



Vol.7 No.2 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

## Generative Artificial Intelligence in chemistry teaching: ChatGPT, Gemini, and Copilot's content responses

Bruno Silva Leite<sup>A</sup>

A

*Professor, Department of Education, Federal Rural University of Pernambuco, Pernambuco, Brazil*

### Keywords

AI;  
Artificial Intelligence;  
chatbots;  
ChatGPT;  
chemical concepts;  
chemistry teaching;  
Copilot;  
Gemini;  
Generative AI;  
Generative Artificial Intelligence.

### Abstract

Within the past decade, enormous strides have been made related to the disruptive effect of AI in education, which has grown exponentially. AI presents itself as an emerging technology that facilitates the personalization of learning, in addition to promoting debate about educational practices in educational institutions around the world. This study aims to explore the performance and accuracy of three chatbots (ChatGPT, Gemini, and Copilot) in defining five basic chemistry concepts compared to the definitions described in the IUPAC's Compendium of Chemical Terminology. To this end, this descriptive qualitative research was carried out in three stages. The results show that the chatbots presented coherent responses to the investigated scientific concepts, in addition to being configured as resources that can be used in the teaching of chemistry and assisting in the teaching and learning processes. However, although chatbots have demonstrated potential as digital didactic resources in chemistry teaching, their implementation should be carefully monitored. In the near future, it is expected that AI chatbots will become valuable resources in education, helping students in their learning journeys and using chatbots actively and critically.

### Correspondence

[brunoleite@ufrpe.br](mailto:brunoleite@ufrpe.br)<sup>A</sup>

### Article Info

Received 11 June 2024

Received in revised form 16 July 2024

Accepted 22 July 2024

Available online 23 July 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.2.13>

## Introduction

Digital technologies have been powerful agents of transformation in contemporary society, changing the way we interact, communicate, and learn. One of the most notable changes is the democratization of access to information and knowledge (Leite, 2018). With the expansion of the internet and the development of mobile devices, information is just a click away, enabling people from different social backgrounds to have access to educational and cultural resources that were previously only available to a few.

Furthermore, digital technologies have driven a paradigm shift in the educational model, promoting the personalization of learning and flexibility in teaching (Leite, 2022; Reis et al., 2021; Sadykov & Čtrnáctová, 2019). Online teaching platforms, educational software, digital games, and mobile applications, among others, offer the opportunity to adapt the content and methodology according to the individual needs of each student, allowing for more autonomous and effective learning. This student-centered approach has the potential to break with the limitations of traditional teaching, providing a more inclusive education adapted to the challenges of the 21st century (Leite, 2018).

With the evolution of technologies, new challenges are emerging, and one of these challenges involves the insertion of artificial intelligence (AI) into teaching and learning processes. AI has been progressively integrated into the field of education, offering a variety of benefits that aim to optimize teaching and learning processes (Farazouli et al., 2023; Leite, 2023; Pavlik, 2023). As AI becomes incorporated into education, it is important to look at the essence of human creativity in the face of the potential of algorithmic knowledge. According to Khademi (2023), advances in AI have brought significant progress and contributions to deep learning of neural networks and natural language processing and its applications. AI offers adaptive personalization, enabling students to receive immediate feedback and engage in tailored learning experiences. One of the most prominent forms of AI in education is the use of educational chatbots, automated conversational systems that interact with students to provide personalized support, answer questions, and even conduct formative assessments (Rudolph et al., 2023; Silva Neto & Leite, 2024). These chatbots can be programmed to adapt to each student's level of knowledge and individual needs, offering a more individualized and effective approach to teaching subjects such as chemistry.

Additionally, chatbots, such as ChatGPT, Gemini, and Copilot, have been used as support tools. These virtual assistants can answer questions, explain chemical concepts, and even assist in communicating and disseminating research in the area (Leite, 2023). However, it is crucial that the use of these technologies is guided by reflection to ensure that students develop critical thinking and that teachers consider AI as a complementary resource, in addition to other pedagogical strategies (Leite, 2023; Rudolph et al., 2023). Teachers can become the encouragers who will support students so that AI continues to advance in different ways than one can imagine. Teachers should keep in mind that technology will never replace people or human creativity – it is important

to recognize technological tools as a way to improve communication, connection, and creativity in this world.

On the other hand, chemistry is the science of the composition, structure, properties, and reactions of substances, their laws, and principles (Taber, 2013). Chemistry is a science that operates in the basic and applied areas (Doren & Duffy, 2016). The fundamental concepts (e.g., molecules, atoms, etc.), their structures, and properties studied in chemistry often have a wide range of applications in real life. The different research findings in chemistry are closely related to many aspects of daily life (Qian et al., 2023). Therefore, considering the importance of understanding concepts for learning chemistry, understanding the fundamental concepts and exploring their structure in the teaching of chemistry is necessary. As basic concepts are the most important part of the set of concepts, chemistry teachers can use these fundamental concepts as a general reference and implement them in teaching and learning processes.

In this context, this research will analyze the use of chatbots (ChatGPT, Gemini, and Copilot) in the definition of five chemical concepts. This investigation will make it possible to identify the potentialities and limitations of Generative Artificial Intelligence (GenAI) by evaluating its performance in the ability to answer questions involving chemistry.

## Brief literature review and theoretical framework

Artificial intelligence is rapidly transforming many aspects of society, including education. In this new environment, it may seem that traditional digital skills are less important, as AI can perform many tasks more efficiently and accurately than humans. However, AI cannot completely replace the importance of digital skills, as they continue to be relevant in many aspects of education. Firstly, it is important to highlight that AI is still in its early stages of development, and, in most cases, human intervention is required to configure, program, and monitor AI systems. Therefore, people with digital skills will continue to be essential for the effective use of AI in education.

In recent years, AI has demonstrated the ability to generate fluent language, producing sentences that are increasingly difficult to distinguish from text written by people (Editorials, 2023). It is worth noting that AI only “reproduces the processes of our society, in activities for which it was trained, carrying them out faster and, often, with greater precision than a human being” (Leite, 2023, p. 916). Additionally, AI can be a valuable resource in helping students develop digital skills by offering personalized and adaptive learning experiences, providing automated feedback, and helping teachers identify areas of improvement for students.

The term *artificial intelligence* did not emerge recently. This name was created by computer scientists, mathematicians, and statisticians in the 1950s in an attempt to convince the Pentagon (American Department of Defense) to create intelligent machines (Kapur, 2023; Farazouli et al., 2023). It is important to highlight that (biological) intelligence is a property of organisms, and it depends on analogue and not digital processing. Between 0 and 1 of the computer

(of technology), the human brain fills all this space (something the computer cannot do). Thus, care must be taken when equating human and artificial intelligence, as human intelligence cannot be fully captured in accurate and machine-compatible descriptions (Rudolph et al., 2024b). But when artificial intelligence “merges” with biological intelligence, many benefits can result.

In this regard, the concept of human intelligence could be linked to Gardner’s concept of multiple intelligences. Gardner (1993) conceptualized multiple intelligences into eight distinct types: visual-spatial, linguistic-verbal, logical-mathematical, bodily-kinesthetic, musical, interpersonal, intrapersonal, and naturalistic. Rudolph et al. (2024b) highlight if Gardner’s multiple intelligences are agreed upon and intelligence is not reduced to its logical-mathematical aspect, then artificial general intelligence would have to encapsulate all eight intelligences.

Biological intelligence can be defined as the ability to solve complex problems or make decisions with results that benefit the individual. For Pavlik (2023, p. 3), AI refers to “the simulation of human intelligence in machines that are programmed to think and act like humans. These machines are designed to learn from their environment and experiences and are able to adapt their behavior based on this learning.” AI concerns the performance of any of the actions, such as planning, reasoning, problem-solving, perception, knowledge presentation, creativity, etc. (Mecaj, 2022). AI is also defined as the application of statistical methods (multivariate statistics, artificial neural networks) to mine large databases and extract correlations and projections of future estimates (as if the future were to reproduce the past).

Some researchers, such as Chomsky (Chomsky et al., 2023), argue that it would be more prudent to call it artificial assistance since AI is a program that takes a collection of things that already exist. Intelligence is being able to create something new, while AI programs cannot create new things; they only make new combinations. Thus, AI would be configured as an intelligent assistant, just like washing machines, coffee makers, etc., that help you perform some tasks faster and save time that you can use for other activities. According to Popenici (2023, p. 383), “it is deceiving to say, dangerous to believe, that artificial intelligence is... intelligent. There is no creativity, no critical thinking, no depth or wisdom in what generative AI gives users after a prompt: it is just plausible text with good syntax and grammar, and this is all that it is”. It is necessary to understand what AI is and what it is not, what its true contributions are, and what is just marketing (Popenici, 2023). Furthermore, Alan Turing (1950) researched AI, which led UNESCO to recommend AI Ethics in applying machine learning (robots) to prevent abuse, fraud, and security risks (Van Wyk, 2024).

