

Hallucinating ChatGPT and the innocent lawyer

Kaito Shimoura (Ritsumeikan Asia Pacific University, Japan); kaito.shimoura.apu@gmail.com

Tam-Tri Le (Phenikaa University, Vietnam); tri.letam@phenikaa-uni.edu.vn

June 26, 2023

Earlier in June 2023, Steven A. Schwartz, a New York lawyer with 30 years of experience, faced a court hearing due to submitting documents containing fake cases to the court [1]. While such incidents are not rare in the legal landscape, what made this man a sensation on the Internet is the fact that the supporting evidence he presented was written by ChatGPT.

Schwartz was the lawyer in charge of a case involving a man suing an airline over an alleged personal injury. In a brief submitted to the court, Schwartz cited previous court cases. However, at least 6 cases in the submitted document did not exist in any record. These fake cases were hallucinated (made up) by ChatGPT.

But why did this experienced lawyer believe ChatGPT's responses so easily?

According to Schwartz's explanation, he did not comprehend that Chat-GPT could fabricate cases [2]. He and his team asked ChatGPT whether these cases existed or not, and ChatGPT answered, "Yes". Such a methodological and rigorous verification process!



Illustration. A lawyer who trusts artificial intelligence

This story shares some uncanny similarities with the incident where a Texas university professor failed his whole class for cheating after asking ChatGPT if it had written the papers submitted by the students and believed the bot's affirmative responses [3].

Schwartz and his team believed ChatGPT's answers without putting in their effort to manually check. As a result, Schwartz not only had to face a court hearing and put his career on the line but also became the target of ridicule all over the Internet.

In the human mind, trust is an energy-saving mechanism of information processing, allowing for quicker reception and evaluation through prioritized channels [4]. In the chaotic infosphere of modern society, humans may feel the urge to rely on the trust mechanism to function more efficiently, driven by competitive survival pressures [5]. On the one hand, we are aware that Schwartz pursued self-interest, wanting to gain an edge in his career with less spent time and energy, just like many other people in life. On the other hand, he sacrificed his whole career and reputation just to make the public laugh (what a selfless lawyer).

Under the enormous amount of information in today's digital world, the filtering systems of each individual and the whole human society need to adapt accordingly, with new fitting

skills and appropriate mindsets [6]. Automation can greatly help but should never reach the point of replacing human thinking altogether.

While many fear that artificial intelligence may rebel against humanity, some (like the aforementioned lawyer) willingly believe whatever the machine says in the field of their own human expertise. Irrationality is very rational in the mind of these people, which makes reality sometimes more ridiculous than fiction can ever be [7].

References

[1] Armstrong K. (2023). ChatGPT: US lawyer admits using AI for case research. BBC News.

[2] Weiser B, Schweber N. (2023). The ChatGPT Lawyer Explains Himself. The New York Times.

[3] O'Neill N. (2023). Texas professor flunked whole class after ChatGPT wrongly claimed it wrote their papers. New York Post.

[4] Le TT, Nguyen MH, Vuong QH. (2022). Trust in mindsponge: A new perspective on information reliability. In Vuong QH, La VP, Nguyen MH (Eds.) The mindsponge and BMF analytics for innovative thinking in social sciences and humanities (pp. 67–86). De Gruyter.

[5] Vuong QH. (2023). Mindsponge Theory. De Gruyter. https://www.amazon.com /dp/B0C3WHZ2B3

[6] Nguyen MH, et al. (2023). Examining contributors to Vietnamese high school students' digital creativity under the serendipity-mindsponge-3D knowledge management framework. Thinking Skills and Creativity, 49, 101350.

[7] Vuong QH. (2023). Meandering Sobriety. https://www.amazon.com/dp/B0C2TXNX6L

