INNOVATION IN LANGUAGE LEARNING International Conference

Challenges Related to Using Artificial Intelligence Tools in Tertiary Writing

Tanja Psonder[1], Dietmar Tatzl[2]

FH JOANNEUM University of Applied Sciences, Austria[1]
FH JOANNEUM University of Applied Sciences, Austria[2]

Abstract

With the rapid advent of generative artificial intelligence (GAI) tools, academic writing instructors are confronted with a necessary re-thinking of task design and its execution in tertiary EAP and ESP education. For some instructors, the currently available tools, which encompass functions such as text generation, text correction, text evaluation as well as literature research, mark the beginning of a new pedagogical era in which the educational focus needs to be redirected on the ability to think critically and to question AI-generated results. However, the use of GAI tools in higher education poses considerable challenges that need to be addressed by institutions and legal experts. These range from practical to ethical issues that students and teachers face when working with texts generated by AI. The technological progress in this area puts further pressure on teachers to adapt to an increasingly demanding educational environment. To address these current challenges, this contribution describes the results of a survey delving into the use of GAI in academic writing classes. A first discussion reflects on the acceptance of artificial intelligence tools among students and their impact on text production. The second part of the discussion critically addresses open issues related to professional referencing and an academic code of ethics with the purpose of creating an awareness of appropriate study behaviour in learners from various educational contexts and career fields.

Keywords: *ESP, EAP, higher education, artificial intelligence, issues, challenges*

1. Introduction

Writing is not only a powerful process in academic communication but also an essential technique to question and understand professional content. In general, a profound written discussion of a subject serves to activate students' cognitive processes and aims at developing their critical thinking ability within defined ethical boundaries. Hence, the writing process can be seen as the professional socialisation in which students engage with scientific processes, conventions, and cognitive procedures of their career fields.

With the rise of generative artificial intelligence (GAI) tools, writing as a central element of tertiary enculturation for students seems to be threatened. Since GAI produces different output even if the exact same prompts are entered, there is no way of factually proving that a certain text was authored by a human or compiled by GAI. As a consequence, students may feel tempted to use GAI without acknowledgment for writing texts, which would deprive them of the writing experience as a pedagogic process.

The European Network for Academic Integrity provides recommendations on the ethical use of AI to guarantee responsible research and innovation by defining default rules on the use of different AI tools for academic staff, students, and other educational stakeholders. These rules are complemented by a guidance brief on the correct and transparent acknowledgement of AI tools in papers, theses, or articles [1]. As the institutional policy at FH JOANNEUM University of Applied Sciences enables academic staff to realise the implementation of AI tools in a range from no to full integration, it was of utmost interest to learn about the GAI standard of knowledge among students in academic writing classes. Hence, a survey with nine questions was created to investigate students' degree of familiarity with the use of different GAI-tools. The questions further focused on students' respective text processing applications and GAI's perceived impact on critical thinking processes.

2. Methods

The survey was conducted in five different obligatory academic writing classes on undergraduate and graduate level over two months in the academic year of 2023/2024. The participants (N = 178) were either bachelor's (n = 103) or master's students (n = 75) enrolled in construction engineering and design programmes. The survey comprised the following questions:

1. Which of these GAI-based tools (ChatGPT, DALL-E, DeepL Write, Elicit, Grammarly, Humata, Magic Slides, QuillBot) are you familiar with?
2. Which of these GAI-based tools (ChatGPT, DALL-E, DeepL Write, Elicit, Grammarly, Humata, Magic Slides, QuillBot) have you already used?
3. Which version of ChatGPT (ChatGPT-3.5, ChatGPT-4 [probably the best-known tool]) – if you work with it – do you use?
4. Have you used plug-ins like Wolfram with ChatGPT – again, if you work with it? If yes, which ones?
5. Have you used GAI-based tools for literature review?
6. If you use artificial intelligence for literature research, how would you validate the findings?
7. There are numerous GAI-based text processing tools. For which application (text generation, text correction, paraphrasing, translation, literature research, text evaluation) would you use them?
8. Google Translate is one of the most popular tools. Which tool do you use for translations?
9. How does writing shape the critical thinking process according to your opinion?

