The Illusion of Scientific Language
Lisa Thomann

As artificial intelligence (AI) continues to advance, it is influencing not only research and teaching, but also the understanding of facts, language and truth. The scientific community faces both the promises and perils of these innovations, prompting fundamental questions about how knowledge is constructed and communicated. This essay serves as a starting point for this discussion, exploring whether and how the rise of AI is transforming the language of science, with the aim of sparking deeper discussions about this critical shift.

The functioning of ChatGPT relies on algorithms that are designed to interpret natural language inputs and generate appropriate responses, which may either be pre-written or newly created by the AI.[1] This technique allows the AI to perform similarly well or even better than students in many subject-specific exams at university level, including a simulated bar exam with a score around the top 10% of test takers.[2] Hence, it seems plausible that ChatGPT has the potential to serve as a beneficial tool for various academic tasks, including automatically generating drafts, summarizing articles, and translating languages. These capabilities could greatly enhance the efficiency and ease of writing work for academic purposes. However, the question remains if the language produced by ChatGPT is capable of attaining all relevant aspects of the language used in science.[3]

THE LANGUAGE OF SCIENCE

Science, it has been taught for decades, aspires to show, not to tell. Yet science has never been speechless, in fact, at the latest since the 17th century, science has been all about communication.[4] Scientists collaborate by sharing data, debate over explanations, prepare lectures, write down results from experiments and exchange ideas with colleagues. No data set can exist in isolation. Scientists must convey their interpretations of the data and provide logical arguments to support their claims. The effectiveness of this professional task is determined by the scientist's mastery of language and rhetoric. [5] Language is therefore just as crucial to science as it is to literature or religion. But there

is an evident distinction between the language chosen in academic contexts and other fields. The language in academic contexts, commonly known as "scientific language," has certain characteristics that distinguishes it from other types of languages.[6]

LANGUAGE LIMITATIONS

Contrary to the common belief, scientific language does not merely report facts but rather involves interpretation of data and development of ideas. Nevertheless, scientists aim to identify the most efficient and objective language to explain phenomena. The primary objective of scientific language is to minimize any potential connotations that may reflect or create cultural biases or emotional attachments. Therefore, scientific language must be rational rather than arbitrary or subjective.[7] In comparison to other uses of language, the words chosen ought to represent some actual reality and aim to establish what is generally true or constant under certain circumstances, and what is broadly applicable across different contexts. While scientific language may not be as emotional or overtly persuasive as poetic, political or religious language, it is still a human-made system, and its creators cannot fully control the impacts their language choices have on others. Moreover, neither terms nor descriptions are capable of entirely capturing a phenomenon. Rather, scientific language is an approximation as precise as possible to the reality investigated. The standards for scientific language correspond with the general standards for science.[8] The scientific language is therefore closely linked to the respective scientific methods, values, and norms.[9] This leads us to the primary feature of scientific language: its existence is inextricably tied to science itself. This is because language, although not being the thing itself, shapes the thing it describes.[10] Accordingly, language and science are mutually dependent and must adapt to each other. This is where the first major discrepancy lies between the language produced by ChatGPT and scientific language. As a language model, ChatGPT can only adopt one part of what is needed for scientific language, namely the language itself. This means that ChatGPT can adopt some of the *formal* characteristics that make up scientific language (just as grammar, plausible sounding, selection of terms), but it lacks the *constitutive* and most important aspect to actually produce scientific language. ChatGPT can imitate scientific language, but not produce it in the sense of science. However, scientific language is an inherent part of the scientific process and not something that can be fabricated in isolation.[11] The language produced by ChatGPT is thus an optical illusion, without the corresponding content. Being a language model, ChatGPT lacks the ability to reflect on the content it generates, leading to the possibility of producing not only misinformation but also perpetuating biases.[12] No information generated by ChatGPT is verified, sources are not cited, and it is not subjected to peer review by the scientific community. To be considered a scientific language, the language and its associated content must meet these criteria.
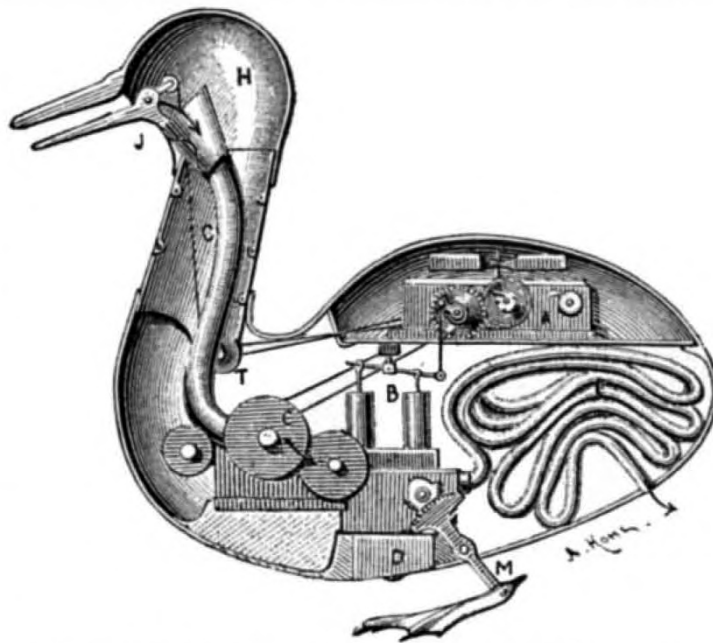
INTERIOR OF VAUCANSON'S AUTOMATIC DUCK.

