

The article does raise some interesting questions about the use of AI in teaching Geosciences. The main points are well made – there is a tension between student use of ChatGPT and its well-known issues of reliability and bias. There is also mention of the ethical considerations of AI use. How can it be ethically used in the classroom to supplement, rather than replace, student learning?

Re: Thank you for your insightful feedback on our manuscript. We appreciate your attention to the ethical considerations surrounding the use of AI in the classroom.

In response to your comments, we agree that our discussion of the ethical usage of AI in education in section 4.3 (pages 417-427) could be expanded upon to provide more depth and clarity. Specifically, we intend to enhance this section by elaborating on various ways in which AI can ethically supplement student learning.

Firstly, we will highlight the benefits of AI in facilitating personalized learning experiences, virtual tutoring, and the creation of teaching materials. By leveraging AI technologies, educators can tailor instruction to meet the unique needs and preferences of individual students, thereby enhancing their learning outcomes. Additionally, automated grading systems can streamline assessment processes, allowing teachers to allocate more time towards personalized interaction with students.

Furthermore, we will emphasize the importance of preparing educators for the integration of AI in the classroom. We will discuss the need for courses aimed at training "educators for tomorrow" who are equipped to navigate the ethical implications of AI use and guide students in utilizing AI as a supplementary tool rather than a replacement for traditional learning methods.

Additionally, we will address the role of collaborative efforts between educators and AI development institutes in promoting ethical usage of technology in the classroom. It is imperative for both parties to work together to establish standards and mechanisms for monitoring and regulating the usage of AI, ensuring that it aligns with ethical guidelines and serves the best interests of students.

Polling student's use and views of AI does add an interesting perspective as it can add to the teacher's own view. The article does produce some (potentially) interesting data gathered from the student responses that could add to the discussion on how to incorporate AI in the classroom. It is this part that needs to be developed.

Re: In response to your suggestion, we plan to incorporate the student responses into our discussion section, with a focus on elucidating how AI can be effectively integrated into the classroom environment.

The student data needs to be developed and contextualised.

Re: In response to your suggestion, we plan to incorporate the student responses into our discussion section, with a focus on elucidating how AI can be effectively integrated into the classroom environment.

The discussion on the student's view can be developed by expanding on the statistics discussed (lines 136-169). The main findings are very quickly discussed without discussing what they actually mean or how they relate to the main purpose of the article.

Re: Our original intent for the manuscript was to explore the pros and cons of ChatGPT's features and subsequently address pedagogical, ethical, and societal implications of AI integration in education, as outlined in the manuscript. However, we recognize that there is a need to better connect the statistical findings with these broader themes.

To address this concern, we will expand our discussion of the student's views by providing a more comprehensive analysis of the statistical data. We aim to explore the implications of these findings in terms of understanding the opportunities and challenges associated with AI integration in educational contexts. Moreover, we will emphasize the ethical dimensions of AI integration in the classroom and examine how the statistical findings contribute to this discussion.

The contextualisation can be developed by exploring what policies and software the institutions use in AI usage. The students are polled, the statistics are mentioned briefly, but there is no attempt to contextualise the students in their institutions. The responses have been isolated from the classroom. If AI offers "clarification outside of traditional classroom hours" (line 67), then how have the institutions used it? Is this use authorised by the institution or do the students do this independently? Given the concerns about bias, false references etc produced by ChatGPT, there needed to be more contextualisation of student use.

Re: Thank you for this insightful suggestion. Upon further inquiry, we discovered that there is currently no specific policy in place across the institutions we examined to regulate AI usage in education. However, through interactions with teachers, we learned that both teachers and students in Geosciences (GS) utilize ChatGPT. Interestingly, teachers are now taking steps to design assignments that limit the reliance on ChatGPT and encourage students to rely more on their own critical thinking skills. This observation is crucial and will be discussed in detail under the section on pedagogy.

