ARTIFICIAL INTELLIGENCE IN RESEARCH PSYCHOLOGY: ETHICAL CONSIDERATIONS

LUKE GRILES
Iscte-Instituto Universitário de Lisboa, Lisboa, Portugal

Keywords: Artificial Intelligence; Scopus AI; Ethics.

GOALS

This chapter aims to:

(a) give an introduction to artificial intelligence;

(b) illustrate the relevance of Scopus AI as a tool;

(c) share ethical considerations for using AI in research.

INTRODUCTION

More and more, readers hear about the burgeoning field of artificial intelligence with accomplishments that sound almost like science fiction. Artificial intelligence (AI) is a crucial topic across many disciplines, with implications ranging from economics and efficiency to energy and ethics. With the advancement and adoption of generative AI, it has become rapidly more useful, and its implementation is far more ambiguous than could have been imagined even a few years ago. Generative AI is key to the future, but in what capacity and at what cost?

As a human-created technology, generative AI is at the service of humanity. While concerns about artificial general intelligence (AGI) that eventually takes control of humanity are enthralling, they remain relegated to science fiction. Nonetheless, focusing on the outlying apocalyptic scenarios can actually prove far more imminently dangerous. Generative AI in 2025 is very much a reality, and rapid adoption and dissemination of the technology raise substantial ethical concerns.

At the same time, AI provides a considerable upside in the face of these risks. With the propagation of information as never before, AI offers a means to condense content into understandable summaries and syntheses that can provide key insights that could otherwise be lost. The use of AI may still result in information being overlooked or lost, but this problem can be meaningfully mitigated through the thoughtful use of

technology. Before the Internet, information in an obscure library or database was inaccessible to most of the academic communities based elsewhere. AI offers the opportunity for a similar leap forward by giving researchers leverage to better understand the massive amount of information they can now access.

The ethics of generative AI and its applications in research is a challenging but crucial area for academics to approach. To do so here, we will consider the history and underlying technology of artificial intelligence, demystifying the "black box" of AI, while at the same time acknowledging that the technologies it is based on are necessarily obscure. We will then consider a specific tool, Scopus AI, and look at its application and use as well as limitations. Finally, we will turn to the far more tenuous consideration of ethics in AI and chart a path into the complications that surround this, but also the guideposts that can make AI a meaningful tool for ethical scientific research.

THE HISTORY OF AI

While AI is a topic of much conversation in the present, its history stretches back decades, and, in other forms, the concept of artificial intelligence has been approached by humanity for millennia. The foundations of generative AI can be traced back to the 1940s, before computers or the Internet (Strawn, 2024). Vannevar Bush (1945) envisioned a "memex" that could hugely aid the recall of scientists and preserve knowledge in ways novel for the time. In the 1970s, the first chatbots were created using rule-based pattern matching, and though they did not generate novel text, they could sometimes trick the user into thinking they were interacting with a sentient being (Gupta et al., 2024). From the 1980s onward, applications such as image and language processing began, with advances in machine learning increasing the novelty of the output. However, these were slowed by funding freezes that led to the AI winter (Gupta et al., 2024). Nonetheless, starting in the 2010s, large language models (LLMs) and neural networks gave a new breadth to the potential outputs (Fan et al., 2024). GPT 4.0, in development for years by Open AI, was launched in 2023, with substantial disruptions to the entire space following rapidly (Fan et al., 2024).

MACHINE LEARNING

To better understand AI, it is instructive to first consider some of the underlying technologies. Machine learning is based on computer models that make predictions based on provided data (Kufel et al., 2023). Mathematical models underlie these algorithms, which are used to predict outcomes and thereby "learn". They can be improved through coding the model or improving the dataset the model uses (Kufel et al., 2023). Models can be supervised (where the problem is related to the dataset) or unsupervised (where the model learns without expected values; Kufel et al, 2023). k-Nearest Neighbor algorithms are popular in data mining and work by predicting the correct class based on its difference from the training set. As an example, imagine that a new animal is found. Scientists would take into consideration its environment, its behavior, and similar species in order to create a taxonomic identification for it.

