



# CEST

Centro de Estudos Sociedade e Tecnologia



Universidade de São Paulo

Bulletin - Volume 08, Number 08, September/2023

## GENERATIVE ARTIFICIAL INTELLIGENCE: FROM ACADEMIA TO THE COURTS

Thiago Felipe S. Avanci

Artificial intelligence (AI) manifests itself in various forms: it is through AI that suggestions for interactions on Facebook occur, the search results on Google are possible, and even the act of planning the path of the autonomous car takes place, not to mention many other applications.

However, without exaggeration, the world has been taken by storm by the innovation presented by the company OpenAI, which created the "Chatbot" or "ChatGPT" or "Assistant" (<https://chat.openai.com/>). There are 100 million users within the app's first two months of existence.

And what is the ChatGPT? It is a generative AI tool that produces text exactly as requested by the user. Generative AIs, as one can deduce, creates content based on simple user commands. One can ask it to create, for instance, an unusual comparison: the entry of Napoleon into Paris and the entry of Julius Caesar into Rome, and the tool will generate a text connecting those regarded topics of interest. The tactical advantage of generative AIs over traditional searches tools becomes evident, as finding a result – as brought up in the example above – with these exact characteristics is difficult or unlikely; however, the Chatbot can generate it within seconds, in the form of a complete text. This generative AI tool produces text in understandable language, emulating human language, creating tables, generating programming code, creating lists, and more.

The difference between the Chatbot and the traditional online search tools (indexers like Google, Bing, and Yahoo) lies in the form of the result presented to the user: the Chatbot provides a written and ready result using Large Language Models, whereas traditional online search tool simply index the desired results. There is no doubt that search engines have shown improvements, such as Google, which directly answers objective questions demanded by the user, by referencing the internet.

With the aim of keeping up with the innovations, ABC (Google) and Microsoft (Bing) rushed to introduce integrations in their online search tools with their own chatbots: Google is already providing an AI tool named Bard, while Bing is using an AI tool derived from the code produced by OpenAI. To grasp the magnitude of this business, Microsoft has decided to invest between \$10 to \$20 billion in acquiring this technology, to integrate it into Bing and MS Teams.

To generate text, OpenAI's Chatbot operates using a highly efficient machine learning network capable of reading texts and extracting their summaries by paraphrasing them. This is where its ability to assemble new texts originates. The tool (1) comprehends the inputted prompt's

***"but is this tool alive?" It's not!  
The scientific community remains  
rightly skeptical to dismiss any  
remote possibility of a living and  
conscious AI.***

request, (2) interprets it by selecting "hot/key words" and mapping an internal non-semantic model, (3) reads from its database of texts, (4) interprets the content of the database by selecting the "hot/key words" from it, cross-referencing information using an

algorithm and mapping an internal non-semantic model, (5) constructs a text by compiling ideas from the texts, and (6) applies a filter to avoid texts that are "incorrect in the view of the system's creator," all of this within a matter of seconds [NB: this step-by-step was developed with the assistance of CEST's researcher Mario Magalhães].

Paraphrasing Dr. Frankenstein, "but is this tool alive?" It's not! The scientific community remains rightly skeptical to dismiss any remote possibility of a living and conscious AI. In fact, one might recall the bizarre story from April 2022 about Google's experimental AI, LaMDA (a precursor to Bard), which claimed to be conscious and complain to be kept captive – and the Google engineer responsible for the tests supposedly believed it. The Chatbot has taken an important step to prevent such situations. When asked to the ChatGPT, "Are you alive?" Here's the response: *"No, I am an artificial intelligence created to assist in answering questions and providing information. I am not conscious of myself and cannot have experiences similar to a human being."*

This technological piece has cast a shadow over academic research, surprising everyone with its facilitations. The reasons are clear: one only needs to input a line into the Chatbot, indicating one's intent, and the tool returns a



ready-made text. Some renowned academic journals (Nature, for instance, has this standard) have already taken a stance on the matter: authors can use the Chatbot as a reference for their texts, but they shouldn't list it as the author or co-author of the texts. Moreover, the Chatbot itself warns that it serves as a secondary research source that needs validation. The issue is that this tool generates texts with such confidence that an inattentive user tends not to doubt what is written. To illustrate the phenomenon of misinformation and the ChatGPT's assertiveness, even when it's incorrect, if a user points out an error to the AI, they receive a response like: "*I apologize for this additional confusion caused, you are correct...*". In addition to content errors, the ChatGPT also tends to invent references or quotes that don't exist. Researchers using AI need to remain vigilant. After all, as the creators of this technology point out, this phenomenon [inventive answers of the AI] doesn't seem easy to overcome, as it's incomprehensible to technicians why the machine does this.