Our article serves to raise awareness among teachers and students about the potential concerns associated with the widespread use of ChatGPT and encourages a shift towards teaching methods that prioritize the development of students' own abilities. We believe that

initiatives such as online courses aimed at preparing the 'teachers of tomorrow' will play a pivotal role in this endeavor, and we intend to highlight this point in our discussion.

Regarding line 67, our perspective is that the impact of ChatGPT on students' education extends beyond the traditional classroom hours, as students often use it in their own space as part of their educational experience.

We will incorporate these in the article.

The ethical and societal implications (lines 384-429) do raise important questions about ethical use of AI and the possible threat it poses to education, but it does not offer any suggestions or answers to the questions. How have the three institutions responded to these questions? What policies and procedures do they have in place? There is a broader discussion on AI use in GS, but the article focuses on an Indian perspective, without contextualising it in an Indian setting.

Re: It is indeed true that the ethical and societal implications of AI usage in education raise important questions, and while the three institutions we examined do not have specific policies in place to regulate AI usage, teachers are cognizant of the potential negative impacts. Teachers are taking proactive steps to address these concerns by designing classroom curriculum that limits reliance on AI and instead emphasizes the development of critical thinking skills among students.

Moreover, the concept of training "teachers of tomorrow" is crucial in preparing educators to navigate the ethical challenges associated with AI integration in education. To provide concrete examples, one professor at IIT Bombay implemented a classroom exercise in a micropaleontology course that involved students collecting samples from nearby beaches, processing them for foraminifera, and studying foraminifera morphology and ecology in the lab. This exercise effectively integrates AI for information gathering while primarily promoting critical thinking and practical application skills.

In our manuscript, we plan to include more examples of classroom exercises that limit AI usage while enhancing critical thinking and practical skills among students. By showcasing such examples, we aim to demonstrate how AI can be integrated into the classroom in an ethical and effective manner, particularly within the Indian educational context.

Section 4 (p.13) reads like the introduction to the article. It contextualises the use of AI in education more broadly, and in GS teaching more specifically. Section 4.2 again beings in pedagogical implications, but this comes after the cramped discussion on student use of AI. Moving section 4 to the beginning would provide a more logical flow from the aims of the study to the broader pedagogical implications, to the specific student use. Unpacking 3.1 next would develop the argument, as it would focus on student engagement.

Re: Our manuscript primarily focuses on examining how GS students, particularly from India, utilize ChatGPT applications and assesses their reliability through tests conducted by our team. Subsequently, we delve into pedagogical considerations and ethical impacts associated with the integration of AI in Geosciences education (final part of discussion). We agree that the responses from students need better explanations. Pedagogical and ethical implications are the final discussions of this manuscript, in which we plan to add ways of integrating AI into GS classrooms as suggested by you.

The sections on testing ChatGPT features (3.2.1) is confusing. Is this how students have used them? If so, this link needs to be made clearer. At present, the section on testing ChatGPT reads like a test of its features in absence of student engagement, perception or learning.

Re: We appreciate your attention to the clarity of the sections discussing the testing of ChatGPT features, particularly in relation to student engagement, perception, or learning.

Section 3.2 serves as the starting point for examining ChatGPT features, following the exploration of student usage in Section 3.1. In Section 3.2, we aim to evaluate the reliability of ChatGPT features through various tests, which are designed to provide students and teachers with insights into the trustworthiness of their applications.

For instance, in Phase 1 of our study, we identified that question-answering is one of the most used applications of ChatGPT by students. Building upon this observation, Section 3.2.1 (part of Phase 2) focuses on exploring methods to test this specific feature of ChatGPT. We have developed tests based on relevant areas of GS, such as solving competitive examination questions, to assess the accuracy and reliability of ChatGPT in providing responses.

We acknowledge the need to clarify the link between student usage and the testing of ChatGPT features to ensure coherence and understanding. In the revised manuscript, we will provide clearer explanations to elucidate how these sections are interconnected, thereby enhancing the overall coherence and readability of our study.