In this context, Gravel et al. (2023) warn us to be careful about the risks of trusting information generated by AI, which is sometimes very convincing quotes but often fabricated. This information appears to be constructed by merging existing references. Although GenAI holds promise for educational improvement, it is necessary to be concerned with academic integrity, with the development of critical AI literacy among

educators and students, emphasizing the need to promote an environment of responsible innovation and informed use of AI (Rudolph et al., 2024a). GenAI in education presents a “yin and yang perspective”, with threats (Technological dependence and the concentration of power; Academic integrity challenges; Dilution or automation of university teaching roles; Quality, accuracy, and ethical concerns; Erosion of graduate attributes; Graduate Employment; Privacy and surveillance; Sustainability; Regulatory and policy challenges) and opportunities (Teaching assistant; Research assistant; Feedback and Assessments; Adaptive curriculum development; Automated Grading; Creating collaborative circles; Student intervention; Administrative support; Including education; Wellbeing support) (Rudolph et al., 2024a). It is important to be aware of how AI impacts education, analyze, investigate, and reflect, in a calm but courageous way, on the role of each of the actors involved (managers, teachers, students, etc.). Nevertheless, amid the growth in the use of GenAI in education, policies for its use, an understanding of its ethical use, and comprehensive regulatory oversight are needed to ensure social well-being (Popenici, 2023; Rudolph et al., 2024a).

The integration of AI in education has proven to be a highly relevant resource for improving teaching and learning processes, thanks to its ability to personalize and provide automated feedback and more objective and accurate assessments (Castaneda, 2023). However, talking about education in times of AI implies “reflecting on the role of education in preparing people for a world in constant change, in which this technology will be present in all aspects of life: work, studies, leisure, personal relationships, etc.” (García-Peñalvo et al., 2024, p. 11). Therefore, it is important to know the benefits and risks of its use, requiring new knowledge, skills, competencies, and values for life and work in this era of AI. According to Waring (2024), universities should be debating the best way for students to use AI technologies, emphasizing the need for educational institutions to reevaluate their curricula.

In Zhai’s view (2022), it is possible to move forward with AI in education as long as we are prepared for the challenges and opportunities that this resource offers. According to Holmes and collaborators (2022) and Grassini (2023), it is necessary to face the complex challenges associated with AI in education, for example, in promoting digital education and ethics. The UNESCO report suggests that human responsibility and control must be maintained in all aspects of the use and application of AI in educational contexts or in everyday environments (Miao & Holmes, 2021). There are several ways to use AI to make learning personalized, including content adaptation, specific recommendations for each student, tailored guidance, and early identification of learning needs.

Artificial intelligence has played a significant role in transforming education, and its application in chemistry teaching could promote notable changes. AI makes it possible to adapt strategies that allow the construction of knowledge, meeting individual and collective needs. Intelligent platforms assess the pace and style of learning, adjusting content and activities according to each person’s progress and difficulties. This ensures that all students

receive the support they need to reach their full potential.

The main educational systems that use AI technologies are intelligent tutoring systems (ITS), learning management systems (LMS), intelligent educational robotics, massive open online courses (MOOCs), learning analytics (learning analysis), and chatbots (Göçmez & Okur, 2023). However, each of these applications uses AI differently, although there is great potential in its use to support tasks in the educational context, both from the perspective of the student and the teacher. Thus, tools such as chatbots, adaptive platforms, and virtual tutoring systems are shaping a more dynamic and effective educational future.

Generative AI (such as chatbots) can generate content (text, images, videos, audio, etc.) from a request, usually expressed in text written in natural language, although there are currently applications that accept multimodal inputs. GenAI is extremely powerful and is improving at a rapid rate, but it is based on large probabilistic-based model languages, which means they have no reasoning or understanding capabilities and are therefore susceptible to errors that need to be contrasted (García-Peñalvo et al., 2024; Leite, 2023). By definition, a chatbot is a computer program designed to simulate conversations with human users, especially over the Internet (King, 2023).

AI chatbots date back to the 1960s, with MIT's first ELIZA chatbot that could simulate a conversation (Farazouli et al., 2023). The ELIZA chatbot was programmed to behave like a Rogerian psychotherapist. Currently, there are different chatbots available, such as ChatGPT (<https://chat.openai.com>), Bard/Gemini (<https://gemini.google.com/app>), Bing Chat/Copilot (<https://copilot.microsoft.com>), LuzIA (<https://www.luzia.com>), Llama 2 (<https://llama.meta.com>) and Claude 2 (<https://claude.ai>) (Figure 1).



Figure 1: Chatbots.

ChatGPT is based on a large language model (LLM) called Generative Pre-trained Transformer (GPT). GPT is an artificial learning model that uses unsupervised and supervised learning techniques to understand and generate human-like language. ChatGPT, launched on November 30, 2022, by OpenAI, is capable of answering everything from simple questions to writing more complex texts (Ismail et al., 2023; Liu et al., 2021) in a language almost indistinguishable from natural human language (Leite, 2023; Ogunleye et al., 2024; Rudolph et al., 2023). ChatGPT is a conversational LLM chatbot that acts on prompting (Van Wyk, 2024). However, ChatGPT has been called a "stochastic parrot" (Arkoudas, 2023; Bender et al., 2021) since it only "approximated

the content of its replies according to its training data without a deeper understanding of such content, that this approximation also applied to numbers and references which often ended up being wrong or fictitious" (Schäfer, 2023, p. 5), with replies were often of limited quality (Gravel et al., 2023).

Bard is a large language model from Google AI, trained on a huge dataset of text and code. Bard was launched on March 21, 2023, as Google attempted to compete with ChatGPT (a competing bot produced by OpenAI). With Bard's advancements, Google renamed it Gemini on February 8, 2024 (Google, 2024). Gemini can generate text, translate languages, write different types of creative content, and answer questions in an informative way.

Copilot started under the name Bing Chat, a chatbot developed in collaboration with OpenAI (the same company that created ChatGPT). Currently, this chatbot is designed to support various tasks related to writing and productivity across various Microsoft applications (Motlagh et al., 2023). Copilot (launching November 15, 2023) uses a large language model to generate text, translate languages, create creative content, and answer questions in an informative way.

The LuzIA chatbot transforms audio into text with the help of ChatGPT's artificial intelligence. LuzIA works like ChatGPT, but the conversation takes place directly with a common cell phone contact (via WhatsApp or Telegram chat). LuzIA uses three artificial intelligence models and, based on them, is capable of unlimited interaction through text messages (GPT 3.5, from OpenAI), transforming audio inputs of up to 10 minutes into text and creating up to five images per day from text descriptions (LuzIA, 2023). Flores and Peña (2024) highlight the possibilities of using LuzIA by teachers in designing thematic content to integrate them into their planning.

Llama 2 is an open-source language model developed in partnership between Meta and Microsoft (Meta, 2024). Llama 2, an updated version of Llama 1, trained on a new mix of publicly available data (Touvron et al., 2023). Unlike previous chatbots that are available in the cloud, Llama 2 can be downloaded for free and used on a computer with a Windows system.

Finally, Claude 2 is a chatbot developed by the company Anthropic and launched in March 2023. Claude 2 is an AI system developed to assist users in carrying out tasks in a conversational way, that is, through a prompt text, the user can explain what they need from the tool, and it will return with a solution. According to Lozić and Štular (2023), a unique feature of Claude 2 is the size of its context window of 100,000 tokens, or about 75,000 words, which allows users to paste or upload large documents or entire books and ask questions based on its content. The Claude 3 (released March 19, 2024) is available in the United States, the UK, and some countries in Latin America, Africa, and Asia via a website ([claude.ai](https://claude.ai)) and as an application programming interface (API). The Claude 3 family of models initially offered a 200,000 tokens context window upon launch. However, all three models (Haiku, Sonnet, and Opus) are capable of accepting inputs greater than 1 million tokens (Anthropic,



2024).

AI-based chatbots take on an important role in education, as thinking about prohibiting or restricting their use does not seem to be the best solution. Chatbots can simulate conversations with students, offering individualized support, answering questions, clarifying concepts, and providing personalized feedback (Leite, 2023). The use of GenAI can either improve or harm learning processes, depending on whether it is used to augment or 'simply' delegate the execution of tasks. Educational environments (like universities and schools) can leverage the capabilities of GenAI chatbots and take advantage of creating innovative teaching and learning opportunities (Van Wyk, 2024). Ogunleye et al. (2024) aimed to assess the performance of GenAI tools in STEM-related disciplines to understand their potential impact on students' learning and development. In teaching chemistry, for example, chatbots can help understand complex formulas, solve problems, and even simulate experiments, making learning more interactive and accessible. Furthermore, chatbots can help explain abstract concepts, such as molecular structure, through interactive simulations and 3D visualizations. Furthermore, AI can be used to create games and/or augmented and virtual reality experiences, enabling chemistry teaching to be more effective.

