



SEÇÃO: IA NOS PROCESSOS DE ENSINO-APRENDIZAGEM

Case study on the use of a chatbot as a pedagogical tool for teaching Diels-Alder reactions in undergraduate education¹

Estudio de caso sobre el uso de un *chatbot* como herramienta pedagógica para la enseñanza de Diels-Alder en el pregrado

Estudo de caso sobre o uso de um *chatbot* como ferramenta pedagógica para ensino de reações de Diels-Alder na graduação

Wilton José Diolindo do Nascimento Júnior², Mayara de Carvalho Santos³,
Paulo César Muniz de Lacerda Miranda⁴, Gildo Giroto Júnior⁵

ABSTRACT

This study investigated the employment of the IQ.QO Assistente chatbot, developed by one of the authors to specialize in organic chemistry teaching with knowledge restricted to reliable sources. The research adopted a case study conducted with 10 third-period Chemistry students during the first semester of 2024. The chatbot's susceptibility to conceptual errors was analyzed in comparison with broad-access language models. Additionally, by content analysis of 30 voluntarily collected prompts, the study sought to identify patterns in the formulation of questions by undergraduates. Results showed that the chatbot presented an error rate of only 11%, significantly lower than that of general models. Prompt analysis

¹ The authors were responsible for translating this article into English.

² Universidade Estadual de Campinas (Unicamp), Campinas, SP, Brazil.
ORCID ID: <https://orcid.org/0000-0002-9964-4572>. E-mail: wiltonjdn@gmail.com

³ Universidade Estadual de Campinas (Unicamp), Campinas, SP, Brazil.
ORCID ID: <https://orcid.org/0000-0003-1204-0184>. E-mail: decarvalho.mayara@gmail.com

⁴ Universidade Estadual de Campinas (Unicamp), Campinas, SP, Brazil.
ORCID ID: <https://orcid.org/0000-0002-6418-3614>. E-mail: pmiranda@unicamp.br

⁵ Universidade Estadual de Campinas (Unicamp), Campinas, SP, Brazil.
ORCID ID: <https://orcid.org/0000-0001-9933-100X>. E-mail: ggirotto@unicamp.br

revealed a tendency toward simplicity, with an emphasis on input data (63.5%) and an absence of context, suggesting students use generative artificial intelligence similarly to search engines. Such findings reinforce the need for digital literacy for the effective use of artificial intelligence tools in educational contexts, promoting the development of digital competencies and the transition to active learning models.

Keywords: artificial intelligence; chatbot; chemistry education; organic chemistry; Diels-Alder.

RESUMO

O estudo em questão investigou o emprego do *chatbot* IQ.QO Assistente, desenvolvido por um dos autores para ser especializado no ensino de química orgânica, com conhecimento restrito a fontes confiáveis. A pesquisa é um estudo de caso realizado com dez estudantes do terceiro período de um curso de Química durante o primeiro semestre de 2024. Analisou-se a suscetibilidade do *chatbot* a erros conceituais em comparação com modelos de linguagem de acesso amplo. Adicionalmente, por meio da análise de conteúdo de trinta *prompts* coletados voluntariamente, buscou-se identificar padrões na elaboração de perguntas pelos graduandos. Os resultados mostraram que o *chatbot* desenvolvido apresentou uma taxa de erro de apenas 11%, significativamente menor do que a de modelos gerais. A análise dos *prompts* revelou uma tendência à simplicidade, com ênfase nos dados de entrada (63,5%) e ausência de contexto, sugerindo que os estudantes utilizam inteligências artificiais generativas de forma semelhante a motores de busca. Tais dados reforçam a necessidade de *literacia* digital para o uso eficaz de ferramentas de inteligência artificial no contexto educativo, promovendo o desenvolvimento de competências digitais e a transição para modelos de aprendizagem ativa.

Palavras-chave: inteligência artificial; *chatbot*; ensino de química; química orgânica; Diels-Alder.

RESUMEN

El presente estudio investigó el uso del *chatbot* IQ.QO Asistente, desarrollado por uno de los autores para especializarse en la enseñanza de química orgánica, con una base de conocimiento restringida a fuentes fiables. La investigación es un estudio de caso realizado con diez estudiantes del tercer período de la carrera de Química durante el primer semestre de 2024. Se analizó la susceptibilidad del *chatbot* a errores conceptuales en comparación con modelos de lenguaje de acceso amplio. Adicionalmente, mediante el análisis de contenido de treinta *prompts* recopilados voluntariamente, se buscaron identificar patrones en la elaboración de preguntas por parte de los estudiantes de pregrado. Los resultados mostraron

que el *chatbot* desarrollado presentó una tasa de error de solo el 11%, significativamente menor que la de los modelos generales. El análisis de los *prompts* reveló una tendencia a la simplicidad, con énfasis en los datos de entrada (63,5%) y ausencia de contexto, lo que sugiere que los estudiantes utilizan inteligencias artificiales generativas de forma similar a los motores de búsqueda. Dichos datos refuerzan la necesidad de la alfabetización digital para el uso eficaz de herramientas de inteligencia artificial en el contexto educativo, promoviendo el desarrollo de competencias digitales y la transición hacia modelos de aprendizaje activo.

Palabras clave: inteligencia artificial; *chatbot*; enseñanza de química; química orgánica; Diels-Alder.

INTRODUCTON

“This is happy to be of service” – this was how the android Andrew Martin, in *Bicentennial Man*⁶, reported to his owners every time he performed a task. It is no coincidence that stories about machines reaching human intelligence have populated the imagination since they first emerged. The conception of an entity capable of reasoning in a human manner, yet with unrestricted access to a vast repository of knowledge, can simultaneously evoke enthusiasm and apprehension. Far from Asimovian⁷ stories and the dystopias in which machines replace humans, it is pertinent to discuss the possibilities of using artificial intelligence (AI) and its implications in the most diverse fields of knowledge.

In this work, we present and discuss the implications of using a chatbot specifically developed to assist students in the Organic Chemistry II course, during the third period of an undergraduate degree at a Brazilian public university. Through testing, we sought to verify how the 10 undergraduates utilized this tool, their usage patterns, and the adequacy of the chatbot's responses. The study focused on the analysis of questions and answers related to content on the Diels-Alder reaction, an important topic that frequently presents challenges in organic chemistry education. This difficulty is primarily due to its three-dimensional nature, which requires spatial reasoning from students to understand the overlap of molecular orbitals, the stereochemistry of products, and the kinetic factors governing the reaction.

⁶ A 1999 film directed by Chris Columbus and based on the work of Isaac Asimov. The plot follows the 200-year journey of the android Andrew (played by Robin Williams), who transcends his initial programming to develop complex characteristics such as creativity, emotions, and the yearning to be recognized as a human being..

⁷ Isaac Asimov (1920-1992) was an American writer and professor of biochemistry, considered one of the greatest science fiction writers. In his works, such as the collection *I, Robot* and the *Foundation* series, Asimov deeply explored the relationship between humans and intelligent machines. Asimov is famous for proposing the Three Laws of Robotics, a code of ethics for artificial intelligence that became a pillar in the genre and influenced the philosophical debate on the subject.

