Student Perspectives on Ethical Academic Writing with ChatGPT: An Empirical Study in Higher Education

Lukas Spirgi\textsuperscript{1a}, Sabine Seufert\textsuperscript{1b}, Jan Delcker\textsuperscript{2c} and Joana Heil\textsuperscript{2d}

\textsuperscript{1}Institute for Educational Management and Technologies, University of St. Gallen, Switzerland
\textsuperscript{2}Chair for Economic and Business Education – Learning, Design & Technology, University of Mannheim, Germany

Keywords: Human-AI-Collaboration, AI Ethics, ChatGPT, Academic Writing.

Abstract: The emergence of ChatGPT has significantly reshaped the landscape of higher education, sparking concerns about its potential misuse for academic plagiarism (Cotton et al., 2023). This study examines the use of ChatGPT in academic writing among students at the University of Mannheim in Germany and St. Gallen in Switzerland, using a proposed Human-AI collaboration framework with six levels of AI-enabled text generation (Boyd-Graber et al., 2023). The survey of 699 students reveals varied ChatGPT usage across all six levels, with Level 3 (Literature Search) being slightly more utilized. Students expressed mixed opinions on ethical issues, such as the declaration of ChatGPT-generated content in academic work and the extent to which ChatGPT is allowed at their university. The results of the study highlight students’ concerns about negative effects on grades, a lack of clarity about university policies on ChatGPT, and fears that hard work will not be rewarded. Despite these issues, most students support open access to ChatGPT. The findings suggest the need for clear ethical guidelines in academia regarding AI use and highlight the potential stigmatization of AI, which could hinder technology acceptance and AI-related skills development.

1 INTRODUCTION

The emergence of ChatGPT indicates the incorporation of Artificial Intelligence (AI) in professional and educational settings. AI appears to be having an escalating impact on people’s lives due to greater interactions between humans and robots (Kim, 2022). AI in Higher Education has been used to provide personalized feedback on academic writing (Knight et al., 2020). The developments in the field of generative AI (such as ChatGPT) are accelerating the transformation in the area of knowledge work (Dell’Acqua et al., 2023). Generative AI can be defined, according to Lim et al. (2023, p. 2), "as a technology that (i) leverages deep learning models to (ii) generate human-like content (e.g., images, words) in response to (iii) complex and varied prompts (e.g., languages, instructions, questions)."

The effectiveness of this AI has led to widespread apprehensions in higher education, especially pertaining to the potential misuse by students for plagiarism through the utilization of AI-generated content in unmonitored academic tasks (Lo, 2023). Consequently, discussions in the public domain frequently emphasize the viewpoints of educators and university administrations. To date, there is a restricted amount of research on the application of AI in higher education (Garrel et al., 2023; Kim, 2022; Lim et al., 2023).

2 AI IN ACADEMIC WRITING

2.1 Framework for Human-AI-Collaboration in Academic Writing

Artificial intelligence tools for academic writing can be described as human-like robots. Initially, the term robot referred to appearance in the sense of physical presence, but it is increasingly used to describe
human-like performance (Murphy, 2019).

Human-like capabilities can be characterized by the breadth and complexity of their functionalities (Dang & Liu, 2022). In the field of AI-based text generation in academic contexts, the model proposed by Boyd-Graber et al. (2023) can serve as a reference framework. The Association for Computational Linguistics, an international research community focused on language models such as ChatGPT, has published guidelines for the ethical use of AI-based writing tools (Boyd-Graber et al., 2023). Within these guidelines, different levels can be defined that indicate increasing levels of AI performance in text generation, which affects the novelty of the content generated.

**Level 1: Assistance Purely With the Language of the Paper.** The AI assistant performs the task of paraphrasing and refining the author's initial content. The human carries out the final correction.

**Level 2: Short-Form Input Assistance.** The AI assistant serves as a writing aid for brief texts, while the human is accountable for examining the produced text.

