

# Generative AI and its Impact on Activities and Assessment in Higher Education: Some Recommendations From Master's Students

**Peter Mozelius**

Mid Sweden University, Sundsvall, Sweden

[Peter.Mozelius@miun.se](mailto:Peter.Mozelius@miun.se)

**Abstract:** The rapid development of generative AI (GenAI) raises new questions in higher education such as: What should be the university policy regarding GenAI? How ought courses be redesigned for fair and resilient assessment? What the added pedagogical and didactical values when involving GenAI in teaching and learning activities? Different universities have rapidly created and presented contradictory standpoints and draft policies, and teachers show different opinions regarding the pros and cons of GenAI. This study has been carried out with a student perspective, where 16 students have been examining their own Master's programme on sustainable information provision. Students have assessed the assessment in their previous courses in the Master's programme. The aim of the study is to investigate how sustainable course activities and assignment are, and to explore how GenAI tools might support and facilitate teaching and learning activities. Moreover, the students were given the task to test detection software on GenAI generated solutions to assignments in chosen Master's courses. Students conducted these tasks as a part of a 7.5 ECTS project course in the same Master's programme as the investigated courses are a part of. For inspiration and for background information on artificial intelligence to the project work students participated in the first Symposium on AI Opportunities and Challenges (SAIOC) in December 2023. Data have been gathered from reports of 3 group projects where 16 students have investigated 5 freely chosen courses in the programme in each group work. Beside from testing GenAI tools in existing activities and assignments students also interviewed the subject matter experts that are responsible for the chosen courses. Results were firstly analysed and presented in group reports, combined with 16 individual reflection essays. Regarding the individual essays students were instructed to bring up ethical perspectives on GenAI in higher education, and also to present and discuss suggestions for how the current course design and assignments better could be redesigned for improved sustainability and fairness. Finally, all the group reports and the individual reflection essays were thematically analysed by the author, who also is the subject matter expert and main teacher for the project course. Findings show that many of the existing assignments in the Master's programme could be partly solved with different GenAI tools. The AI generated solutions showed different levels of quality and correctness for different types of activities and assignments. An ethical concern that many student essays brought up was the relatively poor quality of the tested detection software. A question in one of the essays was if teachers should use detection software with an accuracy rate just above 50% to evaluate student submissions. The recommendations from both the students and the author are to provide clear instructions about when GenAI is allowed and not in course activities, and to redesign the course structure for continuous assessment. With or without GenAI tools, a continuous assessment where the whole study path through a course is assessed, and not only isolated submissions, would strengthen fairness and sustainability. Finally, several students suggest oral examinations as a complement to the existing assessment methods, even if their findings showed that GenAI tools can be used to prepare oral presentations.

**Keywords:** Artificial intelligence, Generative AI, GenAI, Higher education, Sustainable education

---

## 1. Introduction

During the last years there have been a rapid development of generative AI (GenAI), with new GenAI tools that raises new questions in higher education. Research studies have reported on the need of new assessment methods to secure that students submissions are original work (Luo, 2024). The ongoing third spring of AI might be a hype followed by a new AI winter, but the impact on teaching and learning in higher education will remain (Pons, 2023; Whitham et al., 2023). Some universities have rapidly created draft policies contradicting each other, while other universities have kept a passive silence. As discussed in Luo (2024), to strict and banning policies may result in that students become hesitant to use GenAI even for legitimate learning purposes.

Higher education and technological development have a symbiotic relationship that goes way back. On one hand, technological development could be seen as a result of research efforts that has led to technological applications in society. One such field is digital technologies for technology enhanced learning which currently has a strong focus on GenAI technologies. On the other hand, these technologies are also applied in experiments and projects that aim to further develop higher education operations and to investigate how AI impacts learning and teaching in higher education. There is a strong need today for higher education initiatives on how to involve GenAI in course design and students' learning activities. (Mozelius et al., 2024)

Several research studies have investigated the teacher perspective of GenAI in educational contexts (Duc 2023; Stewart, 2023), and pointed out the need for teacher professional development on the use of GenAI in education (Mishra, Warr & Islam, 2023; Chiu, 2024). As highlighted by Chan and Hu (2023) it is important to bring in the student voices in the discussion as well. In this study 16 students have examined their own Master's programme

on sustainable information provision, but that the teacher voices have been involved as well when students have conducted semi-structured interviews with the teachers have designed the ingoing courses. The aim of the study is to investigate the sustainability of course activities and assignment in the Master The aim of the study is to investigate how sustainable course activities and assignment are, and to explore how GenAI tools might support and facilitate teaching and learning activities in the Master's programme.