Specifically regarding the interaction of undergraduates with the created chatbot, we analyzed the quality of the prompts used to understand how students utilize this tool in their personal studies. This quality is a determining factor for the utility of the AI response and, according to Giray (2023), the parameters for this quality are established by the presence and clarity of four essential components: the instruction or task, the context, the input data, and the output indicator.

Although the term prompt has gained new prominence with the popularization of AIs, its use in the educational field is longstanding, traditionally referring to a stimulus or instruction aimed at evoking a specific response from the learner. In the context of generative artificial intelligence (GenAI), this notion is expanded: prompts are instructions or commands, generally formulated as questions, provided to a language model, such as a chatbot, to guide its behavior and generate desired responses. It is the appropriate way to communicate with AIs, guiding the specification of the task, detailing the context, providing input data, and specifying the desired response format (Giray, 2023). In general terms, prompts serve to improve the quality of the responses obtained. They can be used for any task and, when properly prepared, guide the language model to produce more accurate and relevant results for the user.

Given this context, this research was guided by the following objectives:

1. To investigate the propensity of the chatbot for conceptual errors, especially regarding the precision of definitions, reaction mechanisms, and primarily the representations of molecular structures in the Diels-Alder reaction.
2. To analyze the forms of interaction between undergraduates and AI, examining the quality of the prompts they develop.

To achieve these objectives, the methodology was designed to answer the following research questions: Does a knowledge-base-limited chatbot commit fewer errors than other general-purpose language models? Do the analyzed undergraduates develop complete prompts to interact with the chatbot? Does the quality of the prompts interfere with the quality of the responses generated by the limited chatbot?

From the answers to these questions, we expect to contribute to the definition of best practices in the use of AIs in university contexts. Although the study focuses on a topic of organic chemistry, the findings offer pedagogical implications that can be applied to other teaching areas. The need for curation of sources for the AI and for training students in prompt engineering indicates a model of technological integration that can be adapted to different contents and disciplines that require conceptual rigor. Thus, the main contribution of this work

lies in the reflection on how to plan and mediate the use of AIs in higher education, transforming them into effective tools to promote active learning.

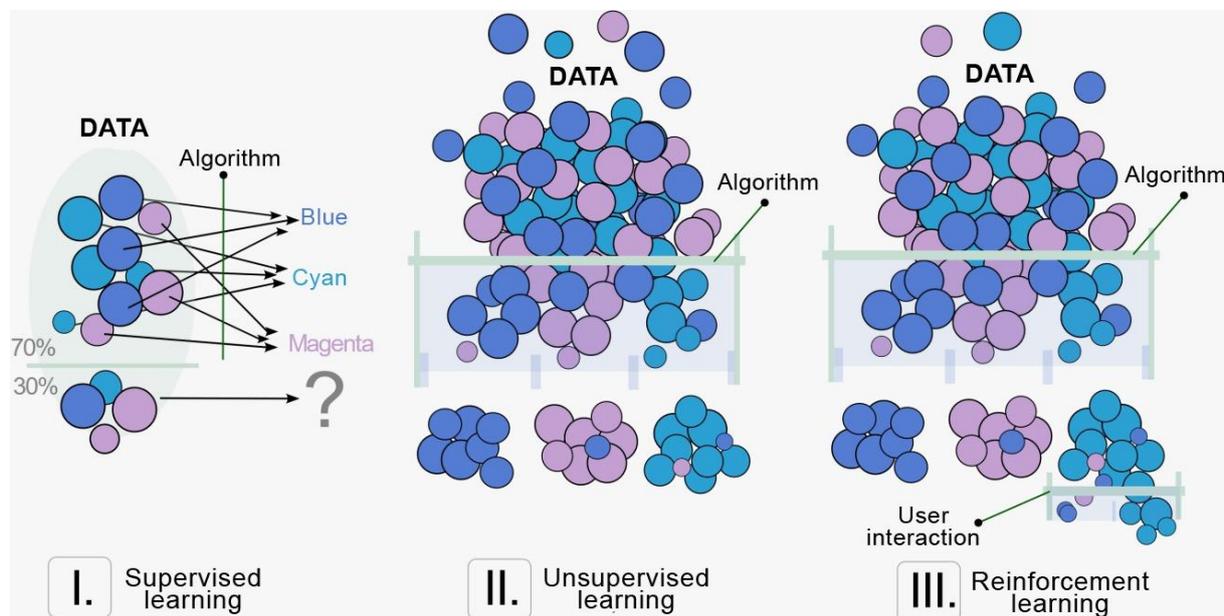
It is important, however, to recognize the limitations of this study to contextualize its findings. The main limitation, as noted, is the scope focused on a single topic of organic chemistry, the Diels-Alder reaction. Although this allowed for an in-depth analysis, the effectiveness and usage patterns of the chatbot in other themes of the discipline were not evaluated. Moreover, the study was conducted with a small sample of participants, which, while characteristic of a qualitative case study, requires caution in generalizing the observed interaction patterns. Finally, as with all research involving AI, the results represent a snapshot of a specific moment and the rapid evolution of these technologies underscores the need for continuous validation and ongoing research.

THEORETICAL FRAMEWORK

After all, what is artificial intelligence (AI)? Generally speaking, artificial intelligence refers to the ability of machines to display intelligent behaviors, such as learning, making decisions based on data, and making predictions. They utilize techniques such as machine learning, natural language processing, and computer vision to perform a wide variety of tasks (Alpaydin, 2010). AIs can be developed considering supervised learning, unsupervised learning, and reinforcement learning (Russel, 2021).

In supervised learning (Figure 1, Block I), a predictive model is created from a dataset and its known results, enabling the machine to make predictions for new data. In unsupervised learning (Figure 1, Block II), the algorithm analyzes data without prior identification to recognize patterns autonomously, which is useful in processing and clustering raw data. Finally, in reinforcement learning (Figure 1, Block III), the model is created based on trends in the data, and subsequently, the user intervenes to indicate improvements that are tested iteratively (Alpaydin, 2010).

Figure 1 – (Block I) Supervised learning scheme. (Block II) Unsupervised learning scheme. (Block III) Reinforcement or semi-supervised learning scheme.



Source: the authors.

Effective communication with GenAIs is a technical skill known as prompt engineering, which consists of drafting clear instructions to guide the behavior of language models. The main characteristics of a prompt are: instruction, context, input data, and output indicator, according to Giray (2023). Instruction defines the task and can be characterized as “write a summary” or “translate the text.” Additional information that helps the model understand the task, such as delimiting the audience or writing style, defines the context. Input data are the specific elements the model is expected to process, which can be text, an image, or even a set of instructions. It is also important to indicate the desired response format, whether by delimiting the number of characters in the text or if the user prefers the data to be presented in the form of a table, for example (Giray, 2023).