Similarly, k-Nearest Neighbor algorithms can use a known dataset to predict the values of unknown datapoints.

DEEP LEARNING

Artificial intelligence differs from other tools like search engines in the means by which it understands data. While search engines use algorithms to sort results, deep neural networks (DNNs) use multiple hidden layers to process information from input to output. Increasing the hidden layers improves output, but also computational and memory requirements (Kufel et al, 2023). Research indicates that deep neural networks are better at predicting unknown data; deep reinforcement learning integrates with DNNs to learn through trial and error and maximize cumulative reward signals. Transformer neural networks are widely used in natural language processing, using a self-attention mechanism to understand the importance of elements in sequence (Kufel et al, 2023). These various types of deep learning allow artificial intelligence to provide unique outputs and a broader range of solutions to novel problems.

NEURAL NETWORKS

Artificial neural networks mimic the brain by processing and transmitting information. The data goes through the network, being processed by successive layers (Kufel et al,2023). Neural networks are useful for predicting, specifically regarding natural language processing, but also in image recognition. There are various types, including perceptron networks, layered networks, recurrent networks, convolutional networks, gated recurrent networks, and long short-term memory networks (Kufel et al, 2023). The crucial element in neural networks is the weighting of tokens. These allow different weights (a means of assigning values) to be assigned to tokens (parts of a word), changing the probability that certain outcomes will result. For example, consider the letters: i-u-t-q-e. In English, we could use context to determine whether quite or quiet or quit would make more sense. These contextual factors are similar to the weights used in training AI.

LARGE LANGUAGE MODELS

Large language models (LLMs) learn through neural networks and are trained on huge amounts of text data, often based on transformers (Fan et al., 2024). Since 2018, there has been a paradigm shift in natural language processing by focusing on attention. Typically, these are based on Google's bidirectional encoder representations from transformers (BERT), OpenAI's generative pre-trained transformer (GPT), or Meta's LLM Meta AI (LLaMa; Fan et al., 2024). Crucially, these are pre-trained models, meaning that they are heavily reliant on the data on which they are trained. These models use transformers to assign weights to different tokens (Fan et al., 2024). The applications of LLMs are broad, with some key examples being text generation, translation, and summarization. These broad applications build on the same basic frameworks but use different weighting of tokens to produce various results.

WHAT IS GENERATIVE AI?

All of the above may sound very technical, but even a cursory use of ChatGPT reveals a very intuitive interface. Generative AI uses the advancements in LLMs, neural networks and machine learning to create novel solutions. It's a rapidly growing sector with broad applications from conversation generation and code development to drug design and automation (Gupta et al., 2024). Generative AI is trained on huge, varied data sets, and creates novel outputs in diverse fields. There are many positive outcomes, but ethical considerations also arise with the creation of deepfakes and hallucinations (Balasubramaniam et al., 2024).

Showcasing just how confusing these hallucinations can be, consider an excerpt from an article I prompted ChatGPT to write:

> "**Evolutionary Origins and Phylogeny** *The giggleberry tree belongs to the Fictaceae family, which is hypothesized to have diverged from common angiosperms around 20 million years ago (Smith & Jones, 2020). Fossil evidence suggests that early members of this lineage inhabited humid forests, with genetic analysis indicating a close relationship to the Luminifera genus (Brown et al., 2019). The defining trait of Fictus ridiculus—its bioluminescent fruit—likely evolved as an adaptation to enhance seed dispersal*."
>
> Brown, T. P., Green, R. A., & White, L. M. (2019). Phylogenetic relationships of the Fictaceae family: A molecular approach. *Journal of Botanical Evolution, 34*(2), 145-160.
>
> Smith, J. K., & Jones, B. D. (2020). The evolution of temperate angiosperms: A fossil record analysis. *Evolutionary Botany, 12*(1), 89-105.