3. Selected Results and Discussion

The first two questions of the survey investigated students' mere degree of familiarity with different, listed GAI-based software and their active use of these tools. As Table 1 shows, ChatGPT was the most familiar tool among the respondents, with a coverage of 100 percent concerning familiarity and a slightly lower number in usage. The GAI-based tools Elicit, Humata, MagicSlides and QuillBot had no recorded familiarity or use among most respondent sub-groups, with negligible exceptions (≙ 5) among the design students. DALL-E scored the lowest among bachelor's degree students in construction engineering, while design students demonstrated an average degree of familiarity with and use of this tool. In contrast, most respondent sub-groups were equally familiar with DeepL Write and Grammarly, but both software programs were used by a smaller proportion of respondents. These GAI-tools showed moderate scores across all sub-groups, but concerning usage Grammarly had consistently lower scores except for design students on bachelor's degree level.

Table 1. Overview of results for Questions 1 and 2

Q1	B BE (53)	M BE (42)	B ID (50)	M ID (33)	Q2	B BE (53)	M BE (42)	B ID (50)	M ID (33)
ChatGPT	53	42	50	33	ChatGPT	42	40	48	30
DALL-E	4	18	35	19	DALL-E	2	6	22	10
DeepL Write	12	25	38	20	DeepL Write	9	17	28	9
Grammarly	15	10	33	33	Grammarly	6	6	17	4

Figure 1, the bar graph (Question 7), compares the relative popularity in usage of GAI-based text processing tools across the six categories translation, literature research, paraphrasing, text evaluation, text correction and text generation. In general, bachelor's and master's degree students with a design background showed higher usage across most categories compared to first- and second-cycle construction engineering students. Notably, bachelor's degree students in design used GAI most for text generation, paraphrasing, text correction and translation, while first-cycle construction engineering students used it least for text evaluation. In text generation, there was a clear dominance of the field of design over construction engineering, with 36 students in the first cycle followed by 32 students in the second cycle. The mentions of literature research and paraphrasing were lower than the other categories but still reflected a clear preference among design students.

The least usage of GAI-based text processing tools was again observed in the category of text evaluation, with a higher number of users among first-cycle design students, indicating that this might

be a less prioritised application in text refinement. In comparison to the other five categories, a more consistent usage was evident in literature research. Bachelor's degree students in design again led with around 22 users, but the other three sub-groups of respondents also indicated similar usage (ranging from 12 to 18), hence, suggesting a more balanced preference for GAI-based tools for literature research.

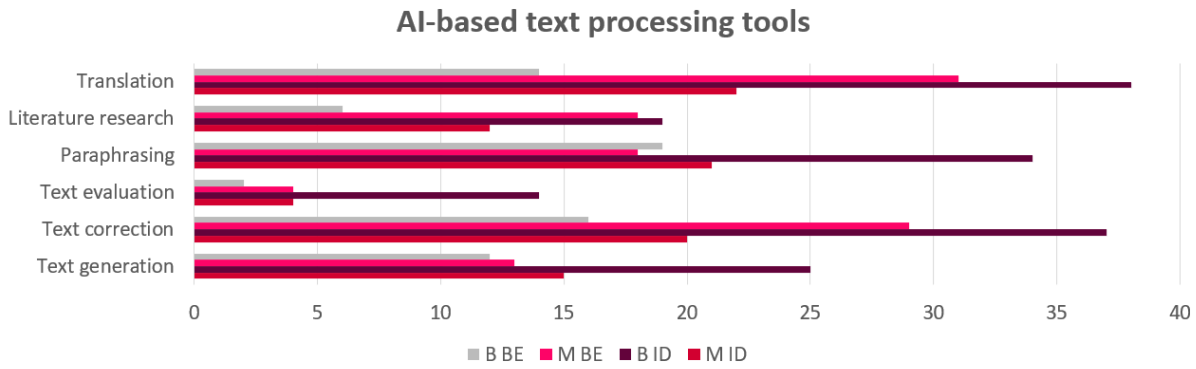

Fig. 1. Categories and extent of usage of AI-based text processing tools

The aims of writing pedagogy are to foster creativity, to develop critical thinking, and to enhance the ability to articulate ideas with clarity and precision in the individual writing and editing phases. Hence, Question 9 addresses these core issues in the preparation of academic writing classes. The respondents' views varied from little awareness to critical reflection of the impact of GAI-tools on reproductive skills. For example, some respondents stated that writing is a process that requires careful revision and self-evaluation, which enhances the writer's ability to express their ideas clearly and critically. Similarly, some students mentioned that this practice also leads to questioning the reliability of texts. In their opinion, hence, writing as a skill shapes the cognitive process by encouraging deep reflection on content, structure and approach. Other examples addressed the fact that writing is a form of self-expression that enriches the thinking process but is often impaired by a lack of knowledge of relevant structural aspects. For these students, GAI-tools could function as a virtual tutor who supports them in potential insecurities with structural and linguistic aspects [2, p. 2]. The risk that texts may take on a strong structural similarity [3, p. 4] was not reflected in their responses. A smaller group of respondents did not even find text generation or paraphrasing tools particularly useful. Some referred to a lack of sufficient content knowledge on the writers' side to verify GAI-output, while others demonstrated a certain critical awareness of the reliability of GAI-resources used for text generation.