In this sense, the use of prompts improves the precision and relevance of the obtained responses, controls the format and content of the output, and automates tasks, making communication with AI more effective. The ability to draft a prompt is intrinsically linked to what is known as Information Literacy (InfoLit), a necessary capability not only in the school context but for the information society as a whole. InfoLit reveals the ability to identify and define what information is needed to answer a question, clarify a doubt, or solve a problem (Kuhlthau, 1991). It is what motivates the planning of strategies to obtain information and is in action when we interact with search engines or AIs, based on the awareness of a knowledge gap and the ability to formulate questions.

To draft prompts with the four characteristics, users may follow these recommendations:

Clear Task: Explicitly define the task to be performed, using verbs such as “explain,” “summarize,” or “compare.” Avoid ambiguities and be as specific as possible.

Contextualization: Provide data about the context, such as your role (e.g., “I am a student...”) and the topic under study (e.g., “we are analyzing...”). This directs the AI to the desired level and specificity of the response.

Precise Data: Present specific and relevant data on the subject, using appropriate technical terms. The clarity and detail of the input data are essential for the quality of the response.

Desired Format: Indicate the format in which the response should be presented, such as a concise summary, an organized list, or a detailed explanation. This step allows for shaping the response according to your needs (Giray, 2023).

Als have been considered relevant technological tools for users in various fields, including the scientific community. In chemistry, a significant increase in publications related to AI is observed, especially in analytical chemistry and biochemistry (Baum et al., 2021), highlighting their functionalities and the expansion of use, which opens space for new integrations and future developments.

Recent studies describe other types of artificial intelligence capable of predicting the three-dimensional structure of proteins based on amino acid residues (Senior et al., 2020) and performing screenings to assist in drug discovery (Sellwood et al., 2018). These systems can also support synthesis planning (Segler; Preuss; Waller, 2018), predict molecular properties (Kuntz; Wilson, 2022), and even contribute to solving complex wave functions (Carleo; Troyer, 2017), among many other possibilities. The utility of AI in these research areas presents great potential, as many of these tasks represent challenges of extremely high complexity that, until recently, required decades of research or were considered computationally intractable.

The use of technologies within the educational scope is not a recent practice. Teachers and students have explored a variety of resources, such as software (Scott; Bohaty; Gadbury-Amyot, 2021), applications (Grando; Cleophas, 2021), videos (Santos et al., 2017), animation (Lopes; Chaves, 2018), digital games (Tauber; Levonis; Schweiker, 2022), and the internet itself (Leite, 2018). This set of tools is known in the educational field as Digital Information and Communication Technologies (DICT) (Leite, 2015), a research field that has been widely explored by researchers in the area (Leite, 2020). Furthermore, there are norms and guidelines that highlight the importance of developing activities that promote the improvement of digital

competencies (Silva; Behar, 2019) for students and teachers (ISTE, 2018; Profuturo, 2020; Unesco, 2024).

Despite the broad spectrum of research in the area, the use of these technological resources in the educational context still generates doubts and insecurities. For instance, the introduction of Wikipedia in 2001 sparked vigorous debates in academic institutions regarding how this tool could potentially foster intellectual passivity in students (Knight; Pryke, 2012). When a teacher proposes an activity, students may seek information directly from websites and blogs, rewriting the texts and attempting to relate them consistently to the requested work, often without considering the reliability of the information sources. AIs can perform the same task almost instantaneously with just one click, depriving the student of this entire search process, which generates a justifiable concern among education professionals (Leite, 2024).

There are also concerns regarding the ethics of AI use by students, ranging from the privacy of their learning data to the risk of algorithmic bias in assessments and the need to ensure academic integrity. Therefore, it is necessary to think about its implementation in the classroom critically, weighing the risks and benefits. Despite these challenges, the pedagogical potential is notable, and according to Luckin et al. (2022), AI can be used primarily in three ways in the educational domain: personal tutoring for each learner, intelligent support for collaborative learning, and intelligent virtual reality.

In the first scenario, AI would monitor the student's progress in real-time, identifying gaps and offering feedback, complementary resources, and formative assessments so they can improve their incipient skills. In the second case, AI would evaluate different profiles in the class, assisting the teacher in forming groups by considering affinities and specific characteristics. Finally, AI-driven instructional materials could be made available to students in a virtual reality format, providing a more immersive sensory experience and facilitating the understanding of abstract concepts common to chemistry.

For this study, the selected chemical topic was the Diels-Alder reaction, a relevant concept for synthetic organic chemistry that presents a high degree of complexity. It is one of the most important tools for constructing six-membered ring molecules, which form the basis for the synthesis of numerous pharmaceutical products, including certain antibiotics and anti-inflammatory drugs, as well as polymers and agrochemicals. (Funel; Abele, 2013; Hermanson, 2013; Sauer, 1967; Yeingst; Helton; Hayes, 2024).

The main didactic challenge of this reaction lies in its intrinsically three-dimensional nature. Students frequently face difficulties in visualizing how molecules approach each other in

space, how their electron clouds interact, and, primarily, how the spatial orientation of the reactants determines the three-dimensional structure of the final product. It is precisely to overcome these barriers of abstraction that the use of AI seems promising. Tools such as the developed chatbot can generate dynamic visual representations, offer step-by-step explanations, and answer specific questions on the stereochemistry of the reaction, providing personalized learning support that complements traditional teaching methods.

METODOLOGY

Chatbot development: IQ.QO Assistente

The chatbot entitled IQ.QO Assistente, developed by one of the authors of this research, was built upon a knowledge base selected by experts in the field of organic chemistry, consisting of 20 internationally recognized academic sources. These sources were chosen for their high degree of technical precision and scientific rigor, including works such as *Advanced Organic Chemistry Part A: Structure and Mechanism* and *Writing Reaction Mechanisms in Organic Chemistry*, as well as virtual libraries like the *Virtual Textbook of Organic Chemistry* and *LibreTexts Chemistry*.

Notably, the default settings of the chatbot do not allow for searches on websites other than those included in its database. In this way, we prevent responses from being generated from unreliable sources or those with greater availability on the internet, bypassing a problem common to conventional chatbots. This approach of careful curation of the chatbot database, while not completely eliminating the possibility of errors, was specifically designed to minimize the occurrence of conceptual inaccuracies and hallucinations—a common challenge in AIs with broader and unverified databases.

Classroom application

The research was conducted as a case study from the perspective of Stake (1995), which values the in-depth analysis of a particular case to understand a broader phenomenon. Participants included 10 students in their third period of the Chemistry program, enrolled in the Organic Chemistry II course at a São Paulo state public university.

Students were given access to IQ.QO Assistente via a link provided by the instructor after a series of lectures on the topic and were encouraged to use it as a study support tool. Except for the method of access, interaction with the chatbot was not directed, allowing students to formulate their prompts freely. Although doubts covered other topics of the course, the selection criterion for this study was interactions related to the Diels-Alder reaction, resulting

in an analytical corpus composed of 30 prompts shared by the 10 participants regarding the content.

It is important to highlight that, even though they created IQ.QO Assistente, the researchers did not have access to any content or student interactions with the chatbot. All students were informed about the research objectives and asked to send the researcher in charge the files of their choice regarding their personal interactions with the chatbot. Therefore, the corpus of this research is entirely composed of materials voluntarily shared by the students who consented to participate.

