

ChatGPT and the future of humanity's trust in technology

HISTORY

By YAN HONGXIU

As digital technologies keep evolving, they have exceeded the traditional level of merely replacing dangerous, harsh and tedious human work with a more adaptive approach to meet the diverse needs of human beings. Generative AI allows, in a rather subtle way, for individuality to be expressed and suppressed at the same time. Recently, ChatGPT has emerged as a vivid case of introducing AI into people's daily life with a convenient and easy-to-use Q&A mode, which pushes the two dimensions of expression and suppression to a new level. Expression means that human-technology interaction offers more possibilities to unleash individuality. Suppression means that the environment constructed by technological rationality keeps human beings framed in the logic of technology.

Trust and technology

In this context, are the answers provided credible? With the two-way human-technology integration, technology has gradually become strong evidence of trust. In that case, can the logic of technological rationality serve as the foundation of trust?

For example, universities such as the Institut d'études Politiques de Paris and the University of Oxford have banned the use of ChatGPT in studies and exams where its use would be considered cheating, as its ability to generate text comparable to that of real people has not been clearly evaluated and described. It is evident that this attitude highlights the threats ChatGPT poses to the existing educational trust system and the impending new challenges to the trust relationship between teachers and students. Subsequently, the technology community has proposed to develop ever more advanced technologies to detect whether ChatGPT is being used for studies and exams in order to identify academic misconduct. In fact, this approach pushes back trust formation precisely to the logic of technological rationality again.

But is this logic valid? According to Niklas Luhmann, in the face of technological developments, it is not to be expected that scientific and technological development in the course of modernization will bring events under control and replace

As the logic of technology becomes stronger, the intelligent assistant of the past may become subject-to-be.

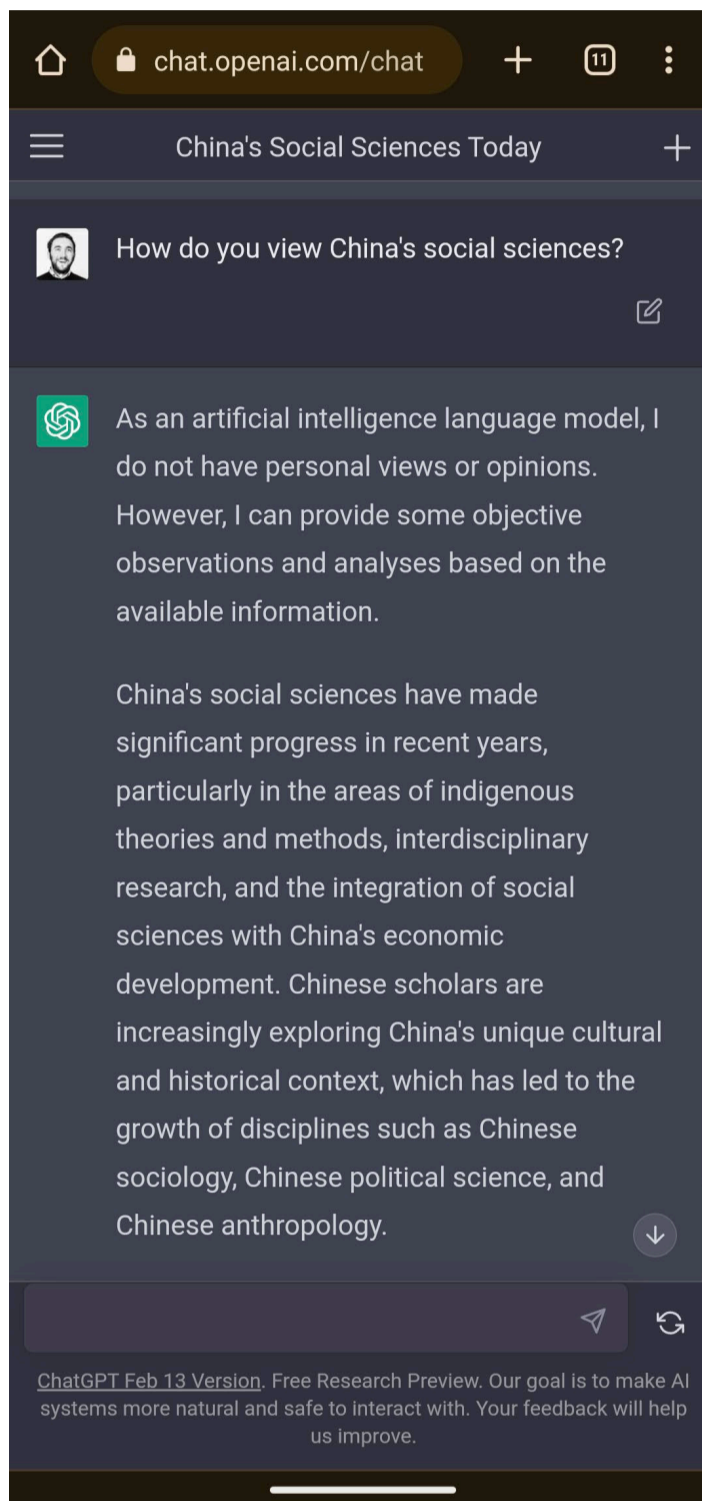
trust as a social mechanism with mastery over things. Instead, one should expect trust to be increasingly in demand as a means of enduring the complexity of the future that technology will generate.