## 2. Method

The overall research approach could be described as educational action research as outlined by Norton (2009). A branch of action research in the British tradition "*that links research to improvement of practice and is education orientated*" Norton (2009, p. 71). Aligned to the idea that research results should have a positive impact on the redesign of the future versions of the investigated courses. As posit by Arnold and Norton (2021), an educational action research that should strive to bring change, innovation and educational development. The educational action research has been combined with the idea of participatory where students should not only be researched objects, but also partners in the research process (Baum, MacDougall and Smith, 2006). Moreover, the study had an overall qualitative approach were interviews and activity evaluations in the student groups have been summarised in group reports and individual assignments.

### 2.1 Research Context

This study was carried out with a student perspective, where students have examined the Master's programme on sustainable information provision that they are a part of. The overall purpose of the Master's programme is to provide advanced training in sustainable information services with special focus on management and design. For this study, the main idea has been that students should investigate the assessment and teaching and learning activities in their previous courses in the Master's programme. Students carried out this work as a part of a 7.5 ECTS project course where they also should plan and conduct a project in smaller groups with 4-6 participants. As inspiration and for background information on GenAI for the project work students participated in the first Symposium on AI Opportunities and Challenges, SAIOC (Jaldemark, Mozelius & Humble, 2023).

Besides from the students' findings, this study also involves the teacher perspective since all groups have interviewed at least one subject matter expert or course designer in one of the chosen Master's courses. Students were also given the task to test different detection software to compare GenAI generated solutions to human generated solutions to assignments in the chosen Master's courses. All interviews were conducted in Swedish and quotes have been translated from Swedish to English by the author.

### 2.2 Data Collection

Data have been collected in group projects where 16 students, divided into 3 groups, investigated 5 courses in each group work. All groups carried out teacher interviews, document studies and various GenAI tool testing. As a part of the tool testing student generated solutions to assignments were compared to AI generated solutions with the use of detection software. Students could freely choose detector tools from the listing in Figure 1 below. This is the actual listing in December 2023 during the course span, and that the current listing in July 2024 involves 10 detector tools with a slightly higher accuracy rate than the percentages in Figure 1.

Winston AI	84%	✓	4.2/5
Originality.AI	76%	✓	3.8/5
GLTR	72%	✓	3.6/5
Sapling	68%	✓	3.4/5
Content at Scale	66%	✓	3.3/5
Copyleaks	66%	✓	3.3/5
Crossplag	58%	✓	2.9/5

Figure 1: The 7 Greatest AI Content Detector Tools (Forbes, 2023, December)

In the individual essays students were also asked to bring up ethical perspectives on GenAI in higher education. Moreover, students should present suggestions on how the current course design and assessment better should

be redesigned for sustainability and fairness. Finally, the author has used notes taken at discussion webinars in the described project course.

### 2.3 Data Analysis

Firstly, results were analysed by students individually, and in groups, before the presentations at course webinars, in written group reports, and in individual reflection essays. The different groups have used different analysis methods, but the overall approach could be described as thematic analyses.

Secondly, the group reports and the individual essays were analysed by the author in an inductive thematic analysis. A process that followed the six-step process outlined by Braun and Clarke (2006) to identify themes, i.e. patterns in the data that can contribute to meeting the aim of the study. The six-steps that guided the analysis were: 1) To become familiar with the data, which started early during the discussions at the course webinars, 2) To generate initial codes, which also started out early with the student analyses, 3) To search for themes, which has been done both by the students and the author, 4) To review themes, which was done in a first iteration by the students, and then in a second iteration when student submissions were aggregated by the author, 5) Defining the final themes, was carried out by the author in an aggregation and reconstruction of the student themes, and finally 6) Writing up and creating a narration to present the findings was done by the author, but based on the student submissions.