Due to the volume of data collected, we filtered the interactions considering only those related to the Diels-Alder topic, resulting in the analysis of materials from 10 undergraduates, totaling 30 prompts. When the accessed material was sent as an image, it was transcribed using the AI service Google Gemini in the Advanced 2.0 Pro Experimental version. The validation process consisted of manual verification of each transcription by two researchers, who compared the generated text with the original images to ensure data reliability.

The corpus was then organized for qualitative analysis using MaxQDA software, following the content analysis methodology (Bardin, 2015). Notably, the comparative data with other AI models, such as ChatGPT, cited in the results section, were extracted from a previous study (Nascimento Junior; Moraes; Giroto Junior, 2024), which employed an analogous methodology for evaluating general-purpose AIs.

The categories created are *a priori*, based on the prompt model by Giray (2023), namely: context (CT), instruction (IT), input data (ID), and output indicator (OI). The iteration (ITE) category was added to specify cases in which student iteration occurred based on the previous question. In addition to these categories, the chatbot's responses were categorized into correct passage (CP) and incorrect passage (IP) regarding the conceptual precision of the Diels-Alder content. Table 1 presents the categories used in the analysis based on the expected characteristics of a prompt.

Table 1 – A priori categories for the research corpus analysis.

Category	Description
Context (CT)	External information or additional context that provides background knowledge to the model, helping it generate more accurate and relevant responses.
Instruction (IT)	A specific task or instruction that guides the model's behavior and directs it toward the desired output.
Input Data (ID)	The input or question we want the model to process and provide an answer for—in this case, data related to Chemistry content.
Output Indicator (OI)	Specifies the desired output format. Helps shape the response by defining if a short answer, a paragraph, or any other specific format is needed.
Iteration (ITE)	Specifies a question derived from the previous question.
Correct Passage (CP)	Specifies the passages of the AI-generated text that are conceptually correct.
Incorrect Passage (IP)	Specifies the passages of the AI-generated text that are not conceptually correct.

Source: adapted from Giray, 2023.

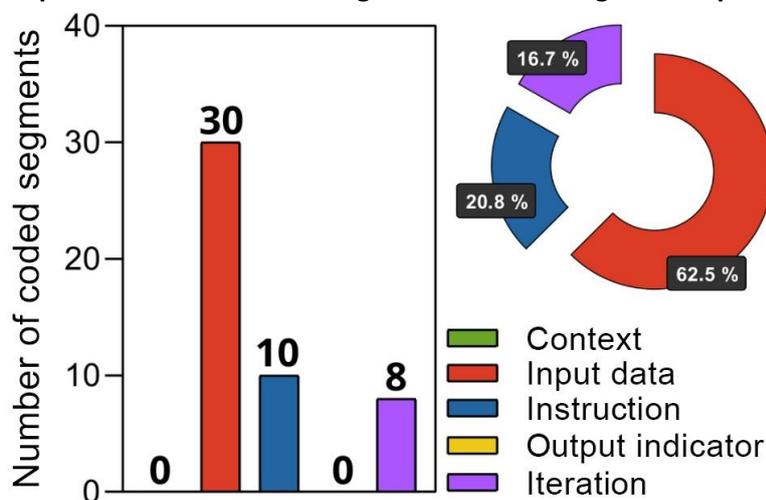
The analysis of these records enabled the identification of usage patterns, the most frequent doubts regarding the chemical content, and the effectiveness of the chatbot as a learning support tool in the educational context.

RESULTS AND DISCUSSION

The content analysis of the students' interactions with the developed chatbot resulted in the identification of 461 segments, distributed among the six pre-established categories: (CT, IT, ID, OI, ITE, CP, and IP). These segments were extracted from the categorization of complete sentences within the prompts developed by the students and the responses generated by the chatbot, specifically regarding the topic of Diels-Alder reactions.

The quantitative distribution of the 461 segments revealed patterns both in the structuring of prompts by students and in the quality of chemical knowledge present in the AI's responses. Graph 1 illustrates this distribution.

Graph 1 – Distribution of categories characterizing student prompts.



Source: the authors. Graph produced with the aid of GraphPad Prisma and MaxQDA softwares.

Regarding the structure of the prompts created by the students, out of 30 analyzed questions, 48 segments were categorized, with a predominance of ID (30 segments, 62.5%), followed by IT (10 segments, 20.8%) and ITE (08 segments, 16.7%). The concentration on ID indicates that students tend to adopt a direct approach in their queries. This means they focus primarily on the specific chemistry content for which they seek information, rather than more sophisticated prompt engineering strategies, such as using context or adding a desired format for the final response. In a study conducted by Tassoti (2024), students tended not to contextualize their questions, as they basically copy and paste them when interacting with the machine. A hypothesis for this behavior is the idea that students use AIs as they use conventional search engines, such as Google, Bing, and DuckDuckGo.

The presence of few iteration questions (16.7%), aligned with the high ID concentration, suggests that students do not perform refinements or follow-ups on previous queries. This may indicate an acceptance of the initial response as true or sufficient, a misunderstanding of the response provided by the AI, or a lack of reflection on the answer. Of the 10 participants in the study, it was found that only a single subject proceeded to rectify information. In Tassoti's (2024) research, the same behavioral phenomenon occurred among students who, despite identifying IPs in the responses, did not correct the machine.

Although this apparent acceptance of responses by students may engage with the debate on technological determinism, this perspective – which is not a consensus – stems from the belief that technological systems inevitably outperform human performance. However, this view of trust coexists with a more skeptical stance from other users and researchers, who warn of the limits and biases of these tools (Correa; Geremias, 2013).

Case study on the use of a chatbot as a pedagogical tool for teaching Diels-Alder reactions in undergraduate education

*Wilton José Diolindo do Nascimento Júnior, Mayara de Carvalho Santos,
Paulo César Muniz de Lacerda Miranda, Gildo Giroto Júnior*

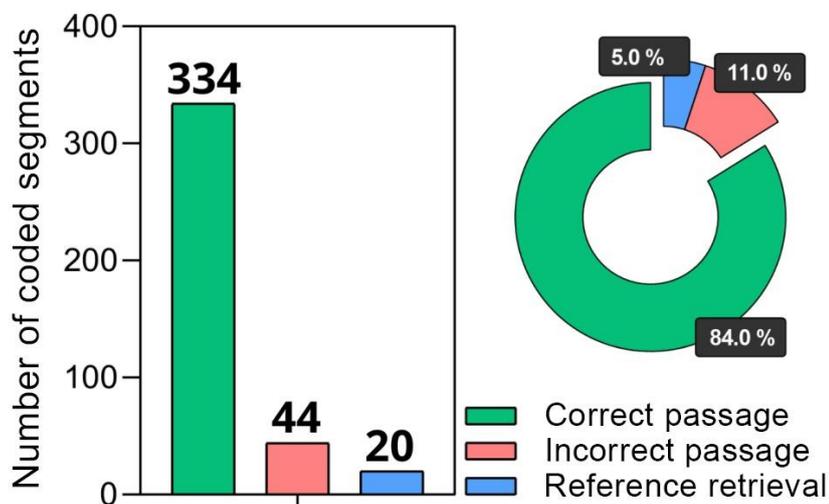
The possible misunderstanding of the generated response, if it occurred, may be related to the complexity of the Diels-Alder reaction content. Even if a restrictive-based AI generates more reliable answers, this does not guarantee they are simplified or easy to understand, as content interpretation depends on the theoretical framework the undergraduate already possesses. In other words, the AI acts as a supporting tool, but its effectiveness is linked to the student's prior knowledge to interpret the response. In this scenario, an undergraduate with an insufficient theoretical basis may not be able to assimilate the AI's explanation, even if it is correct. Ideally, they would continue the iteration with the machine to clear up doubts, but the difficulty in understanding the initial response may discourage this iteration. Subsequent studies may evaluate the occurrence of this possible loss of student motivation.