**Level 3: Literature Search.** The AI assistant acts as a search tool, guiding the user while the human is responsible for searching, reading, and discussing references, similar to a typical search engine (Alshami et al., 2023).

**Level 4: Low-Novelty Text.** The virtual assistant is accountable for producing text that describes widely accepted concepts or presents an automated literature review summary. Subsequently, the human reviewer is responsible for ensuring precision and discerning whether to employ the generated text.

**Level 5: New Ideas.** The AI assistant generates research ideas and model results, while the human develops these further by formulating theses for discussion and defining the research problem. Moreover, humans are tasked with searching for reliable sources to support these ideas.

**Level 6 New Ideas + New Text:** The AI assistant plays a dual role in generating and executing text, whilst the human is responsible for verifying accuracy and deciding whether to adopt the generated text. In addition, the human is tasked with further development, including formulating discussion theses and defining research problems, as well as searching for well-established sources to support these ideas.

### 2.2 Ethical Guidelines for AI

Advancements in artificial intelligence present significant opportunities and substantial challenges, necessitating the ethical and responsible application of AI. In their work, Bao et al. (2022) devised an index to evaluate AI's potential advantages and risks. The ethical application of AI is evidently of paramount significance.

The ethical use of AI has led to the development of various guidelines (Floridi & Cowls, 2019). Jobin et al. (2019) conducted a meta-study examining and comparing existing ethics guidelines for AI. They created an overview of current principles and guidelines for ethical AI to assess whether there is global convergence in the principles of ethical AI and the requirements for its implementation. Their analysis revealed global alignment on five ethical principles: 1) Transparency, 2) Justice and Fairness, 3) Data Protection and Privacy, 4) No Harm and Solidarity and 5) responsible AI development.

The ethical guidelines emphasize the significance of customizing them for specific AI systems and application domains, as suggested by Jobin et al. (2019). To tackle these issues competently, adopting a particular perspective that resonates with the respective stakeholder group is essential.

### 3 THE PRESENT STUDY

The objective of this investigation is to assess the prevalence of using artificial intelligence tools for academic writing. Additionally, this study aims to scrutinize the ethical standards which are deemed crucial by students. To explore the usage and ethics of academic writing when employing ChatGPT, we pose two research questions:

1. How frequently do students use ChatGPT for the different levels according to the Human-AI-collaboration framework in academic writing?
2. How do students perceive ethical guidelines for the use of ChatGPT regarding transparency and fairness?

The ethical principles developed by Jobin et al. (2019) for the use of AI were applied in this study and specifically adapted for higher education. The perspective of students is relevant. Therefore, we focused on two aspects:

**Transparency:** The passages created with these tools are clearly marked as such. The declaration of originality at the end of a written work is adjusted and specifies the use of such tools (with the aim of acknowledging the human’s contribution to AI). These are often new ethical standards at Higher Education Institutions. Consequently, we asked about whether students are afraid of lowered grades for declaring the use of ChatGPT. The consequence of
the ethical aspect might lead to unfair evaluation from the student's perspective. Furthermore, we analyzed the awareness of the extent to which the use of ChatGPT is permitted at the university and how clear the communication is for students.

Justice, Fairness, and Equality: Free access for all learners to avoid social inequalities using AI is an issue many Higher Education Institutions are thinking about whether ChatGPT 4.0 should be offered. As a possible consequence, we wanted to know from students whether the use of ChatGPT at university means that hard work is no longer rewarded. Furthermore, the other way around, we asked how students perceive if teachers correct with ChatGPT in terms of unfair or fair evaluation.