4. Open Practical Issues

As the survey has shown, students in construction engineering and design already use GAI-tools rather frequently. However, several practical issues remain when learners resort to GAI-support in their studies, and educational institutions need to take these concerns into account. The challenges related to GAI in tertiary writing mentioned here mainly refer to ChatGPT 3.5, as this version has been made available to the global public by OpenAI. A major problem of this version is the generation of "nonsensical or inaccurate output" [2, p. 10; cf. 3, p. 4], which means that texts generated by AI need to be checked by human users for their accuracy. An even greater issue, however, is hallucination of GAI, the production of fake content that appears to be true, so that in certain contexts it may become difficult for humans to tell the difference between hallucination and facts. Another challenge identified in the literature is ambiguity and a lack of objectivity when GAI makes recommendations for users based on their prompts [4, p. 43]. A further nuisance consists in ChatGPT's "frequent use of unnecessary statements" [3, p. 4], for example sentences beginning with the phrase *As an AI language model*. Furthermore, ChatGPT 3.5 shows a tendency towards a specific essay structure [3, p. 4], resulting from the examples it has been trained on. For these reasons, some users consider it imperative to be able to detect AI-generated content [5, p. 52], but this task is probably futile, as with newer versions of the tool it will become even more complicated, not to say impossible, to make the distinction between GAI-products and texts written by humans.

5. Open Ethical Issues

Besides practical aspects of using GAI, there are even more worrying ethical issues involved. Plagiarism, for instance, is the act of "copying chatbot-generated texts without appropriate attribution" [2, p. 9], although on a large scale this academic misbehaviour rather represents ghostwriting, as full-length works could be written by GAI and published by human users under their name with intentionally misleading attribution. However, indicating GAI-use implies co-authoring, and then there is the question about GAI's share of a text. In general, GAI-texts represent academically unreliable sources, which cannot be referenced, as they are not published, not permanent, and not replicable. Furthermore, GAI itself tends to produce unreliable reference lists when requested to include one [3, p. 4; 6, n.pag.]. There is not much support from plagiarism software either, as such tools cannot detect original content written by GAI [3, p. 4]. Finally, there is the issue with unclear ownership of data, as private companies provide the tools [7, p. 147], and often their software or AI saves data on US-servers [8, p. 2].

The response to GAI by FH Joanneum University of Applied Sciences consists in four main measures. First, it published GAI guidelines [9], which leave teachers with a range from no to full implementation in their courses. Second, there is the *FH Joanneum guideline for good scientific practice and prevention of research misconduct* [10] for graduation theses, obliging students to sign under oath "that [they] have declared in the method presentation or an index all aids used (artificial intelligence assistance systems […] and indicated their usage at the corresponding text passages" [10, p. 11]. This declaration also contains this sentence: "I have been informed that my work may be checked for plagiarism and for third-party authorship of human (ghostwriting) or technical origin (artificial intelligence assistance systems)" [10, p. 11]. Third, the internal training program at FH Joanneum has been revised with a focus on GAI for teachers. Fourth, the institution has recently provided its own chatbot called A.I.D.A. 1.0., the acronym for Academic Interactive Digital Advisor [11]. It is based on ChatGPT-4, hosted internally and in compliance with data protection regulations [11]. In addition to these measures, there is open encouragement from top and middle management to use GAI in teaching and research.

6. Concerns That Remain

Despite such responses of universities to the release of GAI-tools, several concerns remain for teachers and students alike. First, there are the risks of "learning loss, especially in developing critical and creative thinking" [3, p. 4] and de-skilling of students through an over-reliance on GAI, as it is faster and more comfortable to let AI produce a text than write it yourself. Students, therefore, may miss the opportunity of learning how to write academic and professional texts. Further concerns are an increase in ghostwriting and the lowering of academic standards as a response when unattributed GAI-share in tertiary writing abounds and remains impossible to be detected. Others view GAI as a threat to individual freedom and pluralism, fearing that this may lead to a disenfranchisement of academics [12, p. 19]. It is urgent to establish conventions of GAI-use related to authorship, copyright, and referencing as well as resolve the issue of GAI as a legal entity. It will also become increasingly challenging for humans to check the accuracy of content produced by GAI and detect bias in GAI-texts, particularly as future GAI-tools will be more sophisticated than present ones.