### 2.4 Ethical Considerations

All involved stakeholders have been kept as anonymous as possible, and the ambition has been to not criticise individuals or specific courses. The overall principal has been to leave out all names and more specific details to achieve the aim of investigating how sustainable course activities and assignment are, and to explore how GenAI tools might support and facilitate teaching and learning activities.

## 3. Findings and Discussions

The findings from the analysis have been grouped into the categories of 'Activity types', 'Assignment complexity', 'Detection software', 'Ethical issues', and finally 'Recommendations for redesign'. The found themes are reported and discussed one by one here below.

### 3.1 Activity Types

In a programme with a wide variety of courses and course activities the findings were that most of the activities could be solved, or partly solved with the assistance of GenAI tools. One of the groups that used a wide variety of GenAI tools brought up that also the oral activities such as seminars and discussion activities could be prepared with AI support. The students must of course be the ones that presents and discusses, but that arguments and problematisation could be outlined in chatbot conversations. A conclusion in all group reports is that oral activities and activities where students are supposed to list basic facts or to explain concepts the GenAI support works well. This could be aligned to Bloom's taxonomy that is depicted in Figure 2, and that the lower levels of remembering and understanding are handled without errors by most of the tested tools. Furthermore, to draw connections on the 'Analyze level' of the taxonomy often show relevant output.

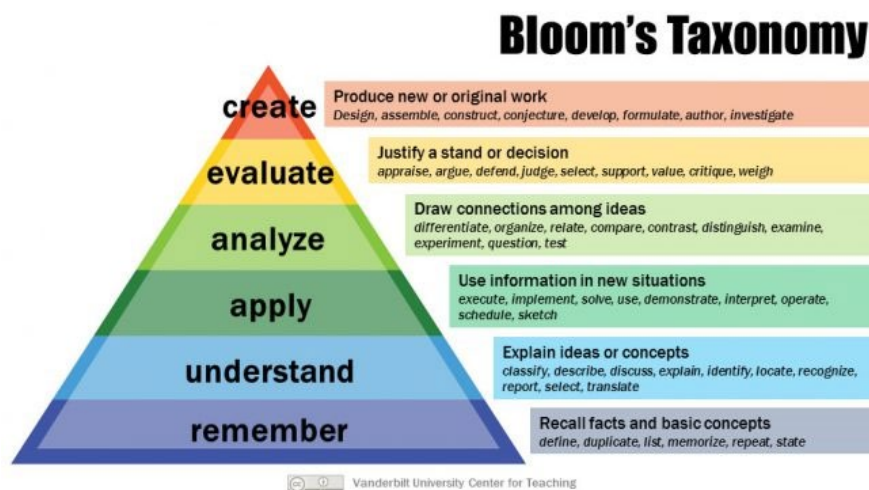


Figure 2: Bloom's taxonomy (Armstrong, 2010)

For other levels such as apply and create, one group report brings up that "The completely AI generated text often show poor quality, with minor errors, superficial analysis and incorrect or missing references". This aligns with the answers from the interviewed subject matter experts, where one of the interviewees brought up that "Two cases where students have used AI tools to solve assignments were mainly detected by the complete lack of references".

In the interview same the informant also pointed out that the use of AI tools for language correction is something to recommend, and something to place at the same level as using the spelling support in a word processing software. The group that had the largest toolbox also tested the free version of 'Grammarly', but their evaluation was that "The language support was not of higher quality than in a word processing software such as Microsoft Word". Another group, where several members have other mother tongues than Swedish, the group report highlighted the benefits of using different software to improve both the written Swedish and the written English in course submissions. Finally, two of the groups found DeepL to be a useful tool for translations between different languages. However, with a slight warning for "losing your own voice" in translations of your own writings.

