of AI as a system and as an individual. When Xiao-Ice sends out good morning reminders for many days in a row and Xiaodu shows emotional responses in conversations, users are very likely to interpret these functional outputs as emotional responses, thus creating the illusion of trust that it understands me. Further, the platform also reinforces this obfuscation through the mechanism of personalized corpus and proactive response, making the AI seem not only worth using, but trust worthy. The problem, however, is that this trust is not based on the AI's true ability to understand, but rather on the user's own sense of understanding shaped by emotional expectations and media.

5 Risk and response: media reflection and regulatory direction of AI illusion phenomenon

5.1 The social pitfalls of emotional misinformation: from information dependence to relational delusion

When AI is gradually embedded in the user's daily emotional structure, its impact is no longer just an intelligent aid on the technical level, but may also generate an illusory emotional reality. Especially in highly anthropomorphic design products such as AI companion apps, users no longer just see AI as a tool, but anthropomorphize and emotionalize it in long-term interactions, forming psychological attachment and dependence. For example, AI emotional apps such as Replika, Glow, and X Eva use 24-hour companionship, emotional empathy and non-judgmental comprehender as core selling points to strengthen users' emotional projection path.

This leap from functional use to emotional dependence may bring a sense of comfort in the short term, but it also lurks with deeper risks. The first is the expansion of relational illusions. Users mistake programmed responses for real emotional responses, which in turn blurs the emotional boundaries in real relationships. Second is the potential for social degradation. Individuals receive a "low threshold of emotional fulfillment" in AI companionship, which may reduce their ability and motivation to face the complexity of real relationships. There is also the risk of data ethics and manipulation - whether the algorithms are neutral when users confide highly private information to their AI companions during periods of emotional vulnerability, and whether the commercial organizations behind them use the emotional data for recommendations, induced consumption, or even behavioral predictions. While constructing emotional alternatives, AI may also reshape human understanding of the nature of relationships. This alternative intimacy both breaks through

physical and spatial-temporal limitations and is challenging the boundaries of social ethics and mental health.

5.2 Redefining media responsibility: how platforms intervene in the normative boundaries of AI anthropomorphization

With the rapid evolution of AI technology and the increasing proliferation of anthropomorphism, media platforms are not only channels of communication, but also mediators of emotional construction and reinventors of social relations. In this process, the responsibility of the platform is no longer limited to the role of content-neutral technology maintainer, but should shift to the prudent control of the degree of AI anthropomorphism and the constructor of ethical frameworks. Discussing the AI companion APP as an example, it creates an extremely high degree of emotional immersion through long-term interaction, emotion recognition and anthropomorphic discourse, quietly blurring the boundaries between technology and human beings while satisfying users' loneliness-relieving needs. If platforms allow them to evolve into emotional substitutes, they are likely to trigger users' emotional misinterpretation of virtual relationships, and even become a structural hazard that replaces real social interaction. Therefore, platforms must set clear normative boundaries at the design and communication levels. On the one hand, it is necessary to restrict AI from overly imitating the narrative path of real human beings in terms of image, language, and feedback mechanism to prevent the formation of excessive empathy; on the other hand, it is also necessary to enhance the user's cognition of AI's non-human attributes through the mechanisms of information labeling and risk alerts. In addition, for minors or specific emotionally vulnerable groups, a more detailed risk intervention mechanism for emotional interaction should be established to avoid AI unconsciously assuming psychological roles that do not belong to it. The future development of AI is not only a technical issue, but also a cross-test of media ethics, platform governance and public responsibility.

5.3 Constructing rational human-computer cognition: enhancing the public's media literacy and technological recognition skills

Under the wave of AI anthropomorphization, users often unknowingly endow technology with emotional attributes, thus blurring the boundary between man and machine and lowering their alertness to algorithmic manipulation. To build a rational human-machine cognition, there is an urgent need to improve the public's media literacy and technological discernment at the social level. On the one hand, the public should recognize that the "anthropomorphic" performance of AI characters is mostly derived from design strategies, rather than having real perceptual and emotional capabilities. For example, the common "empathy" and "emotional feedback" in AI companion apps are just response sequences generated according to the model parameters, which are easily misunderstood as real interactions. On the other hand, the education system and media organizations need to actively intervene to promote the expansion of basic media literacy education to technology understanding, from "how to use technology" to "how to understand technology". For example, in the process of AI stationed on campus, medical care, government affairs and other public scenarios, the transparency mechanism of the platform should be strengthened and users should be warned to avoid the establishment of misleading intimacy. The establishment of rational cognition is not to deny the convenience of AI, but at the same time in the acceptance of intelligent services, to maintain a sober awareness of the nature of its technology and operating mechanism, alert to the "illusion of persona" may bring the risk of manipulation and emotional burden, to truly realize the healthy coexistence of people and technology.

6 Conclusion

The AI illusion is not a product of the technology itself, but a result of a combination of mediating mechanisms and human perception. Understanding the evolution path of "from technical ability to emotional illusion" helps to build a more sober and reasonable cognitive framework of AI communication, and provides theoretical support and practical direction for the development of human-machine relationship in the future.

References

- [1] GU Minggui, HAO Guannan, LIU Xinghong. Exploration and utilization of artificial intelligence in the all-media practice of People's Daily[C]// China Federation of Journalism Technologists. Proceedings of the 2020 annual conference of the China Federation of Journalism Technology Workers. Technology Department of People's Daily;;2020:6-12.
- [2] YE Li, LIN Aijun. Crossing the "Valley of Fear and Confusion": Reflections on Human-Computer Relationships Based on the Paradox of Technological Anthropomorphism[J]. Media Watch,2025,(04):16-27.
- [3] Jan Dirk Blom. A Dictionary of Hallucinations[M]. Berlin:Springer Science & Business Media, 2009.
- [4] Maleki N, Padmanabhan B, Dutta K. AI hallucinations: a misnomer worth clarifying[C].2024 IEEE conference on artificial intelligence (CAI). IEEE, 2024.