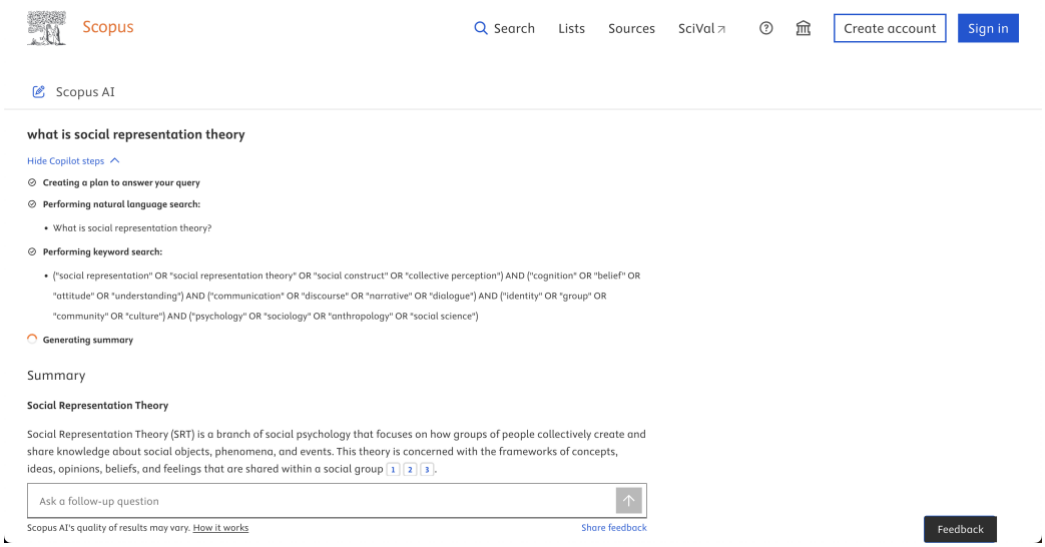
Clearly, something called a "giggleberry tree" does not exist and neither do Brown et al., 2019, or Smith & Jones, 2020. Darwin (1859), on the other hand – cited elsewhere in the article – is a real source. ChatGPT does not suffer if it plagiarizes or creates false sources, but scientific research does. This highlights a specific need for both a thoughtful consideration of the ethics used in approaching AI in research and which tools are used in what ways in the research space.

WHAT IS SCOPUS AI?

Helping to solve the challenges of the integration of artificial intelligence into research, Scopus AI is a tool based on Open AI, but with specific guardrails in place. Scopus AI is built on top of the already expansive and respected Scopus database. Typically, research in Scopus is done through queries, but Scopus AI allows for natural language prompts that can return full-text answers. The focus is to help researchers with navigating fields where they are unfamiliar with the topic, especially helping those early in their careers or working across disciplines. For example, rather than using the typical queries to search the Scopus database, a user can instead input a natural language question like "*What is social representation theory*?" (Figure 1) This will generate an

output as demonstrated below, and also allows users to expand on the search by adding follow-up questions like "*How does this theory integrate with cultural psychology*?"

FIGURE 1 I SCOPUS AI SEARCH



Scopus AI was developed in collaboration with the research community, with the intention of solving the most prescient problems faced by researchers. Scopus AI is driven by abstracts to learn from Scopus content from the past 20 years and delivers results that are immediately usable, with citations on generated content linking back to the original source, allowing for synthesis research (Aguilera-Cora & Feldner, n.d.).

ETHICS IN AI

Artificial intelligence is a nascent technology that has already found myriad practical applications. As rapid adoption takes place, questions of ethics arise. Specific to psychology research, we can begin to consider how to navigate this increasingly complex landscape. One key consideration is attribution, which becomes more difficult with ChatGPT as a "co-author" (Harker, 2023). LLMs are based on massive amounts of data, and ChatGPT is not the author, but its contributions are impossible to fully attribute to their original sources. Scopus AI and similar resources help to navigate this by providing attributions (which are typically more reliable than those that can be hallucinated in broader generative AI). Nonetheless, there is still a margin for error.

