

Editorial: ChatGPT and the ethical aspects of artificial intelligence

Many definitions of artificial intelligence (AI) have been proposed since John McCarthy coined this term in a workshop at Dartmouth College (Hanover, New Hampshire) in 1956. [Bellman \(1978\)](#), for instance, defines AI as “the automation of activities that we associate with human thinking, activities such as decision-making, problem solving, learning.” Another interesting definition of AI is the one made by [Rich and Knight \(1991\)](#): “The study of how to make computers do things at which, at the moment, people are better.” Besides the proper definition, the human kind has always been interested in challenging questions concerning intelligence. The greater processing capacity of small devices and the significant increase in people’s interactions with the Internet (the so-called Internet of things – IoT) led to the intensive development of sophisticated AI techniques. In addition to it, the availability of a huge amount of data – also known as big data – and the development of sophisticated mathematical/statistical modeling approaches also contribute to the literature on machine learning (ML) and AI, giving the emergence of the “data-driven culture”, whereby many decisions are made based on machines’ data processing. The more data, the more a computer can learn and extract hidden relationships that we are unable to process.

From a business and managerial perspective, data-driven decision-making became a competitive advantage. Companies are interested in extracting knowledge from data in order to produce better decisions. This process is performed by data science and data analytics professionals that make use of many AI and ML techniques. At the same time, the development of more sophisticated data-driven AI and ML methods is of great interest. The finest example of this race is the emergence in November 2022 of Chat Generative Pre-trained Transformer (ChatGPT), a chatbot developed by OpenAI, an American AI research laboratory. A chatterbot (or chatbot) is a computer program that tries to simulate a human being in conversation with people. Its goal is to answer questions in such a way that people have the impression that they are talking to another human being and not to a machine. After being asked questions in a natural language, the program queries a knowledge base and then provides an answer that tries to mimic human behavior. The word chatterbot was coined by Michael Mauldin – founder of Lycos, Inc. and creator of the first chatterbot Julia – in 1994 to describe these conversational robots. ChatGPT gained a great deal of attention for its detailed responses and articulate answers across many domains of knowledge. That reveals a high capability of performing adequately many tasks that are generally performed by human intelligence, like writing a song, a poem, a thesis, painting a picture, etc. The chatterbot uses language models and has been fine-tuned (an approach to transfer learning) by using both supervised and reinforcement learning techniques.

One could quote a plethora of benefits from the use of a technology like ChatGPT. In many companies, for example, daily activities can be more productive when using this technology to understand concepts, summarize information, build and evaluate recommendations, optimize processes through algorithms, among others. However, the indiscriminate use of AI



has already been discussed in the recent literature (Dowling & Lucey, 2023; Müller, 2021; Neubert & Montañez, 2020; Safdar, Banja, & Meltzer, 2020) as well, therefore intensifying concerns about the ethical aspects. The use of machinery/automatic knowledge generation excludes some traditional learning acquisition processes, that is, the human ability to think, to reason, to plan, to perceive, to adapt and to perform induction, deduction, logic as well as to communicate. These are ordinary abilities of human intelligence. Not exercising them goes against the ability to innovate, for example. In addition to this philosophical issue, there are also unwanted effects such as the generation of misinformation, the dissemination of fake news and even the promotion of plagiarism. ChatGPT consists of an associative system, whose content is not generated by real knowledge. The use of ChatGPT does not provide information checking (uneven factual accuracy) as well. The technology may generate incorrect information and may produce harmful instructions or biased content, as pointed out by the ChatGPT platform itself [1].

The whole issue reminds us of Searle's "Chinese Room". Searle (1991) describes a room with a person, called the operator, who dominates the English language. Many baskets with Chinese ideograms are placed in the room, as well as a book of rules, written in English, on how to combine Chinese ideograms. The operator receives a sequence of Chinese ideograms through an entrance opening (on the room's wall) and, by consulting the rule book, they combine these input ideograms with others in the baskets, therefore composing a new sequence. This new sequence is then passed through an exit opening (in the room wall). Although the operator does not know it, they are answering questions in the Chinese language.

Searle argues that there is a marked difference between this operator and a person who masters the Chinese language and answers the same questions without using the rulebook. The first one, the operator, is just following syntactic rules. The second is associating semantics (meaning) to what it is being done. Thus, the second is doing much more than the first. Searle's main conclusion is that computers – as syntactic machines – can replace the operator. However, computers can never think, because thinking involves semantics. With that, computers can never be intelligent.

As a matter of fact, Searle's conclusion can be – and is indeed being – revisited. In spite of that, the main challenge is to develop machines that think with the objective of assisting in the progress and evolution of human intelligence. The benefits of disruptive technologies such as ChatGPT are numerous. Even so, ethical aspects must also be taken into account. It is always important to remember that the oracles of Ancient Greece influenced, to some extent, the politics of the time, sometimes even resulting in decisions which culminated in wars.

In scientific research, ChatGPT can play a significant role as a powerful assistant. The question which comes to mind, nonetheless, is as follows: Will we be able to still claim the produced research as our own?

Leandro Maciel

*Faculty of Economics, Administration and Accounting, University of São Paulo,
São Paulo, Brazil*

Note

1. Source: <https://openai.com/blog/chatgpt/>. Access on Feb. 14th, 2023.

References

- Bellman, R. E. (1978). *An introduction to artificial intelligence: Can computers think?*. Boyd & Fraser.
- Dowling, M., & Lucey, B. (2023). ChatGPT for (finance) research: The bananarama conjecture. *Finance Research Letters*, 103662. doi: [10.1016/j.frl.2023.103662](https://doi.org/10.1016/j.frl.2023.103662).

Müller, V. C. (2021). Ethics of artificial intelligence and robotics. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2021 Edition). Available from: <https://plato.stanford.edu/archives/sum2021/entries/ethics-ai/>

Neubert, M. J., & Montañez, G. D. (2020). Virtue as a framework for the design and use of artificial intelligence. *Business Horizons*, 63(2), 195–204. doi: [10.1016/j.bushor.2019.11.001](https://doi.org/10.1016/j.bushor.2019.11.001).

Rich, E., & Knight, K. (1991). *Artificial intelligence*. McGraw Hill.

Safdar, N. M., Banja, J. D., & Meltzer, C. C. (2020). Ethical considerations in artificial intelligence. *European Journal of Radiology*, 122, 108768. doi: [10.1016/j.ejrad.2019.108768](https://doi.org/10.1016/j.ejrad.2019.108768).

Searle, J. R. (1991). *Minds, brains and science - the 1984 Reith Lectures*. London: Penguin Books.