The discussion on integrity was also puzzling. GPTzero is one AI detection platform, but there is no recognised platform that is fool proof. Turnitin has AI detection software built in, but many institutions have not activated it due to the significant false positives it produces. Computer software is not the only way to detect AI. There is no recognition of other AI software or human ways to detect AI use (line 396). Line 106-111 show GPTZero to be accurate, yet line 394 states that AI generated text is undetectable?

Re: It is true that GPTZero serves as one of the AI-detection platforms, and during our testing phase, it was among the most commonly used free platforms available to us. However, we acknowledge that the field of AI is constantly evolving, and new AI detection platforms may emerge over time.

Regarding Turnitin software, we recognize that it is not freely available for usage in this region. Our intention in utilizing GPTZero was to assess its effectiveness in detecting AI-generated content for plagiarism. While GPTZero proved to be fairly reliable in most cases, it is important to note that no detection platform is foolproof, and there may be instances where AI-written content goes undetected.

We also acknowledge your point that computer software is not the only method for detecting AI texts. Teachers often possess a keen understanding of their students' writing styles and may be able to identify AI-generated content through their own observations and experience.

In line 394, where we state that AI-generated text is undetectable, we understand the need to clarify that this statement primarily applies to situations where teachers do not utilize AI-detectors. We will revise this line to better reflect the broader context and acknowledge the role of teachers in detecting AI-generated content through their own expertise.

Overall, we believe that the testing of AI detection platforms serves to benefit both students and teachers by providing insights into the reliability of AI-generated content and the effectiveness of detection methods.

Discussion on ChatGPT 4 (line 52) also raises issues of accessibility. It is subscription based, so excludes students who do not have the means to pay.

Re: True, we will mention that in text.

ChatGPT was given real time access to the internet, meaning the comment on Line 158-9 needs to be updated.

Re: True, this will be updated.

The article has two disparate strands. The first one, which is more in line with the abstract, explores how students use ChatGPT. The second, explores how academics have challenged the reliability of ChatGPT in GS teaching. They do not necessarily follow. The second strand could form an article that does not discuss student usage at all.

Re: We recognize the importance of clarifying the relationship between these two strands to avoid any perceived disconnect. The rationale behind this dual approach is twofold. Firstly,

we delve into students' actual usage of ChatGPT with respect to geoscience education. Subsequently, we subject ChatGPT to reliability tests within these applications and seek to address concerns surrounding the trustworthiness and accuracy of AI-generated content. These tests serve a dual purpose: they inform students about the reliability of their chosen applications and equip them with the knowledge to evaluate and test similar AI technologies in the future.

To better elucidate the connection between these two strands, we will revise the manuscript to provide clearer transitions and explanations, ensuring that the overarching narrative remains cohesive and easily understandable. By doing so, we aim to highlight the interdependence of these aspects and underscore their collective contribution to the broader discourse on AI integration in education.

If the article's focus is on using ChatGPT in the classroom, the argument needs to start with the broader pedagogical discussions, then explore how the three institutions have used it, to then explore the student's use and perceptions. The data collected needs to be better developed and discussed.

Re: Contrary to the perception, the primary focus of our article is on understanding how Geosciences (GS) students generally use ChatGPT and subsequently assessing the reliability of these applications through various tests. Our aim is to facilitate effective and reliable usage of AI platforms.

While we acknowledge your suggestion to start with broader pedagogical discussions before exploring institutional usage and student perceptions, our approach is driven by the need to first understand how students engage with AI technology before delving into broader pedagogical considerations. By examining students' usage patterns and perceptions, we aim to provide a foundation for discussing the pedagogical and ethical implications of AI integration in education.

We agree with your assessment that the data collected from students need to be better developed and discussed in our manuscript. In our revisions, we will dedicate more attention to analyzing and contextualizing the responses we received from students. Additionally, we will expand our discussion on effective and ethical ways of integrating AI into classrooms.