A, clockwork; *B*, pump; *C*, mill for grinding grain; *F*, intestinal tube; *J*, bill; *H*, head; *M*, feet.

Abbildung 1: Just as Vaucanson's mechanical duck mimics digestion, ChatGPT imitates the structure of scientific language. Both captivate with their form, yet neither the duck digests nor does ChatGPT produce genuine scientific knowledge: it remains an illusion of functional depth.

RESPONSIBLE WORDING

Scientific language not only adheres to high formal standards but also carries an essential responsibility as it has the power to shape society, alter perspectives, and push the boundaries of knowledge. Scientific language therefore comes with a significant responsibility, and it is crucial to select words with utmost care. ChatGPT is limited in its ability to produce scientific language that reflects the values and norms of the scientific process. To conduct research, study and teach responsibly, one must be aware of this limitation. However, by understanding these restrictions, one can still benefit from ChatGPT's abilities in regard to the formal aspects of scientific language – but the formal aspects only.

Autor*in
Lisa Thomann studied Cultural Studies at the University of Lucerne and is currently completing her Master's degree in History and Philosophy of Knowledge at the ETH Zurich.

Seminar
A version of this text was produced as part of a collective project called "Text-Generating AI in Learning and Research: Student Perspectives," as part of the course "Digital Ethics and Politics" at ETH Zurich in the Spring 2023 semester.

Redaktionell betreut von
Ines Barner, Niklaus Schneider and Christina Tuggener

Abbildungsverzeichnis
Abbildung 1:
https://commons.wikimedia.org/wiki/File:Digesting_Duck.jpg

Literaturverzeichnis
[1] Michele Salvagno, Fabio S. Taccone, and Alberto G. Gerli: "Can Artificial Intelligence Help for Scientific "Writing?", in: *Critical Care* 27/75 (2023), pp. 1–5, p. 1.
[2] OpenAI: "GPT-4 Technical Report", p. 5-6, https://arxiv.org/pdf/2303.08774 (2023).
[3] "Science" is a very complex and broad field. This paper aims to refer to a general validity of language and science, which, however, is not always and not equally applicable to every scientific discipline.
[4] Jen Christiansen, Lorraine Daston, and Moritz Stefaner: "The Language of Science", in: *Scientific American* 323/3 (2020), p. 26.
[5] In this paper, the focus lies on the written word, as ChatGPT cannot yet speak (other than Siri, Alexa etc.).
[6] Carol Reeves: *The Language of Science*, London and New York: Routledge (2005), pp. 1–2.
[7] Carol Reeves: *The Language of Science*, London and New York: Routledge (2005), pp. 9–10.
[8] These standards include procedure and methods, repeatability, transparency, publication, peer review etc.
[9] Samuel J. McNaughton: "What is Good Science?", in: *Natural Resources & Environment* 13/4 (1999), pp. 513–518, p. 513.
[10] Rudolf Virchow: "Sprache und Wissenschaft", in: Harald Weinrich, Hartwig Kalverkämper (eds.): *Sprache, das heißt Sprachen*, Berlin: Frank & Timme (2022), pp. 293–310, p. 307.
[11] Christopher Tollefsen: "What is 'Good Science'?", in: Marta Bertolaso, Fabio Sterpetti (eds.): *A Critical Reflection on Automated Science: Will Science Remain Human?,* Cham: Springer Nature (2020) pp. 279–292, pp. 279–280.
[12] Haoyi Zheng, Huichun Zhan: "ChatGPT in Scientific Writing: A Cautionary Tale", in: *The American Journal of Medicine*: 136/8 (2023), pp. 725–726.