In the same vein, complementarily, undergraduates using a chatbot to facilitate their studies and optimize learning must – like any study not supported by digital technologies – recognize their need for information. This skill is the core of InfoLit, discussed previously, which manifests as the ability to formulate questions that effectively explore a knowledge gap. Therefore, the low quality observed in the prompts may be a direct reflection of a difficulty in exercising this competence even before interacting with the AI.

The complete absence of context units (CT) and output indicators (OI) reveals that undergraduates, in general, neither contextualize their queries with additional information nor specify the desired format for the responses, suggesting little familiarity with prompt engineering techniques. This is corroborated by the study of Araújo and Saúde (2024), which showed how iterative prompt refinement can improve the quality of ChatGPT responses in generating chemical laboratory protocols, indicating the importance of AI literacy for both educators and students.

Despite the absence of CT and OI in the prompts, a high percentage (84%) of CP was observed in the responses provided by the chatbot (Graph 2). This suggests that the AI model was capable of generating chemically accurate content even from poorly developed prompts.

Graph 2 – Distribution of categories characterizing chatbot responses.



Source: the authors. Graph produced using GraphPad Prism and MaxQDA software.

These results suggest that restricting the chatbot's knowledge sources to reliable websites may have been a key factor in reducing incorrect responses and in the absence of AI hallucinations. Table 2 brings representative excerpts from the undergraduate students' prompts and the chatbot's responses.

Table 2 – Examples of prompt segments and responses extracted from student interactions.

Prompt segment	Response segment
In Diels-Alder, is the kinetic product endo or exo? (Student 1)	The preference for the endo product under kinetic control is explained by the Alder rule. This rule establishes that, under kinetic conditions, the endo product is favored due to secondary interactions between the π orbitals of the diene and the substituents [...].
What happens to the orbitals of the diene and dienophile in an endo Diels-Alder reaction? (Student 3)	[...] The diene provides a Highest Occupied Molecular Orbital (HOMO).
Let's talk about the Diels-Alder reaction (Student 12)	[...] This reaction involves the [4+2] cycloaddition between a conjugated diene and a dienophile, resulting in the formation of a six-membered ring.
When will the demand of the Diels-Alder reaction be inverse? (Student 9)	In the inverse electron-demand Diels-Alder reaction, the diene is electron-poor and the dienophile is electron-rich. This situation occurs when the diene has electron-withdrawing groups and the dienophile has electron-donating groups. [...]

Source: the authors.

The prompts are brief and specific to the Diels-Alder reaction. Prompts such as “In Diels-Alder, is the kinetic product endo or exo?” (Student 1) and “What happens to the orbitals of the diene and dienophile in an endo Diels-Alder reaction?” (Student 3) direct the AI in its search for the knowledge used to generate the response, therefore being classified within the ID category. It is noted that the prompts, even without context and with simple instructions, vary in complexity, from very simple and general prompts like “let’s talk about...” (Student 12) to more elaborate prompts with terms like “orbitals” and “reaction demand” (Student 9).

The responses employ specialized and correct chemical terminology, such as “secondary interactions between π orbitals” (chatbot response to Student 1), “(HOMO)” (response to Student 3), and “[4+2] cycloaddition” (response to Student 12). The AI provides detailed textual explanations, as exemplified in the response to Student 9 regarding inverse electron demand, explaining the difference in electronic distribution between diene and dienophile.

In the topic regarding normal or inverse reaction demand, the responses focused mainly on normal demand, as expected, given that the materials used emphasize this mechanism. Passages such as “the diene provides a Highest Occupied Molecular Orbital (HOMO)” were considered correct. However, a more critical analysis would compel us to consider it an incomplete response. It would be more appropriate if the AI complemented the reasoning, emphasizing that the diene could also use the LUMO orbital in the presence of electron-density withdrawing substituents. The chatbot only addresses inverse demand when directly questioned, as exemplified in Student 9’s prompt.

The AI presented flaws in the graphic representation of reagents and/or products in 27 of the 44 passages classified as incorrect, showing that this was its greatest area of error (61.4%). Tasks involving the representation of compounds have been reported in the literature as challenging for these generative systems. Previous studies show this difficulty: authors performed various tests on different AIs, and in none of them was there a representation free of conceptual errors or possible alternative conceptions (Nascimento Junior; Moraes; Giroto Junior, 2024).

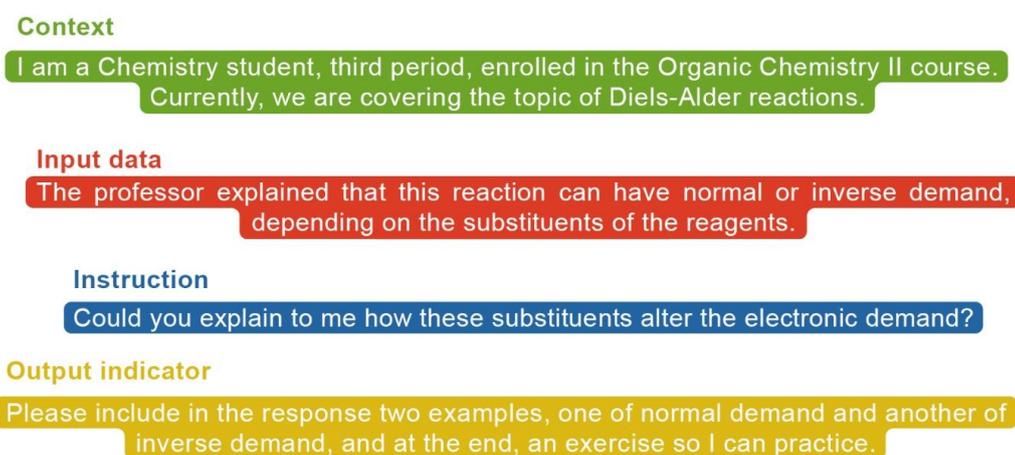
In the aforementioned work, the authors analyzed the precision of different AI models in text generation tasks on the topic of chemical bonds. The results revealed significant variations in performance. ChatGPT 3.5 (free version) presented an error rate of approximately 25%, while ChatGPT 4.0 (paid version) showed slightly better results, with 18% errors. Both tests were conducted with the free versions of the AIs, using the web and available prior knowledge as the information base (Nascimento Junior; Moraes; Giroto Junior, 2024). In contrast, IQ.QO Assistente obtained a lower error rate of only 11%. These data reinforce the importance of restricting GenAI data sources to materials of recognized academic reliability, aiming to

minimize conceptual errors and increase the trustworthiness of results. However it is important to consider that this approach represents a trade-off: one gains in precision but limits the scope of the tool's knowledge, which will be restricted to the topics and perspectives covered in the selected sources.