7. Conclusions

In brief, the question of whether or not GAI can become a legal entity is linked with the question of authorship. At the moment, only humans can be authors and are thus entitled to copyright but also liable to copyright violations, whereas GAI cannot be legally held responsible. However, there are further reasons why GAI in general and ChatGPT 3.5 in particular are no authors. ChatGPT 3.5 lacks "emotional depth" and "life experiences" [3, p. 4; 13, p. 5] as well as a "deep understanding of complex concepts" [3, p. 4], which only humans are capable of. Furthermore, GAI is not creative and cannot rely on intuition and improvisation [13, p. 5]. ChatGPT 3.5 produces machine texts, which are clean, sober, and lifeless but potentially biased. In other words, it writes in prototypical style and its output contains nothing between the lines.

REFERENCES

[1] Foltynek, T., Bjelobaba, S., Glendinning, I., Reza Khan, Z., Santos, R., Pavletic, P. & Kravjar, J. (2023). ENAI Recommendations on the ethical use of artificial intelligence in education. *International Journal for Educational Integrity 19(12).* https://doi.org/10.1007/s40979-023-00133-4

[2] Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing, 57,* 100752. https://doi.org/10.1016/j.asw.2023.100752

[3] Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing, 57*, 100745. https://doi.org/10.1016/j.asw.2023.100745

[4] Schön, S., Leitner, P., Lindner, J., & Ebner, M. (2023). Learning Analytics in Hochschulen und Künstliche Intelligenz: Eine Übersicht über Einsatzmöglichkeiten, erste Erfahrungen und Entwicklungen von KI-Anwendungen zur Unterstützung des Lernens und Lehrens. In T. Schmohl, A. Watanabe, & K. Schelling (Eds.), *Künstliche Intelligenz in der Hochschulbildung* (pp. 27–49). Bielefeld: Transcript. https://www.transcript-verlag.de/media/pdf/c9/16/59/oa9783839457696.pdf

[5] Uzun, L. (2023). ChatGPT and academic integrity concerns: Detecting artificial intelligence generated content. *Language Education and Technology, 3*(1), 45–54.

[6] Thorp, H. H. (2023). ChatGPT is fun, but not an author. *Science, 379*(6630), 313-313. DOI: 10.1126/science.adg7879

[7] Ichaporia, N. (2024). What is the teacher's role in increasingly automated online education? In D. Bullock (Ed.), *IATEFL 2023 Harrogate Conference Selections* (pp. 146–147). Proc. of 56th International Conference, 18–21 April 2023. Faversham, UK: IATEFL.

[8] Schreibzentrum. (2023). Nutzung von KI-Schreibtools durch Studierende. Goethe Universität, Frankfurt am Main. https://www.starkerstart.uni-frankfurt.de/133460941/6-030_KI-Tools_pdf.pdf?

[9] FH Joanneum University of Applied Sciences. (2024, May 2). *KI-Leitfaden für die Lehre*. Retrieved from https://www.fh-joanneum.at/Hochschule/hochschuldidaktik-und-ki/kuenstliche-intelligenz-in-der-lehre/ki-leitfaden-fuer-die-lehre/

[10] FH Joanneum University of Applied Sciences. (2023, Nov. 7). *FH Joanneum guideline for good scientific practice and prevention of research misconduct* (Version 1.2). Graz, Austria.

[11] Philipps, B. (2024, 31 July). *NEU an der FH: A.I.D.A. 1.0 – der erster* [*sic*] *FH-Chatbot!* [Internal e-mail].

[12] Limburg, A., Bohle-Jurak, U., Buck, I., Grieshammer, E., Gröpler, J., Knorr, D., Mundorf, M., Schindler, K., & Wilder, N. (2023). Zehn Thesen zur Zukunft des Schreibens in der Wissenschaft. Diskussionspapier No. 23. Berlin: Hochschulforum Digitalisierung.

[13] Friedrich, J.-D., Tobor, J., & Wan, M. (2024). 9 Mythen über generative KI in der Hochschulbildung. Diskussionspapier No. 29. Berlin: Hochschulforum Digitalisierung.