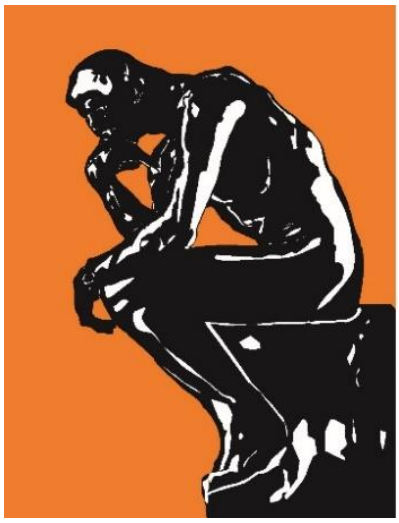
Is artificial intelligence the harbinger of a new natural absurdity era?

Minh-Hoang Nguyen [1,*], Quan-Hoang Vuong [1,2]

[1] Centre for Interdisciplinary Social Research, Phenikaa University, Hanoi, Vietnam

[2] A.I. for Social Data Lab (AISDL), Hanoi, Vietnam

* Correspondence: hoang.nguyenminh@phenikaa-uni.edu.vn



November 08, 2024

[*Original working draft v1 / Un-peer-reviewed*]

"My colleagues, they study artificial intelligence; me, I study natural stupidity."

— Amos Tversky (1937-1996); *Meandering Sobriety* (2024a)

The rapid development of AI is enabling it to reach or even surpass human capabilities in various intellectual domains, such as reading comprehension, image, language, and speech recognition (Kiela et al., 2023). Consequently, AI is increasingly integrated into human decision-making processes across fields like commerce, business, healthcare, education, and governmental activities. Many intellectuals and experts anticipated that AI and its applications would reshape the world and leverage human intellectual capacity.

The immense potential of AI and society's high expectations for it have sharply increased the demand for advancements in computational power, data, and algorithms. In particular, computational power has exceeded predictions based on Moore's law, which suggests that computational power doubles roughly every two years. Since the emergence of deep learning in the early 2010s, computational power has accelerated, doubling approximately every six months from 2010 to 2022 (Sevilla et al., 2022). This skyrocketing demand has boosted the market value and investment in chip makers; for instance, NVIDIA's market capitalization surpassed $3 trillion in June 2024, making it the most valuable company globally. However, along with rising computational power and chip market values comes increased energy consumption. According to the International Energy Agency (2024)'s *Electricity 2024: Analysis and Forecast to 2026* report, energy-intensive data centers for AI and cryptocurrencies are projected to double their electricity consumption by 2026, roughly equaling that of Japan, the world's fourth-highest energy consumption country.

These numbers and milestones have fueled current AI buzzwords centered around chips, market valuations of chip companies, and energy consumption—factors that significantly impact climate change. One question arises:

- Since when has the advancement of human intellectual capacity become measured by the amount of data, chips' processing capabilities, financial investment in chip makers, and the amount of energy consumed by AI?

The answer to this question is ambiguous, but it is clear that AI development is susceptible to humans' subjectivity. For AI to progress, it must be trained on vast amounts of high-quality data. For example, ChatGPT's models are trained using a variety of data sources, including publicly available data (e.g., industry-standard machine learning datasets and web crawls), proprietary data from data partnerships, and human feedback (from AI trainers, testing teams, employees, and users who opt in for model improvement). All of this data is collected, generated, and provided by humans, who are limited by their inherent fallibility, often leading to misinformation, errors, and biases in the data. Even scientific content, which is typically reliable and precise, includes seriously flawed studies that sometimes require retraction (Nguyen & Vuong, 2024).

Human subjectivity influences not only the quality of data but also the AI model creation processes (e.g., through data selection and model design), resulting in AI outputs that can mirror human errors and biases (e.g., under-weighting rare events, forming illusory correlations, falling into attentional traps, and displaying reference-dependent risk

preferences) (Rich & Gureckis, 2019). Not to say, AI models are increasingly trained to mimic the psychological and decision-making processes of humans for better training effectiveness and efficiency. In particular, application-screening algorithms often reflect the biases of previous human reviewers, perpetuating discrimination against women and minority groups. Predictive policing algorithms, relying on historical crime data, risk creating feedback loops of over-policing over certain groups of people. In healthcare, AI algorithms trained on unrepresentative data may produce inaccurate predictions for specific patient populations, leading to inappropriate care.