Studies corroborate the findings of this research by suggesting that, beyond the quality of the sources, GenAI depends on the quality of the prompts used (Tassoti, 2024; White et al., 2023). A well-developed prompt, containing the four essential components (Giray, 2023), maximizes the probability of relevant and accurate responses.

Figure 3 illustrates an example of a meticulously crafted prompt to optimize the AI's response, simulating a possible interaction from a student in this study. This prompt was structured with all the previously discussed essential components, which can be arranged in any desired order for information presentation. The prompt details the user's context as an Organic Chemistry II student, specifies the study topic (Diels-Alder reactions), delimits the task to be performed (explanation of the influence of substituents on electronic demand), and defines the response format (including examples and an exercise).

Figure 2 – Example of a prompt for chatbot interaction.



Source: the authors.

When using the prompt specified in Figure 3, which encompasses all expected elements, in the chatbot used in this research, the response not only presented the precise definition of the Diels-Alder reaction as a [4+2] cycloaddition but also detailed the differences between normal electron demand (NED) and inverse electron demand (IED). IQ.QO Assistente explained that in NED, the diene is characterized by containing electron-donating groups (EDG), while the dienophile contains electron-withdrawing groups (EWG). Highly specific terms like EDG and EWG appeared only in this response and in no other in the analysis corpus,

highlighting the importance of developing prompts for interacting with the tool. Furthermore, the AI provided specific examples, citing the reaction between 1-methoxy-butadiene and maleic anhydride as a case of NED, and that of a 1,3-butadiene with electron-withdrawing groups and styrene substituted with electron-donating groups for IED. Finally, the AI complied with the request by creating an exercise with reagents to complement the study of the topic.

To deepen learning, the next step is to interact actively with the response. This may include detailing the steps for solving the exercise or raising questions that connect different parts of the explanation. This continuous iteration benefits both the machine, by reinforcing learning, and the student, by encouraging them to reflect on their own doubts and improve their skills in organization, writing, and criticality in formulating new prompts. This attitude leads them to abandon a passive stance toward information, leading them to an active stance by modifying and participating in the structuring of responses.

This study also has implications for future research. We suggest its expansion to other chemistry topics and different disciplines to evaluate the generalization of the results. Investigations with larger samples could permit quantitative analyses of the tool's usage patterns. Finally, longitudinal studies would be valuable to track how training in prompt engineering impacts the development of student autonomy and critical thinking in the long run.

FINAL CONSIDERATIONS

This research aimed to investigate whether a chatbot specifically created for organic chemistry teaching, with a knowledge base limited to reliable websites and materials, would show a lower propensity for conceptual errors compared to general-purpose language models. Furthermore, student prompt preparation patterns were analyzed, in addition to the presentation of characteristics of an effective prompt based on the theoretical model and practical example discussed in the text.

The results show that the chatbot developed makes fewer conceptual errors than its general-purpose counterparts do. The error rate of only 11% for IQ.QO Assistente contrasts favorably with the 25% for ChatGPT 3.5 and 18% for ChatGPT 4.0 reported in a previous study. This difference highlights the importance of restricting generative AI data sources to academically validated materials, especially in highly specific knowledge domains such as chemistry.

The analysis of the 48 segments extracted from the undergraduates' prompts revealed patterns notable for their simplicity, with a predominance of input data (62.5%), indicating a direct approach in queries, a low presence of instructions (20.8%) and iterations (16.7%), and

a complete absence of context elements and output indicators. These patterns suggest that students tend to use generative AIs in a manner similar to conventional search engines, without exploring sophisticated prompt engineering techniques.

Thus, the data reinforce the need to create courses that assist students in using artificial intelligence tools within the educational context. Professors can teach with these tools, but they can also teach how to use them, so that students can extract the maximum benefit from them, being empowered by technology.

The research predictions align with the results of a test using a well-structured prompt that included four essential components—contextualization, precise data, a clear task, and a defined output format—demonstrating that careful prompt design can enhance the accuracy and relevance of responses. This finding reinforces the need to teach students how to formulate effective prompts. Educational opportunities range from fostering critical thinking and personalized learning to supporting active learning, developing computational skills, and complementing traditional instruction with learning assistants.

REFERENCES

ALPAYDIN, Ethem. *Introduction to machine learning*. 2 ed. Cambridge, Mass: MIT Press, 2010.

ARAÚJO, José Luís; SAÚDE, Isabel. Can ChatGPT enhance chemistry laboratory teaching? Using prompt engineering to enable AI in generating laboratory activities. *Journal of Chemical Education*, Washington, D.C., v. 101, n. 5, p. 1858-1864, mai. 2024. Disponível em: <https://pubs.acs.org/doi/10.1021/acs.jchemed.3c00745>. Acesso em: 16 nov. 2025.

BARDIN, Laurence. *Análise de conteúdo*. Lisboa: Edições 70, 2015.

BAUM, Zachary J.; YU, Xiang; AYALA, Philippe Y.; ZHAO, Yanan; WATKINS, Stephen P.; ZHOU Qiongqiong. Artificial intelligence in chemistry: current trends and future directions. *Journal of Chemical Information and Modeling*, v. 61, n. 7, p. 3197-3212, 26 jul. 2021. Disponível em: <https://pubs.acs.org/doi/10.1021/acs.jcim.1c00619>. Acesso em: 02 dez. 2025.

CARLEO, Giuseppe; TROYER, Matthias. Solving the quantum many-body problem with artificial neural networks. *Science*, Washington, D. C., v. 355, n. 6325, p. 602-606, fev. 2017. Disponível em: <https://www.science.org/doi/10.1126/science.aag2302>. Acesso em: 16 nov. 2025.

CORREA, Raquel Folmer; GEREMIAS, Bethania Medeiros. Determinismo tecnológico: elementos para debates em perspectiva educacional. *Revista Tecnologia e Sociedade*, Curitiba, v. 9, n. 18, dez. 2013. DOI: <http://dx.doi.org/10.3895/rts.v9n18.2633>. Disponível em: <https://periodicos.utfpr.edu.br/rts/article/view/2633>. Acesso em: 16 nov. 2025.

FUNEL, Jacques-Alexis; ABELE, Stefan. Industrial applications of the Diels-Alder reaction. *Angewandte Chemie International Edition*, Weinheim, v. 52, n. 14, p. 3822-3863, 2013. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1002/anie.201201636>. Acesso em: 16 nov. 2025.