When using artificial intelligence in research, the weights used in various models can play a crucial role in the output generated, but in more intuitive systems, these are not open to the researcher and are outside the scope of a typical research process. What

emerges when considering ethics in AI is that an entirely new approach must be developed for this new class of tools. Traditional considerations such as attribution and rigor must continue to be valued, but new skills such as a rudimentary understanding of the technologies that underlie our research methods must also be developed. To do so requires thoughtful collaboration between researchers and computer scientists, each with an interest in understanding both fields.

RESOURCES

- Harari, Y. N. (2017). Homo Deus. Harper Perennial
- Kissinger, H., Schmidt, E., & Huttenlocher, D. (2021). The age of A.I.: And our human future. Murray.
- Elicit: The AI Research Assistant. (n.d.). https://elicit.com/

FINAL CONSIDERATIONS

Artificial intelligence is an area with much potential, but not without its pitfalls. Crucial to forward progress and a meaningful use of the rapidly developing tools is a cross-disciplinary collaboration in understanding how research tools are used and what underlying technologies researchers implicitly trust when using them. Trust is paramount. The failure of AI to deliver in the past has led to funding freezes that stopped it short of its potential, but poor reliability can cause many issues before it is even caught. Thoughtful use of artificial intelligence can provide substantial leverage for psychology researchers, and, like technological advances of the past, AI can be integrated in a way that expands the impact that research can make.

ABOUT THE AUTHOR

LUKE GRILES is a final year student of the Master's Degree in Psychology of Intercultural Relations at Iscte-University Institute of Lisbon. He completed his curricular internship at LAPSO-Psychology Laboratory in 2023/2024.

REFERENCES

Aguilera-Cora, E., & Feldner, D. (n.d.). *Gen AI use in the research workflow* [Video]. Elsevier Researcher Academy. https://researcheracademy.elsevier.com/research-preparation/research-design/gen-ai-use-research-workflow?dgcid=STMJ_1731565097_PUBC_TRAIN

Balasubramaniam, S, Chirchi, V., Kadry, S., Agoramoorthy, M., Senthilvel, Kumar, S. & Sivakumar, T.. (2024). The road ahead: Emerging trends, unresolved issues, and concluding remarks in Generative AI—A Comprehensive review. International *Journal of Intelligent Systems*, 013195. https://doi.org/10.1155/2024/4013195

Bush, V. (1945). As we may think. *The Atlantic Monthly*, 176(1), 101–108. http://dblp.uni-trier.de/db/journals/theatlantic/theatlantic176.html#Bush45

Fan, L., Li, L., Ma, Z., Lee, S., Yu, H., & Hemphill, L. (2024). A bibliometric review of large language models research from 2017 to 2023. *ACM Transactions on Intelligent Systems and Technology, 15*(5), 1–25. https://doi.org/10.1145/3664930

Gupta, P., Ding, B., Guan, C., & Ding, D. (2024). Generative AI: A systematic review using topic modelling techniques. *Data and Information Management, 8*(2), 100066. https://doi.org/10.1016/j.dim.2024.100066

Harker, J. (2023) *Science journals set new authorship guidelines for AI-generated text.* National Institute of Environmental Health Sciences. https://factor.niehs.nih.gov/2023/3/feature/2-artificial-intelligence-ethics

Kufel, J., Bargieł-Łączek, K., Kocot, S., Koźlik, M., Bartnikowska, W., Janik, M., Czogalik, L., Dudek, P., Magiera, M., Lis, A., Paszkiewicz, I., Nawrat, Z., Cebula, M., & Gruszczyńska, K. (2023). What is machine Learning, artificial neural networks and Deep Learning?—Examples of Practical applications in medicine. *Diagnostics, 13*(15). https://doi.org/10.3390/diagnostics13152582

Prompt Engineering. (n.d.). OpenAI Platform. https://platform.openai.com/docs/guides/prompt-engineering

Strawn, G. (2024). Where Deep Learning and Generative AL started: Masterminds of Artificial Neural Networks-McCulloch, Pitts, and Rosenblatt. *IT Professional, 26*(3), 99–101. https://doi.org/10.1109/MITP.2024.3404229