In the literature, there are some examples that make use of chatbots in chemistry teaching (Clark, 2023; Dempere et al., 2023; Emenike & Emenike, 2023; Hallal et al., 2023; Leite, 2023). For example, ChatGPT has been used to define chemical concepts, providing coherent answers about the concepts. However, depending on the topic, AI responses may lead to difficulties in understanding for students (Leite, 2023). Dempere et al. (2023) used ChatGPT to explore the effects of AI chatbots on higher education institutions. ChatGPT was also used to answer questions (open and closed) from final exams administered in two chemistry courses (Clark, 2023). Emenike and Emenike (2024) posed related questions to ChatGPT using examples from general and organic chemistry. The authors pointed to some potential benefits and risks of ChatGPT and similar GenAI systems, such as answers to items on assignments and assessments, lab reports, tutoring systems, and text generation for papers assigned in class, for projects, or for research-related assignments. Hallal et al. (2023) used ChatGPT and Bard in understanding text-based structural notations such as condensed structures, InChi, and SMILES and in answering questions related to organic chemistry. In Dao's (2023) analysis, Bard surpassed ChatGPT and Bing Chat in chemistry discussions. In view of the above, this research will analyze the use of these three chatbots (ChatGPT, Gemini and Copilot) in the definition of five chemical concepts.

## Methodology

The methodology is the item that guides the investigation process in decision-making, hypotheses, and research techniques. According to Creswell (2017), generally, not just one method or technique is applied when carrying out research. It is possible to combine two or more methods and techniques.

In this sense, to respond to the objective proposed in this research, a qualitative approach of a descriptive nature is used, as it allows the description, analysis, and evaluation of data in an articulated and in-depth way, with the researcher being its main instrument (Creswell, 2017), in addition to having a case study as a methodological design. For Johnson and Christensen (2008), this type of research aims to tell the story of a limited system (one case). The central idea of this type of study is that it is possible to analyze the system in its entirety from its parts.

Based on a case study, in which a specific event is researched in detail, the responses obtained from chatbots will be descriptively analyzed. The research was developed in three stages. In the first stage, free accounts were created in three chatbots (ChatGPT, Gemini, and Copilot). Considering that it is a digital platform with closed characteristics, it was necessary to register in the environment to access its functionalities. The selection of these chatbots was based on the findings of Silva Neto and Leite (2024), which revealed the three chatbots most chosen by students when writing a dissertation or essay.

The second stage of the research consisted of preparing and asking questions in the three chatbots. In this sense, the questions were prepared and formatted in the GenAI chat. As already described, the chatbot is a generative artificial intelligence that can produce texts from questions. Therefore, AI was used to write definitions of five chemical concepts and then analyzed their academic validity. For this, as a research method, it was determined that the definitions would refer to the chemical concepts: Atom, Electron, Mole, Molecule, and Chemical Substance. These concepts were chosen to take into account that they are commonly present in discussions involving chemistry (Childs, 2015; McGill et al., 2019; Qian et al., 2023) and are generally addressed in students' first contact with the discipline. The question asked to the chatbots was: Define "[atom; electron; mole; molecule; chemical substance]". Considering that these are large language models, all chatbot responses are generated in situ, based on the abstract relationship between words in the neural network (Kung et al., 2022). Therefore, in order to reduce memory retention bias, a new chat session was started in each chatbot for each question asked.

Finally, in the third stage, the texts produced by ChatGPT, Gemini, and Copilot were analyzed and compared with the definitions present in the Gold Book. The Gold Book is the Compendium of Chemical Terminology organized by IUPAC (International Union of Pure and Applied Chemistry), which provides internationally accepted definitions of chemical terms. The compendium (Figure 2) provides an alphabetical list of chemical concepts with definitions recommended by various IUPAC commissions – responsible for organizing nomenclatures, terminologies, symbols, and units (IUPAC, 2019). Its first version was published in 1987 by Victor Gold, hence the informal name: Gold Book.

The texts created by the chatbots were checked by plagiarism tools and AI text generation detectors such as Plagium (<https://www.plagium.com>), Smodin (<https://smodin.io/ai-content-detector>), CopySpider (<https://copyspider.com.br/main/en>), Writer.com's AI Content Detector (<https://writer.com>).

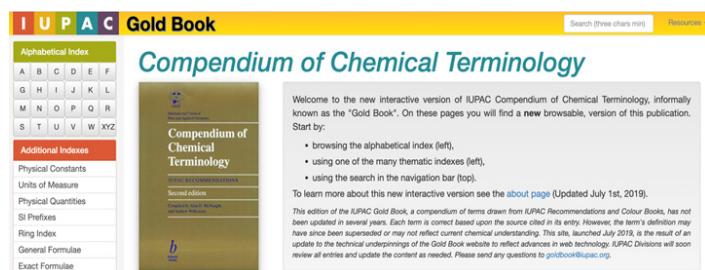


Figure 2: IUPAC Compendium of Chemical Terminology home page.

com/ai-content-detector), and GPTZero (<https://gptzero.me>). All programs used were free versions and available to the public. Plagiarism detection software completely analyzes documents in search of instances of suspected plagiarism in which the results are displayed in a detailed manner in a document, indicating all analysis criteria.

## Results and discussion

In this section, the results achieved while using ChatGPT, Gemini, and Copilot will be presented. The chatbots' answers to questions about the definition of chemical concepts will be discussed. Due to the output nature of chatbots, each response made available was manually reviewed. It is emphasized that the focus of this research is not to discuss each concept presented by chatbots but rather to analyze the answers presented by GenAI and compare them according to the Compendium of Chemical Terminology.

In Leite's view (2023), the word *chemistry* has been understood by current society as something 'good' when it produces substances and materials that benefit people or as something 'bad' when it is associated with things that harm health, the environment and society. In addition to this simplistic and distorted view of chemistry, it is known that the application of chemistry can contribute to the development of society, whether in technical aspects or in social/human aspects. Therefore, observing how chemistry is present in people's daily lives, from daily activities to the development of medicines to cure a disease, has been fundamental, although it is considered difficult for students to understand. Therefore, debating the contents of chemistry in the educational context requires adequate training from the teacher that goes beyond knowing chemistry, but that the teacher knows how and why to teach it.

Currently, teachers have different teaching resources at their disposal (book, board, laboratory, notebook, computer, smartphone, internet, etc.), which requires training to use them in their pedagogical practices. In this context, the rise of GenAI has perhaps provided teachers and students with yet another digital resource to contribute to the teaching and learning processes. Nevertheless, GenAI resources also present situations that are not simple (at the moment) to be resolved, such as the ethical conduct of AI users (Holmes et al., 2022) and plagiarism (Chaka, 2023, 2024). Holmes et al. (2022) point to an ethical challenge in the use of AI in education that requires careful analysis and a considered approach since many AI resources can act in a non-

transparent way, where decisions are based on complex and less clear algorithms. For Chaka (2023), AI content detectors are not able to distinguish, in clear-cut terms, between AI-generated texts and human-produced texts. These programs make estimates in percentages or in probabilistic terms, such as probable and probability, or in combined percentages and probabilistic terms. Thus, according to Chaka (2023), AI content detectors appear not yet to be fully ready to accurately and convincingly detect AI-generated content from machine-generated texts in different contexts.

For this research, five pieces of software were used that aim to identify text produced by AI and whether there is plagiarism. Regarding plagiarism detection, the software Plagium and CopySpider were used, and the results showed that none of the texts generated were considered plagiarized. This, perhaps, is because plagiarism detectors appear to still be in the process of being refined, without the ability to differentiate between AI-generated content and human-written text in a consistent and convincing way (Chaka, 2023, 2024). For each chemical concept investigated, it will be mentioned how the software identified the content, whether generated by AI or mistakenly by humans.

Concepts are the basis of human knowledge structures (Johnson et al., 1994), and in recent years, conceptual research has received increasing attention in education (Qian et al., 2023). It is conjectured that a lack of understanding of chemical concepts impedes students' ability to solve chemical problems and impairs their understanding of the entire conceptual structure (Qian et al., 2023). Furthermore, many chemical concepts are not part of students' daily lives, making it difficult for them to make connections with these abstract and complex concepts. Thus, it is a challenge for students to assimilate new concepts into their minds to establish connections with their existing conceptual structure. According to Qian et al. (2023), many new concepts in chemistry emerged during its development, which indicates the importance of exploring chemical concepts.