4 METHODS

4.1 Online Survey and Sample

An online survey was chosen for the study to comprehensively explore students' experiences with AI. The survey was conducted digitally using the 'Qualtrics' platform from September to October 2023. All questions were single-choice. In total, 699 students from the University of St. Gallen and the University of Mannheim participated. The mean age of the students surveyed was 21.4 years (SD = 2.94). Students from different disciplines were surveyed at both universities.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Absolute</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female Students</td>
<td>348</td>
<td>49.8%</td>
</tr>
<tr>
<td>Male Students</td>
<td>341</td>
<td>48.8%</td>
</tr>
<tr>
<td>Diverse Students</td>
<td>10</td>
<td>1.4%</td>
</tr>
<tr>
<td>First Semester Students</td>
<td>317</td>
<td>45.4%</td>
</tr>
<tr>
<td>Bachelor Students</td>
<td>285</td>
<td>40.8%</td>
</tr>
<tr>
<td>Master Students</td>
<td>97</td>
<td>13.8%</td>
</tr>
<tr>
<td>University of St. Gallen</td>
<td>274</td>
<td>39.2%</td>
</tr>
<tr>
<td>University of Mannheim</td>
<td>425</td>
<td>60.8%</td>
</tr>
</tbody>
</table>

Table 1: Sample.

4.2 Development of Instrument

The questionnaire comprised two parts. In the first part, two specific questions were formulated for each level established in the theoretical framework (Human-AI-collaboration) to assess usage intensity. Respondents were prompted to rate their responses on a seven-point scale, ranging from 1 (never) to 7 (always). The questionnaire explained in detail the frequency of each choice. 'Never' (1) means that ChatGPT is never used this way. 'Rarely' (2) represents a use once per semester. 'Occasionally' (3) means sporadic use, i.e. several times per semester. 'Sometimes' (4) means a of use about once a month. 'Frequently' (5) means using ChatGPT several times a month in the defined way. 'Usually' (6) means once a week. 'Always' (7) means constant use (several times a week).

The study emphasized ethical considerations in the second part, specifically transparency and fairness. The choice to concentrate on these facets arises from their pivotal importance for students. Responses to ethical considerations were gauged employing a five-point Likert scale, spanning from 'strongly disagree' (1) and 'disagree' (2) to 'neutral' (3), 'agree' (4), and 'strongly agree' (5).

A total of 18 items were analyzed for this study.

5 RESULTS

5.1 Internal Consistency

The internal consistency of the constructed indices, designed to assess the frequency of usage at each level, was evaluated using Cronbach's Alpha (Cronbach, 1951). Two questions were combined at each level (1 – 6) to form an index, capturing the nuances of usage patterns among university students. Cronbach's Alpha is a measure of internal consistency, reflecting the extent to which the items within an index are correlated. The values obtained for each index are all above 0.7, indicating an acceptable to good level of internal consistency (Cronbach, 1951). This suggests that the selected items within each index reliably measure the intended construct of usage frequency among university students.

Table 2 presents the computed indices for all six levels, with the Cronbach's Alpha value. The index calculated reflects the average usage by students.

<table>
<thead>
<tr>
<th>Level</th>
<th>Index (SD)</th>
<th>Cronbach's Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.56 (1.79)</td>
<td>α = 0.75</td>
</tr>
<tr>
<td>2</td>
<td>2.58 (1.77)</td>
<td>α = 0.83</td>
</tr>
<tr>
<td>3</td>
<td>2.65 (1.79)</td>
<td>α = 0.78</td>
</tr>
<tr>
<td>4</td>
<td>2.59 (1.74)</td>
<td>α = 0.70</td>
</tr>
<tr>
<td>5</td>
<td>2.21 (1.58)</td>
<td>α = 0.81</td>
</tr>
<tr>
<td>6</td>
<td>2.45 (1.69)</td>
<td>α = 0.89</td>
</tr>
</tbody>
</table>

Table 2: Frequency of use index.
The index values have been calculated to range from 2.21 to 2.65, signifying a frequency of use between 'rarely' and 'occasionally'. The highest index value is observed at Level 3, indicating that ChatGPT is most used for literature searches.

Although the general average usage of ChatGPT across all levels is low, the data suggests that a significant number of respondents frequently use ChatGPT for academic writing.