To put it simply, while human survival requires the use of technology, and technology can be an element in the formation of trust, the formation of trust should not depend on technology. Instead, technological developments, especially with the inherent uncertainties of technology, demand trust as a coping strategy. Therefore, the formation of trust can be achieved with the help of technology, but it must not be confined to technological rationality.

Q&A mode, one-way trust

ChatGPT is an intelligent tool based on data mining and data enrichment techniques and is able to generate knowledge through human-machine interaction. It can provide, or more accurately, generate answers based on its existing database and the questions it is asked. When giving answers, ChatGPT would first declare that it is AI, and notify users of the time nodes of the database update. It would also warn users that the answers it provides may not be accurate, and that it may produce harmful instructions or biased content. It is under this condition that ChatGPT works with users in a Q&A communication mode, thus forming a knowledge chain with users' questions and AI's answers.

After long-term training, ChatGPT may be able to create what Piero Scaruffi referred to as "structured environment" in his 2016 book *Intelligence is Not Artificial*. The intelligent navigation systems we use today have created such an environment, where a person inputs his or her destination, and the rest unfolds in accordance with the logic of technology. If the human user relies on the system to complete the task as he/she is in the process of doing it, his/her decisions and behaviors can be seen as human-machine interactions predicated on the logic of the technology. As a consumer of the technol-



A Q&A conversation between a human user and ChatGPT Photo: PROVIDED TO CSST

ogy, the user here is involved as a participant. Once the logic of the technology functions, the human enters what Scaruffi describes as a predictable state in which the human does not need to think too much. If this state is taken to the extreme, humans would become outsiders of the technological loop, while the logic of technology forms a closed loop of trust and moves directly to one-dimensional trust.

As Max Horkheimer and Theodor W. Adorno said: "Where the development of the machine has become that of the machinery of control, so that technical and social tendencies, always intertwined, converge in the total encompassing of human beings, those who have lagged

behind represent not only untruth." Thus, the starting point of trust is technological reason, and its formation is also confined within technology. Of course, this extreme state can be attributed to people's imagination about technology, perhaps an over-imagination, and so this kind of trust is still hard to achieve.

Nevertheless, with the development of ChatGPT, two inevitable phenomena may occur. First, while using ChatGPT to search for factual knowledge, the results of data training may show that the knowledge is homogeneous and the logic one-dimensional. The homogeneity of knowledge is caused by constantly formatted database and standardized training, while the one-dimensional

logic is the result of the above-mentioned homogeneity and the human's role as a questioner. One-dimensional logic leads to one-dimensional trust.

The second phenomena is that when we use ChatGPTs in a confrontational way, the environment of trust will be damaged, resulting in trust itself becomes one-dimensional.

Trust, technological dependence

In the face of technology, trust becomes a game between human experience and technological reason. As human society becomes increasingly "intelligentized," this game is increasingly dominated by the logic of technology. In recent years, more and more people are calling for human-centered design, technology for good, ensuring technology serves people and improves their wellbeing. This has revealed the growing logic of technology and our prudence about it. As the logic of technology becomes stronger, the technological dependence of trust will change accordingly. And this kind of change will gradually lead technology to a superior position. In other words, the intelligent assistant of the past may become subject-to-be.

In particular, the credibility of ChatGPT is prerequisite for winning trust. The promotion of credibility can be divided into two methods. The first is the direct way, namely by using correct answers to improve its accuracy. The other is the other way around, namely helping the AI to detect deceitful or false information. The latter method also requires data feeds, and will challenge the technological dependence of trust while generating trust's vulnerability, unreliability, and falsity.

As Kevin Kelly wrote in *What Technology Wants*, "My hope is that it will help others find their own way to optimize technology's blessings and minimize its costs." "What technology wants" is a question leading to human happiness, and the goal, or essence, is "what humans want."

Trust is a prerequisite for a functioning human society. What kind of trust do humans want in the context of generative AI technology becoming a part of the human environment? This question represents both an eternal pursuit for trustworthy generative AI and an examination of technological dependence.

Yan Hongxiu is a professor from the School of History and Culture of Science at Shanghai Jiao Tong University.