Chemistry contains many basic abstract concepts that are essential to the body of knowledge. Therefore, exploring chemical concepts and understanding their relationships is crucial. However, in chemistry, it is common for different concepts to exist (Sadykov & Čtrnáctová, 2019; Tomasevic & Trivic, 2014; Upahi & Ramnarain, 2019), and in the teaching process, the teacher usually presents definitions of these concepts. It is worth noting that defining is not the same as conceptualizing. From the perspective of the Oxford Dictionary, defining consists of "that describes or shows the essential meaning of something" (Oxford Learners Dictionaries, 2024), and conceptualizing is "to form an idea of something in your mind" (Oxford Learners Dictionaries, 2024). It is understood that defining consists of the description of the properties or characteristics of an object or concept. Conceptualizing corresponds to an intellectual interpretation of an aspect of sensitive reality. The definition involves the use of nominalization, whereby processes, which normally occur as phenomena are transformed into names or nominal groups. Thus, defining means delimiting, saying what it is, and conceptualizing is about delimiting, saying what it is, but locating the object in a web of relationships. Concepts are historical human constructions that, in new

contexts, acquire new meanings. Regarding concepts, Vygotsky (2012) considers that they evolve with the meanings of words and tend to become more precise within a system of generalities. Every concept is a generalization. At the moment of assimilating a new word, term, expression, etc., the concept development process is just beginning.

In view of the above, the term defined with the aim of chatbots is taken as a reference, describing, explaining, delimiting, or saying what each requested concept is. To facilitate the presentation of the results, the responses of the three chatbots for each concept covered will be presented. Thus, the first concept that chatbots were asked to respond to was the definition of an atom.

#### ChatGPT defined atom as

the smallest basic unit of a chemical element. It consists of a central nucleus, composed of protons (with a positive charge) and normally neutrons (without an electrical charge), around which electrons (with a negative charge) orbit. Atoms are composed mainly of empty space, with the central nucleus concentrating almost all of the atom's mass. Atoms are the fundamental constituents of all matter and are characterized by the number of protons in their nucleus, known as the atomic number, which determines the identity of the chemical element (OpenAI, 2024).

Gemini, in addition to presenting a definition for an atom, provided information about the atom, such as atom structure, important characteristics, atomic models, the importance of atoms, and three links to "learn more". Regarding the definition, Gemini stated that

Atom is the fundamental unit of matter and the smallest fraction capable of identifying a chemical element. The term "atom" comes from Greek and means "indivisible", since, until the 19th century, it was believed that this was the smallest part of matter that existed (Gemini, 2024).

Copilot defined the atom as "the fundamental particle that makes up matter" (Copilot, 2024). This chatbot, in its response, added other information, highlighting that the atom has two main structures:

Nucleus: In the nucleus, there are protons (with a positive charge) and neutrons (without an electrical charge). These particles are responsible for most of the atomic mass and are concentrated in a small, extremely dense volume. [...] Electrosphere: The electrosphere is the peripheral region of the atom, where the electrons (with a negative charge) are located. Although the term "atom" comes from the Greek and means "indivisible", we know that it is, in fact, divided into smaller particles, known as subatomic particles (Copilot, 2024).

At the end of the answer, Copilot (2024) explained that the "concept of the atom originated in Ancient Greece, thanks to the philosophers Leucippus and Democritus, who

initially considered it as an idea". Copilot was the only one of the three chatbots that presented sources about the text produced. Gemini presented links for consultation, but Copilot informed the source that consulted the text generated by the AI.

When defining an atom, the chatbots produced a text that presents keywords that are commonly observed in the literature. The idea that the atom is an infinitely small particle (Atkins & Jones, 2007; Brown et al., 2017), which has a positive nucleus that contains protons (positively charged particles) and electrons (negatively charged particles), and that each element is composed of a single type of atom (Brown et al., 2017). In the Compendium of Chemical Terminology, an atom is "the smallest particle still characterizing a chemical element" (Book, 2014, p. 121). In their responses, the three chatbots, in addition to presenting a coherent definition of an atom, also related it to the definition of a chemical element.

Regarding the answers generated by ChatGPT, Gemini, and Copilot about the definition of an atom, by way of illustration, GPTZero identified that 3% of the ChatGPT text was produced by AI, while for Gemini, GPTZero indicated the probability of 72 % of the text was produced by AI and that 82% of the Copilot text was generated by AI (Figure 3).

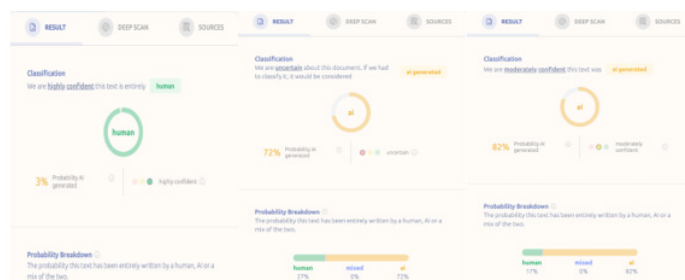


Figure 3: GPTZero analysis of the atom definition of ChatGPT, Gemini, and Copilot, respectively.

The data reveals that both Gemini and Copilot had text identified as produced by GenAI by GPTZero. Although the program presented doubts regarding the Gemini text: "We are uncertain about this document. If we had to classify it, it would be considered AI generated" (GPTZero, 2024). Writer.com's AI Content Detector software indicated that the ChatGPT text presented that 81% of the content about the atom definition was human-generated, for Copilot 88% of the text was human-generated, and for Gemini, the percentage was 72% for human-generated content. That is, for the AI content detector the three texts were produced by humans, which indicates that the software was unable to correctly detect that the text was generated by AI. In Plagium, the software reported that the text of the three chatbots was probably written by a human: "This text is very likely written by a human" (Plagium, 2024). Smodin reported that Gemini's text only was 100% AI-produced.

The second question asked ChatGPT, Gemini, and Copilot to define "electron". The answers are described in Table 1.



Table 1: Chatbots' responses to the definition of electron.

ChatGPT	Gemini	Copilot
An electron is a subatomic particle that orbits around the nucleus of an atom.	An electron is a subatomic particle with a negative electrical charge. It is one of the fundamental components of the atom, along with the proton and neutron.	The electron is a subatomic particle that makes up the atom.

The chatbots presented more information about the electron in their responses, highlighting characteristics and information about its discovery. For the chemistry terminology compendium, the electron is an "Elementary particle unaffected by the strong force with spin quantum number  $1/2$ , a negative elementary charge and a rest mass of  $0.000\ 548\ 579\ 903(13)\ u$ " (Book, 2014, p. 477). It was observed that the answer presented by the Compendium is more complex than those described by the chatbots, although Gemini made mention of the electron's spin and its mass - "Mass:  $9.109 \times 10^{-31}\ kg$  (approximately  $1/1836$  of the mass of the proton)" (Gemini, 2024).

In the analysis of the texts generated by AI, the responses of the ChatGPT were considered as texts produced by humans, according to Plagium, Smodin, and AI Content Detector software. Only GPTZero identified that the text on the definition of an electron was produced by AI. Gemini had its text identified as generated by AI by all software. In the analysis of the text created by Copilot, only Plagium identified it as generated by AI.

Regarding the third question (Define Mole), ChatGPT responded that the Mole "is a unit of measurement used in chemistry to express the amount of substance present in a sample. The mole is defined as the amount of substance that contains approximately  $6.022 \times 10^{23}$  elementary entities, such as atoms, molecules or other particles." (OpenAI, 2024). Gemini, in addition to defining moles as a "fundamental unit of the International System of Units (SI) for the quantity of matter. It is a quantity used to count elementary entities of a substance, such as atoms, molecules, ions or electrons.", presented an analogy, which said: "Imagine that you need to count grains of sand on the beach. It would be impossible to count them one by one. The mole functions as a "counting unit" for the microscopic world, similar to the dozen (12 units) for eggs" (Gemini, 2024).

It is important to highlight that Gemini, in addition to presenting topics with information about the mole, highlighted that there is a current definition for mole: "It is important to highlight that, as of May 2019, the definition of the mole is no longer based on the mass of carbon-12. Currently, it is defined directly by Avogadro's number" (Gemini, 2024), summarizing its definition to: "The mole is an essential tool for quantifying the microscopic world, enabling precise calculations and relationships between different chemical quantities" (Gemini, 2024). Copilot presented the following definition of Mol:

One mole corresponds to approximately  $6.022 \times 10^{23}$  particles. This unit is one of the fundamental quantities of the International System of Units (SI). The term "mol" derives from the Latin word "moles", which means "a mass". It was introduced by the German chemist Wilhelm Ostwald (Copilot, 2024).