Table 3 illustrates the frequency of usage for various scenarios, categorized from levels 1 to 6. The table displays the percentage of respondents who utilize ChatGPT in the described manner for each defined type of use. The original 7-point scale has been condensed into a 4-point to ensure clarity in the table. Respondents who selected 'never' (1) in the usage questionnaire are also represented as 'never' in Table 3. 'Rarely' (2) and 'occasionally' (3) have been merged into 'sporadically', indicating that ChatGPT is used in this manner once or several times per semester. Similarly, 'sometimes' (4) and 'frequently' (5) are combined as 'often', indicating that ChatGPT is used once to several times monthly. Those who indicated 'usually' (6) and 'always' (7) are grouped as 'very often', indicating ChatGPT usage once or several times a week. The most significant proportion of students for all types of use is 'never', but there are always at least 20% who 'usually' or 'mostly' use ChatGPT in the ways described.

The two categories, 'generating keywords for literature searches (brainstorming)' and 'using AI to define terms and explain concepts', have the highest proportion of students who say they use ChatGPT often (28%) or very often (13%). This means that these students use ChatGPT in the way described at least once a week. The categories 'using AI for concept development and design' and 'integrating AI-generated concepts seamlessly into your text' have the highest proportion of students who say they 'never' use these methods (54% and 58%).

Table 3: Frequency of use of ChatGPT (N = 699).

<table>
<thead>
<tr>
<th>Lev.</th>
<th>Type of use of ChatGPT</th>
<th>Never</th>
<th>Sporadically</th>
<th>Often</th>
<th>Very often</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Spell and grammar check</td>
<td>39%</td>
<td>27%</td>
<td>23%</td>
<td>11%</td>
</tr>
<tr>
<td></td>
<td>Translate text</td>
<td>46%</td>
<td>27%</td>
<td>19%</td>
<td>7%</td>
</tr>
<tr>
<td>2</td>
<td>Develop coherent text based on provided keywords</td>
<td>34%</td>
<td>29%</td>
<td>25%</td>
<td>12%</td>
</tr>
<tr>
<td></td>
<td>Apply the AI-corrected text directly in one's writing</td>
<td>53%</td>
<td>25%</td>
<td>16%</td>
<td>6%</td>
</tr>
<tr>
<td>3</td>
<td>Generate keywords for literature searches (brainstorming)</td>
<td>34%</td>
<td>25%</td>
<td>28%</td>
<td>13%</td>
</tr>
<tr>
<td></td>
<td>Identify pertinent literature sources with AI</td>
<td>51%</td>
<td>26%</td>
<td>18%</td>
<td>5%</td>
</tr>
<tr>
<td>4</td>
<td>Utilise AI to define terms and explain concepts</td>
<td>33%</td>
<td>26%</td>
<td>28%</td>
<td>13%</td>
</tr>
<tr>
<td></td>
<td>Incorporate AI-generated concepts seamlessly into your text</td>
<td>54%</td>
<td>26%</td>
<td>15%</td>
<td>5%</td>
</tr>
<tr>
<td>5</td>
<td>Use AI for concept development and design</td>
<td>58%</td>
<td>23%</td>
<td>16%</td>
<td>4%</td>
</tr>
<tr>
<td></td>
<td>Use AI for data analysis to generate new ideas</td>
<td>46%</td>
<td>29%</td>
<td>20%</td>
<td>5%</td>
</tr>
<tr>
<td>6</td>
<td>Use AI to draft comprehensively on given topics and goals</td>
<td>42%</td>
<td>31%</td>
<td>19%</td>
<td>7%</td>
</tr>
<tr>
<td></td>
<td>Enhance AI-generated drafts with more precise prompts</td>
<td>46%</td>
<td>27%</td>
<td>21%</td>
<td>6%</td>
</tr>
</tbody>
</table>

Table 4: Ethical Aspects Transparency and Fairness (N = 699).