GIRAY, Louie. Prompt engineering with ChatGPT: a guide for academic writers. *Annals of Biomedical Engineering*, v. 51, n. 12, p. 2629-2633, dez, 2023. DOI: <https://doi.org/10.1007/s10439-023-03272-4>. Disponível em: <https://link.springer.com/article/10.1007/s10439-023-03272-4>. Acesso em: 16 nov. 2025.

GRANDO, John Wesley; CLEOPHAS, Maria das Graças. Análise de aplicativos móveis de realidades digitais para o ensino de química a partir de um modelo heurístico. *Revista de Investigação Tecnológica em Educação em Ciências e Matemática*, Cuiabá, v. 1, p. 152-165, 2021. Disponível em: <https://revistas.unila.edu.br/ritecima/article/view/3195>. Acesso em: 16 nov. 2025.

HERMANSON, Greg T. The reactions of bioconjugation. In: *Bioconjugate Techniques*. [S.l.]: Elsevier, 2013. p. 229-258.

ISTE. *The International Society for Technology in Education (ISTE) Standards for students*. [S.l.], 2018. Disponível em: <https://unevoc.unesco.org/home/Digital+Competence+Frameworks/lang=en/id=17#tbar>. Acesso em: 9 out. 2025.

KNIGHT, Charles; PRYKE, Sam. Wikipedia and the University, a case study. *Teaching in Higher Education*, v. 17, n. 6, p. 649-659, dez. 2012. DOI: <https://doi.org/10.1080/13562517.2012.666734>. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/13562517.2012.666734>. Acesso em: 16 nov. 2025.

KULTHAU, Carol C. Inside the search process: information seeking from the user's perspective. *Journal of the American Society for Information Science*, v. 42, n. 5, p. 361-371, 1991. DOI: [https://doi.org/10.1002/\(SICI\)1097-4571\(199106\)42:5%3C361::AID-ASI6%3E3.O.CO;2-%23](https://doi.org/10.1002/(SICI)1097-4571(199106)42:5%3C361::AID-ASI6%3E3.O.CO;2-%23). Disponível em: <https://asistdl.onlinelibrary.wiley.com/doi/10.1002/%28SICI%291097-4571%28199106%2942%3A5%3C361%3A%3AAID-ASI6%3E3.O.CO%3B2-%23>. Acesso em: 02 dez. 2025.

KUNTZ, David; WILSON, Angela K. Machine learning, artificial intelligence, and chemistry: how smart algorithms are reshaping simulation and the laboratory. *Pure and Applied Chemistry*, v. 94, n. 8, p. 1019-1054, ago. 2022. Disponível em: https://www.degruyterbrill.com/document/doi/10.1515/pac-2022-0202/html?srsId=AfmBOooo8kgcBbDB7jaBehW6liytbvFwGGI98VFjc29V_W-w0ByaPlyb. Acesso em: 16 nov. 2025.

LEITE, Bruno Silva. *Tecnologias no ensino de química: teoria e prática na formação docente*. Curitiba: Appris, 2015.

LEITE, Bruno Silva. Aprendizagem tecnológica ativa. *Revista Internacional de Educação Superior*, Campinas, v. 4, n. 3, p. 580-609, mai. 2018. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/riesup/article/view/8652160>. Acesso em: 16 nov. 2025.

LEITE, Bruno Silva. Tecnologias digitais e metodologias ativas no ensino de química: análise das publicações por meio do corpus latente na internet. *Revista Internacional de Pesquisa em Didática das Ciências e Matemática*, Itapetininga, e020003, jul. 2020. Disponível em: <https://periodicoscientificos.itp.ifsp.edu.br/index.php/revin/article/view/18>. Acesso em: 16 nov. 2025.

LEITE, Bruno Silva. *Análise da inteligência artificial ChatGPT na proposição de planos de aulas para o ensino de química*, Vigo, v. 23, p. 473-497, 2024. Disponível em: <https://dialnet.unirioja.es/servlet/articulo?codigo=9903754>. Acesso em: 16 nov. 2025.

LOPES, Auxiliadora Cristina Correa Barata; CHAVES, Edson Valente. Animação como recurso didático no ensino da química: capacitando futuros professores. *Revista de Estudos e Pesquisas sobre Ensino Tecnológico (EDUCITEC)*, Manaus, v. 4, n. 07, jun. 2018. DOI: <https://doi.org/10.31417/educitec.v4i07.256>. Disponível em: <https://sistemascmc.ifam.edu.br/educitec/index.php/educitec/article/view/256>. Acesso em: 16 nov. 2025.

LUCKIN, Rosemary; CUKUROVA, Mutlu; KENT, Carmel; DU BOULAY, Benedict. Empowering educators to be AI-ready. *Computers and Education: Artificial Intelligence*, v. 3, p. 100076, 2022. DOI: <https://doi.org/10.1016/j.caeai.2022.100076>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2666920X22000315?via%3Dihub>. Acesso em: 02 dez. 2025.

NASCIMENTO JÚNIOR, Wilton José Diolindo; MORAIS, Carla; GIROTTO JÚNIOR, Gildo. Enhancing AI responses in chemistry: integrating text generation, image creation, and image interpretation through different levels of prompts. *Journal of Chemical Education*, Washington, D. C., v. 101, n. 9, p. 3767-3779, set. 2024. Disponível em: <https://pubs.acs.org/doi/10.1021/acs.jchemed.4c00230>. Acesso em: 16 nov. 2025.

PROFUTURO. *The Global Framework for Educational Competence in the Digital Age*, 2020. Disponível em: <https://unevoc.unesco.org/home/Digital+Competence+Frameworks/lang=en/id=6#tbar>. Acesso em: 9 nov. 2025.

RUSSEL, Stuart. *Inteligência artificial a nosso favor: como manter o controle sobre a tecnologia*. São Paulo: Companhia das Letras, 2021.

SANTOS, Marcos Eduardo Miranda; BATISTA, Wanda dos Santos; OLIVEIRA, João Victor França; JANSEN, Isabel Conceição Carvalho; SANTOS, Kelly Fernanda de Sousa; SANTO, Eliane Coelho Rodrigues dos. Ações educativas para o combate ao mosquito *Aedes aegypti* em uma escola da região metropolitana de São Luís. *Caderno Pedagógico*, Curitiba, v. 14, n. 1, jun. 2017. DOI: <https://doi.org/10.22410/issn.1983-0882.v14i1a2017.1317>. Disponível em: <https://ojs.studiespublicacoes.com.br/ojs/index.php/cadped/article/view/1372>. Acesso em: 16 nov. 2025.

SAUER, Doz J. Diels-Alder reactions II: the reaction mechanism. *Angewandte Chemie International Edition in English*, v. 6, n. 1, p. 16-33, 1967. DOI: <https://doi.org/10.1002/anie.196700161>. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/anie.196700161>. Acesso em: 16 nov. 2025.

SCOTT, JoAnna M.; BOHATY, Brenda S.; GADBURY-AMYOT, Cynthia C. Using learning management software data to compare students' actual and self-reported viewing of video lectures. *Journal of Dental Education*, v. 85, n. 10, p. 1674-1682, 2021. DOI: <https://doi.org/10.1002/jdd.12633>. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1002/jdd.12633>. Acesso em: 16 nov. 2025.