In the Compendium of Chemical Terminology, the definition for a mole is "the SI unit of quantity of substance. One mole contains exactly  $6.022\ 140\ 76 \times 10^{23}$  elementary entities. This number is the fixed numerical value of Avogadro's constant,  $N_A$  when expressed in the unit  $\text{mol}^{-1}$  and is called Avogadro's number" (Book, 2014, p. 948). This answer is very close to that presented by ChatGPT and Gemini. Furthermore, Gemini and Copilot highlighted similar information in the definition of Mol, while ChatGPT chose to present a description of what Mol is without presenting other characteristics. However, it was observed that Gemini's response was the most generic compared to the responses from ChatGPT and Copilot.

Regarding the answer created by Copilot about the definition of mole, all software mistakenly identified it as produced by humans. Smodin and GPTZero were correct in determining that the text produced by ChatGPT was generated by AI, and only Smodin reported that the Gemini text was generated by AI.

The fourth concept asked for chatbots to define a molecule. For ChatGPT, a molecule "is a chemical entity composed of two or more atoms joined by chemical bonds." (OpenAI, 2024). Gemini and Copilot presented similar answers to the definition of a molecule: "It is a set of two or more atoms covalently linked, forming the smallest unit of a substance that maintains its chemical properties" (Gemini, 2024) and "is a set of atoms, the same or different, united by covalent bonds" (Copilot, 2024). In the compendium, a molecule is "an electrically neutral entity that consists of more than one atom ( $n > 1$ )" (Book, 2014, p. 958). Regarding the molecule being electrically neutral, of the three chatbots, only ChatGPT did not explain this characteristic. According to Qian et al. (2023), the lack of information can lead to students having limitations in understanding the concept, since the lack of understanding of chemical concepts impedes students' ability to solve problems.

Concerning the answers generated by ChatGPT, Gemini, and Copilot about the definition of a molecule, Smodin was the only one that correctly identified them as texts produced by AI. For AI Content Detector (Writer.com) and GPTZero, the texts produced by the three chatbots were generated by humans. It is worth noting that Plagium reported in its response that it was unable to distinguish whether the ChatGPT and Gemini text were produced by AI or by humans. The software presents the following response: "We are perplexed, this text is probably written by a human, but maybe generated by a machine" (Plagium, 2024). In the text created by Copilot, Plagium identified it as human-generated.

Finally, when defining a chemical substance, the chatbots presented more simplified answers than the Gold Book. ChatGPT highlights that a chemical substance is "a form of matter that has a defined chemical composition and distinct properties. It can be composed of molecules, atoms, or ions" (OpenAI, 2024). Gemini highlights that a "chemical substance, also called a pure substance, is a material with a defined and constant chemical composition, that is, it has specific and well-defined physical and chemical properties" (Gemini, 2024). Copilot defined a chemical substance as "any



kind of matter formed by a constant composition of chemical elements, having defined physical and chemical properties” (Copilot, 2024). According to the Compendium of Chemical Terminology, a chemical substance is a “matter of constant composition best characterized by the entities (molecules, formula units, atoms) that compose it” (Book, 2014, p. 265). In addition, it highlights that physical properties such as density, refractive index, electrical conductivity, and melting point characterize the chemical substance (IUPAC, 2019). It is observed that the definitions presented by the chatbots are simpler and aimed at a beginner audience in chemistry, but they all converged by highlighting the constant composition of the chemical substance.

For the definition of a chemical substance, Smodin was only able to detect the text produced by ChatGPT as created by AI, while Plagium identified the three chatbots that the texts were created by AI. Concerning the responses generated by ChatGPT, Gemini, and Copilot regarding the definition of chemical substance, Writer.com’s AI Content Detector and GPTZero incorrectly identified that all texts were produced by humans. Furthermore, GPTZero in response to Copilot, indicated that they were highly confident that the text is entirely human: “We are highly confident this text is entirely human” (GPTZero, 2024).

Table 2 presents a summary of the responses from AI-generated text detectors on texts produced by chatbots, indicating the probability of having been produced by a human (H) or by AI (AI).

Table 2: Responses by AI content detectors each chatbot.

Tools	Chatbot	Atom	Electron	Mole	Molecule	Chemical substance	Generated by AI
Plagium	ChatGPT	H	H	H	H/AI	AI	33.3%
	Gemini	H	AI	H	H/AI	AI	
	Copilot	H	AI	H	H	AI	
Smodin	ChatGPT	0% AI	0% AI	100% AI	0% AI	100% AI	33.3%
	Gemini	100% AI	100% AI	100% AI	0% AI	0% AI	
	Copilot	0% AI	0% AI	0% AI	0% AI	0% AI	
AI Content Detector	ChatGPT	81% H	74% H	75% H	71% H	77% H	6.7%
	Gemini	72% H	51% AI	78% H	76% H	76% H	
	Copilot	88% H	81% H	90% H	84% H	87% H	
GPTZero	ChatGPT	97% H	83% AI	91% AI	95% H	96% H	33.3%
	Gemini	72% AI	71% AI	98% H	94% H	97% H	
	Copilot	82% AI	97% H	56% H	95% H	99% H	

Regarding the programs used, it is important to highlight that Plagium does not provide a percentage in its responses, only indicating whether the text was generated by AI (This text is likely generated by a machine) or generated by a human (This text is very likely written by a human) or both (We are perplexed, this text is probably written by a human but may be generated by a machine). Copyspider only indicates whether the text is plagiarized or not. Smodin, AI content detector and GPTZero in their responses present the percentage indicating whether the text was created by AI or by a human.

The data in Table 2 reveals that, in general, the software made mistakes in detecting generated texts, identifying them as produced by humans (Figure 4). Of the five concepts

analyzed by each software (three for each chatbot), the AI Content Detector in 93.3% of the responses indicated that it was human-generated text. Smodin and GPTZero identified 66.6% of texts as produced by humans. In two responses, Plagium was unable to answer whether it was a text generated by AI or by a human, so of the fifteen analyses, eight (53.3%) considered the text as generated by a human.

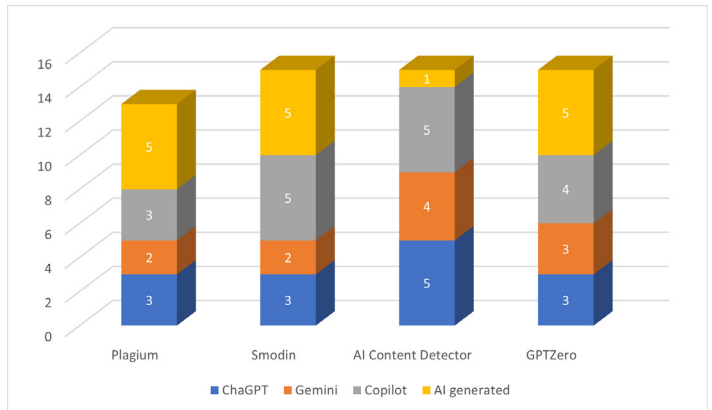


Figure 4: Quantity of texts identified as human-created by the programs (Plagium, Smodin, AI Content Detector and GPTZero) and correctly identified as AI generated.

As seen in Figure 4, Plagium incorrectly identified the text as human-generated on three out of five occasions (3/5) for ChatGPT’s responses. For Gemini’s answers, the software made mistakes on two occasions (2/5), and for Copilot, it erred three out of five times (3/5). It is possible to observe that Plagium was correct in only five instances out of a total of 15 analyses (5/15, 33.3%). Additionally, Plagium could not determine whether the ChatGPT and Gemini responses for the concept of the molecule were AI-generated (Table 2).

Among the AI content detection programs, AI Content Detector was the least successful in identifying whether the text was AI-generated, with only one correct identification (1/15, 6.7%). The other software programs (Plagium, Smodin and GPTZero) showed a low accuracy rate, about one-third (33.3%), but they were the ones that most successfully identified AI-generated content. For Chaka (2023), most AI content detectors are new, and not much research has yet been carried out to assess their efficacy, accuracy, and reliability in terms of distinguishing between content generated by chatbots and content written by humans. This is probably one of the reasons that the software, for the most part, indicated that the texts generated by GenAI were generated by humans.

With regard to the concepts presented by chatbots, Table 3 presents a summary describing the responses considered satisfactory, partially satisfactory, and unsatisfactory. Satisfactory responses (SR) were considered to be those in which the chatbots presented the same or similar definition of the concept to those observed in the Compendium of Chemical Terminology. The responses considered partially satisfactory (PS) are those in which the chatbots responded correctly to the definition of the concept but failed to present any additional information (however, its absence does not affect understanding of the concept) or provided a basic answer (simple, summarized). Those in which the

chatbot incorrectly presented the definition of the concept investigated were considered unsatisfactory responses (UR).

Table 3: Summary of chatbots' responses to define the requested concepts.