<table>
<thead>
<tr>
<th>Lev.</th>
<th>Item</th>
<th>Strongly disagree</th>
<th>disagree</th>
<th>neutral</th>
<th>agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>In my opinion, ChatGPT should only be allowed if the generated passages are marked as such.</td>
<td>6%</td>
<td>19%</td>
<td>34%</td>
<td>31%</td>
<td>11%</td>
</tr>
<tr>
<td></td>
<td>I am afraid that the teachers will lower my work if I declare that I use ChatGPT.</td>
<td>4%</td>
<td>11%</td>
<td>21%</td>
<td>42%</td>
<td>22%</td>
</tr>
<tr>
<td></td>
<td>I am not currently aware of the extent to which the use of ChatGPT is permitted at university</td>
<td>4%</td>
<td>15%</td>
<td>27%</td>
<td>39%</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td>In my opinion, open access to ChatGPT for all learners is essential</td>
<td>4%</td>
<td>10%</td>
<td>29%</td>
<td>39%</td>
<td>17%</td>
</tr>
<tr>
<td></td>
<td>In my opinion, using ChatGPT at university means that hard work is no longer rewarded.</td>
<td>19%</td>
<td>34%</td>
<td>21%</td>
<td>20%</td>
<td>5%</td>
</tr>
<tr>
<td></td>
<td>I would find it unfair if teachers corrected my work using ChatGPT.</td>
<td>6%</td>
<td>19%</td>
<td>26%</td>
<td>30%</td>
<td>19%</td>
</tr>
</tbody>
</table>
5.2 Ethical Aspects

Table 4 presents the students' views on specific ethical aspects. The table shows the proportion of students who agree or disagree with the statement. The statements are divided into the criteria 'transparency' and 'fairness'.

All statements relating to the ethical aspect of transparency are approved. This means that the proportion of students who agree or strongly agree with the statement is greater than the proportion of students who disagree or strongly disagree with it. The statement 'I am afraid that the teachers will lower my work if I declare that I use ChatGPT' has the highest agreement rate (agree = 42% and strongly agree = 22%).

In the Fairnaiss category, two statements are more likely to be agreed and one statement is more likely to be disagreed. The statement 'In my opinion, the use of ChatGPT at university means that hard work is no longer rewarded' is the one that most strongly disagrees with. 53% of respondents tend to disagree, and only 5% strongly agree.

6 DISCUSSION

Students' average usage of ChatGPT is currently quite diverse; most students follow low or medium frequency. On the one hand, a subgroup of students consistently refrain from using ChatGPT at every level (between 20% to 41%). This may be attributed to the relatively restrictive regulations imposed by universities. Furthermore, students might be opting for alternative AI tools like Deepl Write for assistance in writing, which corresponds to a specific level of usage (Level 2) in our framework. On the other hand, a small segment of students (about 4-13%) consistently utilize ChatGPT across all levels, including the most advanced level, where ChatGPT functions similarly to a co-author by generating new ideas and text. This indicates a high frequency, almost to the point of being a regular pattern or habit.

On average, students use ChatGPT the most at Level 3 (Index 2.65). At this level, ChatGPT is predominantly employed for keyword searches in literature research. ChatGPT is particularly suitable for brainstorming, as the factual accuracy of the output is less critical than, for instance, when explaining theories. There are no significant differences in the frequency of ChatGPT usage across the various levels of the Human-AI Collaboration Framework.

When analyzing the data, it is noticeable that some students (15%) do not use ChatGPT in any of the usage scenarios described. This means that despite the considerable hype surrounding AI text generators, some students do not yet have confidence in this new technology and do not use it.