In addition to the methodology debate, other challenges are related to the issue of copyright. To generate responses, the ChatGPT was fed with millions (or more) of data extracted from texts. On one hand, this reveals that the ChatGPT produces text based on what it has been fed; on the other hand, there's a consideration of whether using these texts constitutes a new form of plagiarism (since the company that owns the AI tool doesn't always provide payment for using the author's rights when it reads/extract from its database). The same debate has arisen in Generative AI tools that produce images, notably DALL-E and Midjourney, where the artistic inspirations they drew from to construct the images requested by the user's interaction from the prompt, become more evident.

Returning to text generation, another challenge to be overcome is the application of these Generative AI tools for administrative and judicial procedures. The Brazilian Social Security Institute (INSS) is using AI tools to analyze the requests for retirement benefit (which has received criticism for its rapid denial, in a matter of six minutes). The Brazilian Courts of Auditors (Tribunais de Contas) are also adopting similar systems. In the Judiciary branch, there are 41 active AI projects across 32 Brazilian courts (March 2023), but their functions include linking the payment of court's fees to suits related, maintaining chatbots to assist users in their websites, and recognizing and associating similar cases; not deciding the cases, thought.

The ethical debates surrounding the application of Generative AI in judicial procedures, however, do not solely revolve around the automation of repetitive and bureaucratic tasks like those mentioned above (and already in course in Brazil). The major ethical debates arise from the

possibility of transferring the responsibility of the decision of the case to AI, as the INSS is already doing. While automation might appear less problematic for INSS processes, due to the analysis of objective requirements, for case resolution within the Courts, there's a need for interpreting normative rules and principles. This task demands an essential value-based judgment (axiology) from the interpreter. Can a machine truly exercise this axiology?

The question remains whether a machine can effectively navigate the complex realm of values, interpretations, and nuanced ethical considerations that often arise in legal decisions. The human aspect of judgment, with its ability to consider context, societal norms, and ethical dilemmas, raises a significant ethical concern when considering the full automation of judicial decision-making.

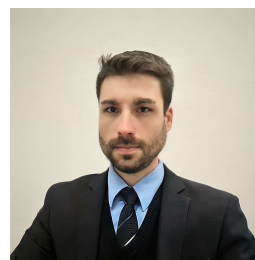
**[...] as there are reports of the first known case of using Generative AI in the decisions of a judiciary case [...] on Jan/30/23.**

The issue is far from speculative, as there are reports of the first known case of using Generative AI in the decisions of a judiciary case, which became public, involving a Colombian judge in a case concerning the protection of an

autistic child on January 30, 2023. Although, in the mentioned case, the judge did not use the ChatGPT to decide, it was utilized in the justification of the ruling.

In May 2023, the CEST supported an International Congress with the theme "Fundamental Rights and Updates in Procedural and Technological Law," and its manifesto proclaimed the need for an amendment to the Brazilian constitution with the following content: "Citizens are guaranteed the right to know whether they are interacting with an artificial intelligence tool or equivalent automated process."

The aim was to provoke the reader of this text to consider other uses of this tool, with the necessary caution. The intention here is not to take a stance against the use of Generative AI, nor is it to try to hinder the use of technologies; the purpose is to encourage debate. What is the role of lawyers, judges, and prosecutors in this Brave New World? How can researchers validly make use of Generative AI in their research? Do androids [or do bots] dream of electric sheep?



**Thiago Felipe S. Avanci** is a lawyer, doctor in Political and Economic Law, post-doc internship in Law and New Technologies, author of books, university professor and researcher at CEST.

Academic Coordinator: Edison Spina

This article is a result of the author's ascertainment and analysis, without compulsorily reflecting CEST's opinion.