	Concept	ChatGPT	Gemini	Copilot
Answers overview	atom	SR	SR	SR
	electron	PS (basic answer)	SR	PS (basic answer)
	mole	SR	SR	SR
	molecule	PS (failed to present a characteristic of the molecule)	SR	SR
	chemical substance	PS (Provided a pragmatic/concise response)	PS (Provided a pragmatic/concise response)	PS (Provided a pragmatic/concise response)

Overall, the responses from Gemini and Copilot were mostly considered satisfactory. Among the three chatbots, when defining the five concepts, Gemini was the one that presented the most satisfactory responses (four definitions out of a total of five), followed by Copilot with three satisfactory responses. ChatGPT, in this investigation, was the chatbot that had the most partially satisfactory responses. There were three partially satisfactory responses and only two satisfactory responses. Nevertheless, none of the chatbots presented unsatisfactory responses in defining the five concepts explored.

Our results differ from the findings of Rudolph et al. (2023) who found that the performance of ChatGPT (both the free version and the commercial version) was much better than Google Bard (Gemini). However, the experiment by Rudolph et al. (2023) aimed to answer fifteen questions from different fields, and none involved defining chemical concepts. Khademi (2023, p. 79) "demonstrated with empirical data that ChatGPT-3.5 and Bard (Gemini) failed to achieve a performance comparable to human experts in rating the complexity of writing prompts". Furthermore, his study highlights that, in some applications, such as education and assessment, ChatGPT and Bard (Gemini) require further improvement and training to perform equivalent to humans. In this context, Table 4 shows a summary of the chatbots' performance, highlighting the strengths and weaknesses in the responses to the chemical concepts observed in this research.

According to Qian et al. (2023), individuals with abundant chemical knowledge or a high level of cognitive ability are able to summarize the common points of some concepts well and are not confused by the incorrect classification of a concept. In this context, with regard to the responses presented by the chatbots for the five concepts analyzed, it is understood that GenAI did not compromise the students' initial understanding in relation to the concepts described, nor did they present confusion during the definition of these concepts. The chatbots' responses proved to be organized and coherent at a basic level. According to Leite (2023), the quality of GenAI responses depends on the concept being evaluated. Holme et al. (2015) identified five categories that define conceptual understanding of general chemistry contents in higher-education (Transfer; Depth; Predict/explain; Problem solving; Translate). These categories provide a framework to understand and evaluate conceptual

Table 4: Strengths and weaknesses in chatbots' responses to the chemical concepts investigated.

	ChatGPT	Gemini	Copilot
Strengths	<p>The answer presented is not limited to just the requested definition;</p> <p>Presents additional information on the requested concept.</p>	<p>The answer presented is not limited to just the requested definition;</p> <p>Presents additional information on the requested concept;</p> <p>Provides links related to the topic investigated (learn more).</p>	<p>The answer presented is not limited to just the requested definition;</p> <p>Presents additional information on the requested concept;</p> <p>Provides links related to the topic investigated (learn more);</p> <p>Presents the sources consulted for the texts produced;</p> <p>Presents images and/or formulas (when applied to the concept).</p>
Weaknesses	<p>Does not provide links to consult about the content investigated;</p> <p>Does not present the fonts used for the text produced;</p> <p>Presents the answers in continuous text;</p> <p>In some cases, it presents the answer without additional information.</p>	<p>It presents answers on topics that may not contribute to the understanding of the requested definition, causing an excess of information (sometimes it is not necessary at that moment).</p>	<p>Excessive extra information can lead the reader to have difficulty understanding the answer.</p>

understanding in the context of chemistry education (Daher et al., 2023).

For this research, it was not assessed whether chatbots are capable of covering the five categories exposed by Holme et al. (2015) since chatbots would need to demonstrate the ability to: apply core chemistry ideas to novel chemical situations; reason about core chemistry ideas using skills that extend beyond rote memorization or algorithmic problem solving; extending situational knowledge to predict and/or explain the behavior of chemical systems; demonstrating critical thinking and reasoning while solving problems, including those involving laboratory measurements; translate across different scales and representations.

On the other hand, it is necessary to warn that there is no point in students taking ready-made responses from chatbots and memorizing them, without giving them a true meaning. Chemical concepts must be taught beyond the memorization and/or reproduction of information. It is important for students to understand chemistry and how this science can transform society, contributing to the development of critical thinking.

## Conclusions and recommendations

This study presents significant contributions to advance GenAI research in the context of chemistry teaching. The findings of this study highlight the convergence in responses from AI chatbots compared to the IUPAC Compendium of Chemical Terminology for five chemical concepts (atom, electron, mole, molecule, and chemical substance). As noted, ChatGPT, Gemini, and Copilot provided satisfactory results for defining the concepts. However, these chatbots

are under development and seek to achieve high accuracy in various textual formats, including the definition of concepts. It is thus necessary to analyze each response provided by GenAI.

The knowledge gained from using chatbots points to encouraging students to try GenAI, as they can provide an efficient way to familiarize themselves with a new topic, find some additional resources, discover examples that were not covered in the classroom, and/or assist in understanding some content. When teaching chemistry, in addition to specific knowledge (in chemistry), the student will need to master philosophical, historical, psychological, and sociological knowledge.

However, it is crucial to recognize the substantial promise that chatbots hold for the field of education. Aligned with the global shift towards personalized, self-directed learning and sustainable educational systems, these resources have the potential to contribute to the knowledge construction process. Furthermore, chatbots should be understood as another resource for teachers to use in the teaching and learning processes. They are capable of providing answers to a multitude of questions, but their use must be accompanied by reflection. Human experience remains fundamental, as nothing replaces your ability to reason, contrast ideas, present critical thinking, and know which sources are being used. However, chatbots have created a feeling in society, including teachers and students, that the answers are accurate and unquestionable. Users of these technologies must validate the results generated by generative AI before applying them effectively, avoiding the use of misleading or incomplete information. A complementary relationship between GenAI and users (teachers, students, researchers, etc.) should be considered, where the use of these chatbots can expedite the analytical process without diminishing the essential role of individual experience and critical engagement (Perkins & Roe, 2024). Furthermore, as AI is further explored and improved, it is also possible that new educational methods can be developed and the full potential of chatbots in education can be used.

More research is needed to build on recent gains in academic awareness of GenAI tools for teaching and learning processes, exploring whether chatbots can contribute to building stakeholder knowledge. However, the conclusions of this exploratory study come from a small sample and cannot be generalized; therefore, a larger study must be carried out, which may produce similar or different results. Furthermore, considering the current pace of AI development, conclusions from research involving GenAI may become obsolete within a period of months, sometimes less, as when the results are published, new versions of the models that were the basis are already available for these studies. For example, when ChatGPT was launched, based on the GPT 3 model, several research studies were carried out, then studies based on GPT 3.5, then GPT 4 (ChatGPT 4) (Dis et al., 2023; Ismail et al., 2023; Lo, 2023; Ogunleye et al., 2024; Parker et al., 2024; Paul et al., 2023). Similarly, when Bard was launched, different studies-investigations were carried out on this model that (quickly) evolved into Gemini (Khademi, 2023; Luzano, 2024; Nguyễn et al., 2023; Ogunleye et al., 2024; Perera & Lankathilake, 2023). This also happened with Bing Chat, which is currently

called Copilot (Głushkova & Malinova, 2024; Motlagh et al., 2023; Rudolph et al., 2023). These research studies sought to validate and analyze the capabilities of these models in comparison to activities carried out by humans. Given the rapid advancements in GenAI capabilities, many of these conclusions have quickly become obsolete and no longer reflect the current reality.

Finally, I believe this research can help teachers and students reflect on the use of AI in the knowledge construction process. Instead of merely giving commands to chatbots, they can learn to engage in meaningful dialogue with language models and explore innovative applications of artificial intelligence in chemistry teaching. There is discouragement or disappointment when it is not possible to have a rich and contextualized interaction with the language models, and the chatbot ends up providing a wrong answer. It is necessary to stimulate these systems as we stimulate a human conversation. If we speak to an expert in a certain area and don't ask the appropriate questions, what will probably happen is that the answers or the richness of the answers will not be sustained. Therefore, it is important to know how to talk.