SEGLER, Marwin H. S.; PREUSS, Mike; WALLER, Mark P. Planning chemical synthesis with deep neural networks and symbolic AI. *Nature*, Londres, v. 555, n. 7698, p. 604-610, mar. 2018. Disponível em: <https://www.nature.com/articles/nature25978>. Acesso em: 16 nov. 2025.

SELLWOOD, Matthew A.; AHMED, Mohamed; SEGLER, Marwin H. S.; BROWN, Nathan. Artificial intelligence in drug discovery. *Future Medicinal Chemistry*, v. 10, n. 17, p. 2025-2028, set. 2018. DOI: <https://doi.org/10.4155/fmc-2018-0212>. Disponível em: <https://www.tandfonline.com/doi/full/10.4155/fmc-2018-0212>. Acesso em: 16 nov. 2025.

SENIOR, Andrew W.; EVANS, Richard; JUMPER, John; KIRKPATRICK, James; SIFRE, Laurent; GREEN, Tim; QIN, Chongli; ŽÍDEK, Augustin; NELSON, Alexander W. R.; BRIDGLAND, Alex; PENEDONES, Hugo; PETERSEN, Stig; SIMONYAN, Karen; CROSSAN, Steve; KOHLI, Pushmeet; JONES, David T.; SILVER, David; KAVUKCUOGLU, Koray; HASSABIS, Demis. Improved protein structure prediction using potentials from deep learning. *Nature*, Londres, v. 577, n. 7792, p. 706-710, jan. 2020. Disponível em: <https://www.nature.com/articles/s41586-019-1923-7>. Acesso em: 16 nov. 2025.

SILVA, Ketia Kellen Araújo da; BEHAR, Patrícia Alejandra. Competências digitais na educação: uma discussão acerca do conceito. *Educação em Revista*, Belo Horizonte, v. 35, n. e209940, 2019. DOI: <https://doi.org/10.1590/0102-4698209940>. Disponível em: http://educa.fcc.org.br/scielo.php?script=sci_abstract&pid=S0102-46982019000100419&lng=pt&nrm=iso. Acesso em: 16 nov. 2025.

STAKE, Robert E. *The art of case study research*. California: Sage Publications, 1995.

Case study on the use of a chatbot as a pedagogical tool for teaching Diels-Alder reactions in undergraduate education

Wilton José Diolindo do Nascimento Júnior, Mayara de Carvalho Santos,
Paulo César Muniz de Lacerda Miranda, Gildo Giroto Júnior

TASSOTI, Sebastian. Assessment of students use of generative artificial intelligence: prompting strategies and prompt engineering in chemistry education. *Journal of Chemical Education*, Washington, D. C., v. 101, n. 6, p. 2475-2482, mai. 2024. Disponível em: <https://pubs.acs.org/doi/10.1021/acs.jchemed.4c00212>. Acesso em: 16 nov. 2025.

TAUBER, Amanda L.; LEVONIS, Stephan M.; SCHWEIKER, Stephanie S. Gamified virtual laboratory experience for in-person and distance students. *Journal of Chemical Education*, Washington, D. C., v. 99, n. 3, p. 1183-1189, jan. 2022. Disponível em: <https://pubs.acs.org/doi/10.1021/acs.jchemed.1c00642>. Acesso em: 16 nov. 2025.

UNESCO. *AI competency framework for teachers*. [S.l.]: UNESCO, 2024.

WHITE, Jules; FU, Quchen; HAYS, Sam; SANDBORN, Michael; OLEA, Carlos; GILBERT, Henry; ELNASHAR, Ashraf; SPENCER-SMITH, Jesse; SCHMIDT, Douglas C. *A prompt pattern catalog to enhance prompt engineering with ChatGPT*. 2023. DOI: <https://doi.org/10.48550/arXiv.2302.11382>. Disponível em: <https://arxiv.org/abs/2302.11382>. Acesso em: 16 nov. 2025.

YEINGST, Tyus J.; HELTON Angelica M.; HAYES, Daniel J. Applications of Diels-Alder chemistry in biomaterials and drug delivery. *Macromolecular Bioscience*, Weinheim, v. 24, n. 12, p. 2400274, dez. 2024. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1002/mabi.202400274>. Acesso em: 16 nov. 2025.

Wilton José Diolindo do Nascimento Júnior

Holds a bachelors degree in Chemistry from Universidade Federal de Minas Gerais (UFMG) and a Master degree from Universidade Estadual de Campinas (UNICAMP). He is a PhD student at UNICAMP, researches digital technologies in organic chemistry education, including animations, augmented/virtual reality, and artificial intelligence. He is completing a sandwich PhD at the University of Ottawa (Canada), investigating systems thinking in chemistry education.

wiltonjdn@gmail.com

Mayara de Carvalho Santos

Holds a bachelor's degree in Chemistry from Universidade Federal Fluminense (UFF) and a PhD from Universidade Estadual de Campinas (Unicamp). She is a CNPq Junior Postdoctoral Researcher at Unicamp, where she develops a project on the systematization of knowledge in science communication. She worked at Instituto Butantan (2021–2024) and has collaborated with the Samsung/CENPEC Solve for Tomorrow Award since 2020, serving on its evaluation committee since 2021.

decarvalho.mayara@gmail.com

Paulo César Muniz de Lacerda Miranda

Associate Professor (MS-5.1) at Universidade Estadual de Campinas (UNICAMP) since 2008. He holds degrees in Industrial Chemistry, as well as a Bachelor's and a Teaching degree in Chemistry from Universidade Federal Fluminense (UFF), a Master's degree from Universidade Federal do Rio de Janeiro (UFRJ), and a PhD and *livre-docência* [Habilitation degree] from UNICAMP. His research focuses on organic synthesis, with emphasis on the total synthesis of natural products, bioactive analogues, molecular organic geochemistry, and QSAR/QSPR methods.

pmiranda@unicamp.br

Gildo Giroto Júnior

Holds a *livre-docência* [Habilitation degree] position at the Institute of Chemistry of Universidade Estadual de Campinas (UNICAMP). He earned a teaching degree from Universidade Estadual Paulista (UNESP/Araraquara) and a Master's and PhD in Chemistry Education from Universidade de São Paulo (USP). After working as a basic education teacher for nine years, he joined UNICAMP in 2017. His research addresses teacher education, digital teaching competencies, STEM/STEAM education, science communication, and applications of artificial intelligence in education.

ggirotto@unicamp.br

How to cite this document – ABNT

NASCIMENTO JÚNIOR, Wilton José Diolindo do; SANTOS, Mayara de Carvalho; MIRANDA, Paulo César Muniz de Lacerda; GIROTO JÚNIOR, Gildo. Case study on the use of a chatbot as a pedagogical tool for teaching Diels-Alder reactions in undergraduate education. *Revista Docência do Ensino Superior*, Belo Horizonte, v. 15, e058046, p. 1-23, 2025. DOI: <https://doi.org/10.35699/2237-5864.2025.58046>.