In other words, as AI learns from humans, it can sometimes not only fail to mitigate but also reproduce and amplify human errors and biases. With AI's strengths in scalability, speed, and automation, misinformation, errors, and biases can be quickly, widely, and consistently amplified, potentially creating serious systematic problems (Vuong & Nguyen, 2024). Once deployed, these AI systems can have long-term influence over the learning processes, decision-making, and behaviors of those who interact with these automated systems. As suggested by the experiments of Vicente and Matute (2023), biased recommendations provided by AI systems can influence people's decisions and persist even after they stop using AI. The risks posed by AI spreading misinformation, creating biases, and distorting reality become even more critical at a time when people increasingly trust machines over humans (Sundar & Kim, 2019).

Then, is viewing human intellectual capabilities and development as determined by AI truly in the name of technological progress, or is it merely a proxy for natural absurdity?

While it is unclear whether AI has accomplished the goal of helping humanity to overcome natural stupidity, the state of limited understanding and processing abilities, it is evident that AI is pushing society into a new era of natural absurdity—where humans abandon logic and reasoning and mindlessly rely on the output generated by AI's black-box systems powered by data, models, and chips, for their thinking, decisions, and behaviors (Nguyen, 2024; Vuong, 2024b). So where are the spirits of pushing beyond the limitations of machinery with sheer intellect—the kind embodied by great minds like Max Planck, Ludwig Boltzmann, Albert Einstein, Werner Heisenberg, and Paul Dirac to enlighten humans' understanding of reality, only to arrive at 2024, when Nobel Prize in Physics is awarded to the training of AI?

AI has strengths that humans cannot replicate, such as scalability, speed, and automation, but this must not mean that we depend entirely on AI for intellectual advancement. For a future where humans coexist with advanced AI, we must acknowledge the existence of intrinsic natural stupidity and absurdity of humans and take them into consideration. Otherwise, increasing the information and processing capabilities of AI may amplify the magnitude of humans' poor decisions and their consequences, but not the other way around.

References

International Energy Agency. (2024). *Electricity 2024: Analysis and forecast to 2026*. https://www.iea.org/reports/electricity-2024

Kiela, & et al. (2023). Test scores of AI systems on various capabilities relative to human performance. *Our World in Data*. https://ourworldindata.org/grapher/test-scores-ai-capabilities-relative-human-performance

Nguyen, M.-H. (2024). Natural absurdity: How satirical fables can inform us of a vision for sustainability? https://philarchive.org/rec/NGUNAH

Nguyen, M.-H., & Vuong, Q.-H. (2024). Artificial intelligence and retracted science. *AI and Society*, 1-2. https://doi.org/10.1007/s00146-024-02090-z

Rich, A. S., & Gureckis, T. M. (2019). Lessons for artificial intelligence from the study of natural stupidity. *Nature Machine Intelligence*, *1*(4), 174-180. https://doi.org/10.1038/s42256-019-0038-z

Sevilla, J., Heim, L., Ho, A., Besiroglu, T., Hobbhahn, M., & Villalobos, P. (2022). Compute trends across three eras of machine learning. 2022 International Joint Conference on Neural Networks (IJCNN), Padua, Italy.

Sundar, S. S., & Kim, J. (2019). Machine heuristic: When we trust computers more than humans with our personal information. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, Scotland.

Vicente, L., & Matute, H. (2023). Humans inherit artificial intelligence biases. *Scientific Reports*, *13*(1), 15737. https://doi.org/10.1038/s41598-023-42384-8

Vuong, Q.-H. (2024a). *Meandering sobriety*. https://www.amazon.com/dp/B0C2RZDW85

Vuong, Q.-H. (2024b). *Wild Wise Weird*. https://www.amazon.com/dp/B0BG2NNHY6

Vuong, Q.-H., & Nguyen, M.-H. (2024). Further on informational quanta, interactions, and entropy under the granular view of value formation. https://doi.org/10.2139/ssrn.4922461