Looking at the broader pedagogical implications of GenAI most of the tested tools could better be used in brainstorming activities and preparation. The very nature of GenAI is to collect and present what already is known with the possibility to select and aggregate from huge amounts of data. However, since there also are several recent research reporting on the darker side of GenAI with issues such as bias and deepfakes (Cooper and Tang; Ferrara, 2024), it seems important, as suggested by Van Dis et al. (2023), to hold on to human verification by teachers and subject matter experts. Looking at this from the student perspective the issues of authorial identity of solutions and accountability are essential (Duah and McGivern, 2024). On the brighter side of these new conditions is that, it could act as a catalyst for the revision of traditional assignments that only involves the lower levels of Bloom's taxonomy (Figure 2).

### **3.2 Assignment Complexity**

As mentioned in the previous section more complex assignments involving the higher levels of Bloom's taxonomy were more difficult to solve correctly. An example that is investigated and discussed in one of the group reports is a two-part assignment on document handling plans. In part one of the assignment the students should describe and discuss what document handling plans are, what they could help out with, what they could be used for, how useful they are, and what pre-knowledge that is required for using document handling plans. Part two of the assignment is about a specific area, where specific document handling plans are given, and where students should discuss and reason around the added value of document handling plans. Two web resources that contain content to answer these assignments are given, and that they both should be referenced to in the solutions.

Students asked a GenAI tool to generate the 2000-word report that was required in the assignment, based on the document handling plans for two Swedish authorities (The Ministry of Justice and The Taxing Authority). The generated output consisted of 485 words where the prompted questions only were partly answered, and without any references at all. In another prompt, the chatbot was asked to elaborate, and to add the required references. Some more information was added, but not much and this type of assignment structure must be seen as 'AI resilient'. When Bard (now Gemini), was prompted another time to elaborate and to add references, the software completely crashed. In the detection software Smodin, the 458 words were judged as AI generated with a probability of 71.8 %. Moreover, the solution of part 2 was also incomplete and the 537 generated words were evaluated as " Human and AI generated" with a probability of being 60.8% AI generated. All group reports criticize the output from the tested detection software which bridges to the theme of the next section.

### **3.3 Detection Software**

Due to the limited course budget students had to use detection software without any licence fees. Some of them performed relatively poor and one group report states that "According to our tests Copyleaks had a higher degree of correctness than GPTZero". Where Copyleaks at that time had an estimated accuracy rate of 66% (Figure 1). The same group report also indicates that human generated texts with spelling errors and grammatical flaws more easily pas the AI detection. Not a surprise, and what was more surprising was a teacher answer to an interview question about the use of AI detector software: "I don't really understand the question if it considers student texts in general, or specifically AI generated texts. Student texts are always run in plagiarism checking software. If AI generated texts get detected more frequently, I don't really know". What this teacher, and many other teachers hope for is that the AI detection should be a built-in part of the plagiarism

control, which now is the case in systems such as Turnitin. However, even if Turnitin has a higher accuracy rate than most of the software listed Figure 1, the accuracy rate was between 75-85% when tested. Where Turnitin stands out is regarding false positives where the test result was 0.00% (Weber-Wulff, 2023)

Another group made tests with the detection software tools Smodin and Copyleaks, where they found both false positives and false negatives. "The Smodin tool showed variations in the evaluations between different executions of the same text, and evaluated large part of an AI generated text as human generated, and the same for Copyleaks". The same group wrote in their report, "this aligns with the conclusions in Farrelly and Baker (2023), that is today not possible to trust the detection tools. There are too low accuracy rates for using the evaluations as proof in the judgement of student submissions". The same recommendation is also given by Tate et al. (2023), where this study points out that the current detection tools are not trustworthy when it comes to decisions about if a text is generated by a human or by an AI tool. The third group made tests with ZeroGPT and found accuracy rates between 46.6 - 67.7% for AI generated texts, to compare with the 99% that can be found on the providers web page. In the middle of the course span the university had the first case of suspected AI generated student submissions, where one student was convicted in the university discipline board. In media it was stated that the conviction was based on an evaluation with an AI detection software. Two of the group reports posit their doubts about using detection software in this kind of cases.