The following three topics focus on the ethical guideline 'transparency' and possible consequences for students following this issue:

Marking ChatGPT Passages: This survey data reflects a range of opinions on whether ChatGPT-generated passages should be marked. While a significant portion of respondents are neutral, there is a notable presence of both agreement and disagreement, suggesting a nuanced and mixed viewpoint on this issue. Further research and context may be needed to understand the reasons behind these opinions and their potential implications. Some students might believe marking is essential for transparency and accountability, helping users distinguish between human and AI-generated content. Furthermore, marking could allow accurate assessment of a student's own understanding. Marking might empower users to make informed decisions about engaging with AI-generated content.

On the other hand, opponents argue that marking restricts creative freedom and experimentation with AI tools. Concerns about grading or assessment biases against AI-generated content may influence opinions (see item: Influence on the grade). The approving position is the most substantial group. Concerns about the transparency of academic accomplishments may arise, as it may become difficult to distinguish between work produced solely by students and work assisted by AI.

In discussions about technology and ethics, neutrality can often be seen as a balanced and cautious approach (Green, 2021). Respondents in the neutral group may be taking a middle-ground position, considering both the potential benefits and concerns associated with marking AI-generated content.

Fear of Lowered Grades for Declaring the Use of ChatGPT: A substantial majority of respondents express concerns about their work being negatively affected by declaring the use of ChatGPT. This group constitutes 65% of the total respondents (agree and strongly agree) and is the highest value of all six ethical topics. Some students may worry that using ChatGPT could be viewed as a form of cheating or academic dishonesty, which could result in penalties or lower grades. Furthermore, students might be concerned that teachers or evaluators could have biases against AI-generated work, leading to unfair
assessment or grading. Worries about how disclosing ChatGPT usage might affect teachers' perceptions of students' capabilities and dedication to their work. Educational systems often place high expectations on students to excel. The fear of potentially lower grades could add to the pressure students already feel.

**Awareness of the Extent to Which the Use of ChatGPT is Permitted at University:** A significant proportion of respondents (19%) indicate that they are not aware of the extent to which the use of ChatGPT is permitted at their university. Both universities provide guidelines to the students on how ChatGPT could be used for academic writing. However, University policies on AI tool usage could be complex and challenging to understand fully. The policies have recently been introduced, giving students insufficient time to become aware. Furthermore, it could be an indicator that more than communication is needed. Students should be provided with training on responsible AI tool usage.

The ethical guideline 'fairness' is discussed with the following three aspects:

**Importance of Open Access to ChatGPT for all Students:** The data suggests a notable level of support and a significant neutral stance towards permitting ChatGPT usage at universities. While there is some opposition, it is not the dominant viewpoint. Optimistic respondents might view ChatGPT as a valuable tool in academic studies. ChatGPT can be tailored to individual needs, allowing students to receive personalized assistance and support in their coursework. Some students might appreciate ChatGPT's ability to assist in improving writing skills and generating content for assignments. Supportive respondents may believe that exposure to AI technology is essential for students to be prepared for future career opportunities as AI becomes increasingly prevalent in many professions.

**ChatGPT Impact on Rewarding Hard Work:** About a quarter of the students (25%) express agreement with the idea that ChatGPT usage may reduce the rewards for hard work. This group believes that technology may make it easier to achieve academic success without putting in as much effort. Some respondents may worry that using ChatGPT could be seen as a form of academic dishonesty or cheating, which could undermine the value of their hard work. Concerns may arise about the fairness of evaluating students when some have access to AI tools that can generate high-quality content, potentially giving them an advantage over those who do not use such tools. Students who put significant effort and time into their coursework may feel that the availability of AI-generated content devalues their hard work and dedication. There could be concerns that AI-generated work might disrupt the meritocratic nature of education, where success is traditionally based on individual effort and abilities. Some students may worry that relying on AI tools for assignments could hinder the development of critical thinking and problem-solving skills, which are essential aspects of the learning process. There may be concerns that students feel pressured to use AI tools like ChatGPT to keep up with their peers, even if they prefer not to. Additionally, some students may worry that using AI tools could conflict with the educational values of effort, learning, and personal growth.