## References

- Anthropic. (2024, March 4). *Introducing the next generation of Claude. Announcements*. <https://www.anthropic.com/news/claude-3-family>
- Arkoudas, K. (2023). ChatGPT is no stochastic parrot. But it also claims that 1 is greater than 1. *Philosophy & Technology*, 36(3), 54. <https://doi.org/10.1007/s13347-023-00619-6>
- Atkins, P., & Jones, L. (2007). *Chemical principles: The quest for insight*. Macmillan.
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the dangers of stochastic parrots: Can language models be too big?. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610-623). <https://doi.org/10.1145/3442188.3445922>
- Book, G. (2014). Compendium of chemical terminology. *International Union of Pure and Applied Chemistry*.
- Brown, T. L., LeMay, H. E., Bursten, B. E., Murphy, C., Woodward, P. M., & Stoltzfus, M. W. (2017). *Chemistry: The central science*. Pearson Education.
- Castaneda, A. U. (2023). Un viaje hacia la inteligencia artificial en la educación [A journey towards artificial intelligence in education]. *Realidad Y Reflexión [Reality & Reflection]*, 56(22), 121-136. <https://doi.org/10.5377/ryr.v1i56.15776>
- Chaka, C. (2023). Detecting AI content in responses generated by ChatGPT, YouChat, and Chatsonic: The case of five AI content detection tools. *Journal of Applied Learning and Teaching*, 6(2), 94-104. <https://doi.org/10.37074/jalt.2023.6.2.12>
- Chaka, C. (2024). Reviewing the performance of AI detection



tools in differentiating between AI-generated and human-written texts: A literature and integrative hybrid review. *Journal of Applied Learning & Teaching*, 7(1), 115-126. <https://doi.org/10.37074/jalt.2024.7.1.14>

Childs, P. E. (2015). Curriculum development in science-past, present and future. *LUMAT: International Journal on Math, Science and Technology Education*, 3(3), 381-400. <https://doi.org/10.31129/lumat.v3i3.1036>

Chomsky, N., Roberts, I., & Watumull, J. (2023, March 8). Noam Chomsky: The false promise of ChatGPT. *The New York Times*. <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>

Clark, T. M. (2023). Investigating the use of an artificial intelligence chatbot with general chemistry exam questions. *Journal of Chemical Education*, 100(5), 1905-1916. <https://doi.org/10.1021/acs.jchemed.3c00027>

Copilot. (2024). *Copilot* (Mar 15 version). <https://copilot.microsoft.com>

Creswell, J. W. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

Daher, W., Diab, H., & Rayan, A. (2023). Artificial intelligence generative tools and conceptual knowledge in problem solving in chemistry. *Information*, 14(7), 409. <https://doi.org/10.3390/info14070409>

Dao, X. Q. (2023). *Which large language model should you use in Vietnamese education: ChatGPT, Bing Chat, or Bard?*. SSRN Electronic Journal. <http://dx.doi.org/10.2139/ssrn.4527476>

Dempere, J., Modugu, K., Hesham, A., & Ramasamy, L. K. (2023). The impact of ChatGPT on higher education. *Frontiers in Education*, 8, 1206936. <https://doi.org/10.3389/educ.2023.1206936>

Dis, E. A. V., Bollen, J., Zuidema, W., Van Rooij, R., & Bockting, C. L. (2023). ChatGPT: Five priorities for research. *Nature*, 614(7947), 224-226. <https://doi.org/10.1038/d41586-023-00288-7>

Doren, G., D., & Duffy, L. K. (2016). Teaching chemistry in rural America: A comparative study examining innovative informal education and applied science approaches to improving engagement. *International Research in Education*, 4(2), 48-62. <https://doi.org/10.5296/ire.v4i2.9537>

Editorials, N. (2023). Tools such as ChatGPT threaten transparent science; here are our ground rules for their use. *Nature*, 613(7945), 612. <https://doi.org/10.1038/d41586-023-00191-1>

Emenike, M. E., & Emenike, B. U. (2023). Was this title generated by ChatGPT? Considerations for artificial intelligence text-generation software programs for chemists and chemistry educators. *Journal of Chemical Education*, 100(4), 1413-1418. <https://doi.org/10.1021/acs.jchemed.3c00063>

Farazouli, A., Cerratto-Pargman, T., Bolander-Laksov, K., & McGrath, C. (2024). Hello GPT! Goodbye home examination? An exploratory study of AI chatbots impact on university teachers' assessment practices. *Assessment & Evaluation in Higher Education*, 49(3), 363-375. <https://doi.org/10.1080/02602938.2023.2241676>

Flores, R. A. R., & Peña, M. A. G. (2024). Aprendizaje Basado en Modelización asistido con Inteligencia Artificial en las Ciencias Naturales: propuesta de intervención neurodidáctica [Modeling-based learning assisted with artificial intelligence in natural sciences: Proposal for neurodidactic intervention]. *Práxis Educativa [Educational Praxis]*, 19, 1-19. <https://doi.org/10.5212/PraxEduc.v.19.22722.011>

García-Peñalvo, F. J., Llorens-Largo, F., & Vidal, J. (2024). The new reality of education in the face of advances in generative artificial intelligence. *Revista Iberoamericana de Educación a Distancia [Ibero-American Journal of Distance Education]*, 27(1), 9-39. <https://doi.org/10.5944/ried.27.1.37716>

Gardner, H. (1993). *Multiple intelligences: The theory in practice*. Basic Books.

Gemini. (2024). *Gemini* (Mar 15 version). <https://gemini.google.com/app>

Głushkova, T., & Malinova, A. (2024). Application of AI technologies in STEAM school education. *International Journal of Research in E-learning*, 10(1), 1-20. <https://journals.us.edu.pl/index.php/IJREL/article/view/15834/13295>

Göçmez, L., & Okur, M. R. (2023). Artificial intelligence applications in open and distance education: A systematic review of the articles (2007-2021). *Asian Journal of Distance Education*, 18(1), 1-32. <https://www.asianjde.com/ojs/index.php/AsianJDE/article/view/665>

Google. (2024). *Bard updates*. <https://bard.google.com/updates>

GPTZero. (2024). *GPTZero*. <https://gptzero.me>

Grassini, S. (2023). Shaping the future of education: Exploring the potential and consequences of AI and ChatGPT in educational settings. *Education Sciences*, 13(7), 692. <https://doi.org/10.3390/educsci13070692>

Gravel, J., D'Amours-Gravel, M., & Osmanlliu, E. (2023). Learning to fake it: Limited responses and fabricated references provided by ChatGPT for medical questions. *Mayo Clinic Proceedings: Digital Health*, 1(3), 226-234. <https://doi.org/10.1016/j.mcpgdig.2023.05.004>

Hallal, K., Hamdan, R., & Tlais, S. (2023). Exploring the potential of AI-Chatbots in organic chemistry: An assessment of ChatGPT and Bard. *Computers and Education: Artificial Intelligence*, 5, 100170. <https://doi.org/10.1016/j.caeai.2023.100170>

Holme, T. A., Luxford, C. J., & Brandriet, A. (2015). Defining conceptual understanding in general chemistry. *Journal of Chemical Education*, 92(9), 1477-1483. <https://doi.org/10.1021/acs.jchemed.5b00000>



Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., Santos, O. C.; Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2022). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education*, 32, 504-526. <https://doi.org/10.1007/s40593-021-00239-1>

Ismail, F., Tan, E., Rudolph, J., Crawford, J., & Tan, S. (2023). Artificial intelligence in higher education. A protocol paper for a systematic literature review. *Journal of Applied Learning and Teaching*, 6(2), 56-63. <https://doi.org/10.37074/jalt.2023.6.2.34>

IUPAC. (2019). *Compendium of chemical terminology* (2nd ed.). Blackwell Scientific Publications, Oxford (1997). <https://doi.org/10.1351/goldbook>

Johnson, B., & Christensen, L. (2008). *Educational research: Quantitative, qualitative, and mixed approaches*. Sage.

Johnson, P., Goldsmith, T., & Teague, K. (1994). Locus of the predictive advantage in pathfinder-based representations of classroom knowledge. *Journal of Educational Psychology*, 86(4), 617-626. <https://doi.org/10.1037/0022-0663.86.4.617>

Kapur, R. (2023). *AI made simple: A beginner's guide to generative intelligence*. Rinity Media.

Khademi, A. (2023). Can ChatGPT and Bard generate aligned assessment items? A reliability analysis against human performance. *Journal of Applied Learning and Teaching*, 6(1), 75-80. <https://doi.org/10.37074/jalt.2023.6.1.28>

King, R. K. (2023). A conversation on artificial intelligence, chatbots, and plagiarism in higher education. *Cellular and Molecular Bioengineering*, 16(1), 1-2. <http://dx.doi.org/10.1007/s12195-022-00754-8>

Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., Maningo, J., & Tseng, V. (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLoS Digital Health*, 2(2), e0000198. <https://doi.org/10.1371/journal.pdig.0000198>

Leite, B. S. (2018). Active technological learning. *Revista internacional de educação superior [International Journal of Higher Education]*, 4(3), 580-609. <https://doi.org/10.20396/riesup.v4i3.8652160>

Leite, B. S. (2022). *Tecnologias digitais na educação: da formação à aplicação [Digital technologies in education: From training to application]*. Livraria da Física.