### **3.4 Ethical Issues**

In the group reports the most frequently mentioned issue is about detection software, and the lack of trust in detection software. Several students use the Swedish word "rättssäkerhet", a term that DeepL translates to legal certainty or legal security, but a term that in Swedish also involves transparency and citizens trust of a legal process. In the individual essays several students bring up the question of that "GenAI could be misused and that students then don't acquire the skills and knowledge that are stated in the learning objectives". However, in the same essay where the quote above was taken from it was also questioned "Is it that wrong to use GenAI in studies?" and "Isn't AI also a knowledge in itself?". Another essay brings up the ethical issues of Deepfakes, and the importance of keeping higher education free from Deepfakes. This essay also highlights how GenAI can be a double-edged sword in programming education, generating code that the students cannot understand. Moreover, this essay also points out the problems with: "GenAI substituting teachers", and "the ownership of AI generated content".

One essay that suggests that assignments that involves more of practical real-world problems would be harder to GenAI solve, also compares GenAI to other technologies. Like Einstein said about nuclear energy, it can be used both for good and for bad. Another essay that brings up the issue of "chatbot hallucinations" also discusses the issue of biased training data, and that chatbots only generate content based on what their neural networks have been fed with. Other essays discuss the issue of integrity, and how much, and which personal information that is stored on different Internet servers. One student wrote the essay partly around a conversation with a chatbot, where the chatbot was asked to bring up ethical issues. In the reply there were three issues that also could be found in several other essays: bias, lack of transparency and the earlier mentioned integrity. This also align with the findings in Zhai (2022), a study where ChatGPT was prompted to write a paper about challenges with GenAI in educational contexts.

### **3.5 Recommendations for Redesign**

One of the groups created a list of recommendations for subject matter experts and instructional designers that is worth considering:

- Encourage students to use GenAI tools to explain complex concept, but that their submissions must be complemented with references to other verifying sources
- Check all the references in essays and other textual submissions
- The textual assignments should be complemented with oral examination
- Test all assignments on GenAI tools before giving them to students to evaluate how easily they can be solved with AI support
- Create assignments where the analysis should be filtered by the course literature and involve own opinions
- Continue to use traditional methods such as to discuss the use of AI with students, instead of only using detection software
- Require students to declare if GenAI tools have been used in a solution, and which tool and in which parts of the solution

- Involve assignments where students should find innovative solutions to real-world problems in areas where students are supposed to work in the future
- Encourage students to interdisciplinary research on AI, technology and education, and that researchers in higher education should have a dialogue with companies and organisations that develop GenAI tools

The last point in the list has its origin in the group's report from the first Symposium on AI Opportunities and Challenges that was a mandatory part of the course (Jaldemark, Mozelius & Humble, 2023). Another group has similar suggestions and sees oral examinations as a kind of active learning events that could reinforce the learning outcomes. Moreover, they do not want to ban GenAI tools, but they want "clear guidelines and a university policy that states when and where AI tools are permitted". The third group also has similar recommendations about new examination forms, new ethical guidelines, and that the university should evaluate the quality of detection tools.

#### **4. Concluding Discussion and Future Research**

Data were collected from students examining their own Master's program which could introduce bias, as students may have preconceived opinions about the program and its assessments. This could affect their evaluations and the study's outcomes as in all educational action research studies where data are collected by researchers with dual roles. Moreover, the relatively small sample size limits the generalizability of the findings to a broader population of students or other educational programs. This study should be seen as a first explorative step in the educational action research process that most departments need to carry out to redesign courses for a higher GenAI resilience.

Some course related ethical aspects were brought up from the student perspective in the result section, but there are also broader ethical implications and other stakeholder involved in the use of GenAI in higher education. As brought up in the study on GenAI research ethics by Van Dis et al. (2023), it could be important to widen the debate to involve more stakeholders, hold on to human verification, and to develop rules for accountability. Other ethical aspects to tackle include bias in the AI models, stewardship and transparency. In a European context it would also be highly important to provide a data privacy that aligns with the GDPR law (Leboukh, Aduku, and Ali, 2023).

The conclusion resembles the one in Mozelius (2024), a majority of the current assignments in the investigated Master's programme could be partly solved with different GenAI tools. At the same time, the solutions to the more complex assignments showed to be incomplete as well as incorrect. A new finding is that students and teachers have a consensus on that GenAI should not be banned, but that there is a need for a general policy and specific guidelines for all courses in the programme. Moreover, the tested detection software showed poor results leading to the ethical concern of using this kind of software in higher education disciplinary boards. The recommendation is to revise the current assignment design in courses to more varied process that involves more of real-world problems and creativity. The revised assignments should also provide guidelines for which GenAI tools are allowed and when, and on the other hand that students should declare if GenAI tools have been used in a solution, and which tools and in which parts of the solution.