**Unfairness if Teachers Correct With ChatGPT:** The data shows a wide range of opinions on whether using ChatGPT to correct work is considered unfair. This indicates that the topic of AI tool usage in educational assessment is complex, and opinions vary widely among respondents. The 'Agree' and 'Strongly Agree' categories collectively make up 49% of respondents, indicating that almost half of the respondents find it unfair if teachers rely on ChatGPT to correct their work.

Some students may believe that using ChatGPT for corrections could lead to generic, automated feedback lacking the personal touch and tailored guidance teachers can provide. Concerns about the accuracy of AI tools like ChatGPT in assessing and correcting complex or subjective assignments may lead to perceptions of unfairness. Students might worry about AI bias in assessments, as AI systems may not account for diverse perspectives, cultural nuances, or individual learning styles (Jobin et al., 2019). Concerns that AI-generated corrections might inadvertently introduce bias or reinforce existing biases in evaluations. Students may feel that relying on ChatGPT for corrections undermines the expertise and knowledge of teachers, potentially diminishing the value of their education. Worries that students may not learn as effectively if AI tools are used for corrections, as they might not receive explanations or insights into their mistakes. There may be concerns that students' engagement and motivation to improve their work could decrease if they receive automated corrections without the opportunity for meaningful interaction with teachers. Overall, it might reduce the teacher-student connection and the potential for mentorship and guidance.
7 CONCLUSIONS

Balancing fears (e.g., fear that using AI tools may be perceived as academic dishonesty, leading to lower grades, unfair grading by AI-based correction tools) and potential positive effects (e.g., free use of ChatGPT as a powerful tool for academic studies) is essential for responsible AI integration in education.

The overarching ethical aspect ‘transparency’ is crucial in addressing these concerns and ensuring responsible AI integration in education. Additionally, the ethical principle of ‘fairness’ is central to discussions about equal access, the impact on hard work, and the potential biases associated with AI tools. To alleviate concerns and promote responsible AI usage in education, universities should provide clear guidelines, educational resources, and open discussions to empower students to make informed decisions and navigate the evolving landscape of AI in academia.

Limited communication or education around the ethical and practical use of AI tools in education can contribute to these concerns. Students may feel that they lack guidance on how to navigate this issue responsibly.

Developing norms and guidelines for the ethical use of generative AI for academic writing currently presents a significant and complex challenge for universities. The requirement to label AI-generated content in academic work can contribute to strengthening and upholding ethical, academic, and pedagogical standards. Clear marking helps preserve academic integrity by distinguishing between students’ own work and machine-generated content (Boyd-Graber et al., 2023). It aids in adhering to ethical standards in academic work. Teachers can better assess the quality of AI-generated content and evaluate how well students use and understand these AI systems. This measure could also promote students’ awareness of responsible AI use and its impact on their learning processes.

However, the results of our studies reveal substantial arguments against labelling AI-generated passages in academic work. Labelling could stigmatize the use of AI in academic work, implying that its use is inherently less valuable or legitimate. Mandatory labelling could discourage students from exploring and using new technologies, inhibiting technology acceptance and the development of necessary AI-related competencies. Regarding human contribution, defining precisely what constitutes AI-generated content may be challenging, especially when students heavily edit and customize AI outputs. Demanding labelling could be interpreted as distrust in students' ability to handle AI independently and responsibly. From students' perspective, there is also a valid concern that open communication about using AI in their work might lead to less favourable evaluations or a loss of trust on the part of teachers.

A significant dilemma appears between establishing ethical academic integrity standards by declaring ChatGPT-generated outputs and nurturing students' AI competencies to learn how to utilize AI tools effectively. In further research efforts, we aim to delve deeper into this student perspective to explore solutions that enable AI's ethical and responsible use in higher education while simultaneously supporting the development of necessary AI competencies rather than hindering them.

REFERENCES