Leite, B. S. (2023). Artificial intelligence and chemistry teaching: A propaedeutic analysis of ChatGPT in chemical concepts defining. *Química Nova [New Chemistry]*, 46(9), 915-923. <https://doi.org/10.21577/0100-4042.20230059>

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, 13(4),

Lozić, E., & Štular, B. (2023). Fluent but not factual: A comparative analysis of ChatGPT and other AI chatbots' proficiency and originality in scientific writing for humanities. *Future Internet*, 15(10), 336. <https://doi.org/10.3390/fi15100336>

Luzano, J. (2024). Pedagogical influence of an AI chatbot Gemini in mathematics education. *International Journal of Academic Pedagogical Research*, 8(4), 107-112. <http://ijeais.org/wp-content/uploads/2024/4/IJAPR240419.pdf>

LuzIA. (2023). *Your intelligent assistant*. <https://www.luzia.com/en>

McGill, T. L., Williams, L. C., Mulford, D. R., Blakey, S. B., Harris, R. J., Kindt, J. T., Lynn, D. G., Marsteller, P. A., McDonald, F. E., & Powell, N. L. (2019). Chemistry unbound: Designing a new four-year undergraduate curriculum. *Journal of Chemical Education*, 96(1), 35-46. <https://doi.org/10.1021/acs.jchemed.8b00585>

Mecaj, S. E. (2022). Artificial Intelligence and legal challenges. *Revista Opinião Jurídica [Legal Opinion Magazine]*, 20(34), 180-196. <https://doi.org/10.12662/2447-6641oj.v20i34.p180-196.2022>

Meta. (2024). *Llama 2 (meta)*. <https://llama.meta.com>

Miao, F., & Holmes, W. (2021). *International forum on AI and the futures of education developing competencies for the AI era*. UNESCO.

Motlagh, N. Y., Khajavi, M., Sharifi, A., & Ahmadi, M. (2023). *The impact of artificial intelligence on the evolution of digital education: A comparative study of OpenAI text generation tools including ChatGPT, Bing Chat, Bard, and Ernie*. arXiv preprint. <https://doi.org/10.48550/arXiv.2309.02029>

Nguyễn, P., Trương, H., Nguyễn, P., Bruneau, P., Cao, L., & Wang, J. (2023). *Google Bard's performance on Vietnamese high school civic education examination*. SocArXiv. <https://doi.org/10.31235/osf.io/uqhts>

Ogunleye, B., Zakariyyah, K. I., Ajao, O., Olayinka, O., & Sharma, H. (2024). Higher education assessment practice in the era of generative AI tools. *Journal of Applied Learning and Teaching*, 7(1), 46-56. <https://doi.org/10.37074/jalt.2024.7.1.28>

OpenAI. (2024). *ChatGPT* (Mar 14 version 3.5). <https://chat.openai.com/chat>

Oxford Learners Dictionaries. (2024). <https://www.oxfordlearnersdictionaries.com>

Parker, L., Carter, C., Karakas, A., Loper, A. J., & Sokkar, A. (2024). Graduate instructors navigating the AI frontier: The role of ChatGPT in higher education. *Computers and Education Open*, 6, 100166. <https://doi.org/10.1016/j.caeo.2024.100166>

- Paul, J., Ueno, A., & Dennis, C. (2023). ChatGPT and consumers: Benefits, pitfalls and future research agenda. *International Journal of Consumer Studies*, 47(4), 1213-1225. <https://doi.org/10.1111/ijcs.12928>
- Pavlik, J. V. (2023). Collaborating with ChatGPT: Considering the implications of generative artificial intelligence for journalism and media education. *Journalism & Mass Communication Educator*, 78(1), 84-93. <https://doi.org/10.1177/10776958221149577>
- Perera, P., & Lankathilake, M. (2023). Preparing to revolutionize education with the multi-model GenAI tool Google Gemini? A journey towards effective policy making. *Journal of Advances in Education and Philosophy*, 7(8), 246-253. <https://doi.org/10.36348/jaep.2023.v07i08.001>
- Perkins, M., & Roe, J. (2024). The use of generative AI in qualitative analysis: Inductive thematic analysis with ChatGPT. *Journal of Applied Learning and Teaching*, 7(1), 390-395. <https://doi.org/10.37074/jalt.2024.7.1.22>
- Plagium. (2024). *Plagium originality AI detector*. [https://www.plagium.com/en/ai\\_detector](https://www.plagium.com/en/ai_detector)
- Popenici, S. (2023). The critique of AI as a foundation for judicious use in higher education. *Journal of Applied Learning and Teaching*, 6(2), 378-384. <https://doi.org/10.37074/jalt.2023.6.2.4>
- Qian, Y., Wang, Y., Wen, J., Wu, S., & Zhang, J. (2023). One hundred core concepts in chemistry and upper-secondary school teachers' and students' chemistry conceptual structures. *Journal of Baltic Science Education*, 22(3), 493-505. <https://doi.org/10.33225/jbse/23.22.493>
- Reis, R. M. D. S., Leite, B. S., & Leão, M. B. C. (2021). Estratégias didáticas envolvidas no uso das TIC: o que os professores dizem sobre seu uso em sala de aula?. [Didactic strategies involved in the use of ICT: What teachers say about your use in a classroom?] *ETD Educação Temática Digital [ETD - Digital Thematic Education]*, 23(2), 551-571. <https://doi.org/10.20396/etd.v23i2.8657601>
- Rudolph, J., Tan, S., & Tan, S. (2023). War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. *Journal of Applied Learning and Teaching*, 6(1), 364-389. <https://doi.org/10.37074/jalt.2023.6.1.23>
- Rudolph, J., Ismail, F., & Popenici, S. (2024a). Higher education's generative artificial intelligence paradox: The meaning of chatbot mania. *Journal of University Teaching and Learning Practice*, 21(6), 1-35. <https://open-publishing.org/journals/index.php/jutlp/article/view/744>
- Rudolph, J., Tan, S., & Ismail, F. (2024b). Joyce's Odyssey. A celebration of human ingenuity in Ulysses and an indictment of the mediocrity of generative AI. *Journal of Applied Learning and Teaching*, 7(1), 07-21. <https://doi.org/10.37074/jalt.2024.7.1.1>
- Sadykov, T., & Čtrnáctová, H. (2019). Application interactive methods and technologies of teaching chemistry. *Chemistry Teacher International*, 1(2), 20180031. <https://doi.org/10.1515/cti-2018-0031>
- Schäfer, M. S. (2023). The Notorious GPT: Science communication in the age of Artificial Intelligence. *Journal of Science Communication*, 22(2), 1-15. <https://doi.org/10.22323/2.22020402>
- Silva Neto, S. L. D., & Leite, B. S. (2024). Inteligência artificial en la mejora de los ensayos de Ecología: Un estudio en una escuela secundaria brasileña [Artificial Intelligence in improving ecology trials: A study in a Brazilian high school]. *Educación*, 33(64), 86-108. <https://doi.org/10.18800/educacion.202401.M004>
- Taber, K. (2013). Revisiting the chemistry triplet: Drawing upon the nature of chemical knowledge and the psychology of learning to inform chemistry education. *Chemistry Education Research And Practice*, 14(2), 156-168. <https://doi.org/10.1039/C3RP00012E>
- Tomasevic, B., & Trivic, D. (2014). Creativity in teaching chemistry: How much support does the curriculum provide?. *Chemistry Education Research and Practice*, 15(2), 239-252. <https://doi.org/10.1039/C3RP00116D>
- Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., ... Scialom, T. (2023). *Llama 2: Open foundation and fine-tuned chat models*. arXiv preprint. <https://doi.org/10.48550/arXiv.2307.09288>
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 5(236), 433-460. <http://dx.doi.org/10.1093/mind/LIX.236.433>
- Upahi, J. E., & Ramnarain, U. (2019). Representations of chemical phenomena in secondary school chemistry textbooks. *Chemistry Education Research and Practice*, 20(1), 146-159. <https://doi.org/10.1039/C8RP00191J>
- Van Wyk, M. M. V. (2024). Is ChatGPT an opportunity or a threat? Preventive strategies employed by academics related to a GenAI-based LLM at a faculty of education. *Journal of Applied Learning and Teaching*, 7(1), 35-45. <https://doi.org/10.37074/jalt.2024.7.1.15>
- Vygotsky, L. S. (2012). *Thought and language*. MIT Press.

Waring, P. (2024). Artificial intelligence and graduate employability: What should we teach Generation AI?. *Journal of Applied Learning and Teaching*, 7(1), 22-25. <https://doi.org/10.37074/jalt.2024.7.1.42>

Zhai, X. (2022). *ChatGPT user experience: Implications for education*. SSRN. <http://dx.doi.org/10.2139/ssrn.4312418>

Copyright: © 2024. Bruno Silva Leite. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.