As future research it would be interesting to the presented recommendations for redesign in the development of a new course, and to try to make the course GenAI resilient from start. This should be a course on AI in education where GenAI tools will be used to a relatively high degree. At the time of writing the author has started the development of a new course on AIED where some of the questions that were unanswered in this study will be asked again. Finally, the broader pedagogical implications of involving GenAI in higher education need further exploration.

#### **References**

- Armstrong, P. (2010). Bloom's Taxonomy. Vanderbilt University Center for Teaching. Retrieved (03/07/2024) from <https://cft.vanderbilt.edu/guides-sub-pages/blooms-taxonomy/>
- Arnold, L. and Norton, L., 2021. Problematising pedagogical action research in formal teaching courses and academic development: A collaborative autoethnography. *Educational Action Research*, 29(2), pp.328-345.
- Baum, F., MacDougall, C. and Smith, D., 2006. Participatory action research. *Journal of epidemiology and community health*, 60(10), p.854.
- Chan, C.K.Y. and Hu, W., 2023. Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20(1), p.43.

- Chiu, T.K., 2024. Future research recommendations for transforming higher education with generative AI. *Computers and Education: Artificial Intelligence*, 6, p.100197.
- Cooper, G. and Tang, K.S., 2024. Pixels and Pedagogy: Examining Science Education Imagery by Generative Artificial Intelligence. *Journal of Science Education and Technology*, pp.1-13.
- Duah, J.E. and McGivern, P., 2024. How generative artificial intelligence has blurred notions of authorial identity and academic norms in higher education, necessitating clear university usage policies. *The International Journal of Information and Learning Technology*, 41(2), pp.180-193
- Duc, A.N., 2023. Exploring the Potential Impact of Generative AI in Education—A reflective analysis from nine IT courses. In *Norsk IKT-konferanse for forskning og utdanning* (No. 2).
- Farrelly, T. and Baker, N., 2023. Generative artificial intelligence: Implications and considerations for higher education practice. *Education Sciences*, 13(11), p.1109
- Ferrara, E., 2024. GenAI against humanity: Nefarious applications of generative artificial intelligence and large language models. *Journal of Computational Social Science*, pp.1-21.
- Forbes 2023. "The 10 Greatest AI Content Detector Tools" Retrieved (15/12/2023) from <https://www.forbes.com/sites/technology/article/best-ai-content-detector-tools>
- Jaldemark, J., Mozelius, P. and Humble, N., 2023. Symposium on AI Opportunities and Challenges: Education will never be the same again. ACI Academic Conferences International.
- Leboukh, F., Aduku, E.B. and Ali, O., 2023. Balancing ChatGPT and data protection in Germany: challenges and opportunities for policy makers. *Journal of Politics and Ethics in New Technologies and AI*, 2(1), pp.e35166-e35166.
- Luo, J., 2024. A critical review of GenAI policies in higher education assessment: A call to reconsider the "originality" of students' work. *Assessment & Evaluation in Higher Education*, pp.1-14.
- Mishra, P., Warr, M. and Islam, R., 2023. TPACK in the age of ChatGPT and Generative AI. *Journal of Digital Learning in Teacher Education*, 39(4), pp.235-251.
- Mozelius, P., 2024. Generative AI and Assessment in Higher Education: The Student Perspective. In *SAIOC 2* (Vol. 2, p. 23). ACI Academic Conferences International.
- Mozelius, P., Cleveland-Innes, M., Håkansson Lindqvist, M., Jaldemark, J., and Sundgren, M., 2024. Educating the educators on artificial intelligence in higher education. In *Proceedings of the 4th International Conference on AI Research, ICAIR 2024*. Lisbon, Portugal. Academic Conferences International.
- Norton, L., 2009. *Action research in teaching and learning: A practical guide to conducting pedagogical research in universities*. Routledge.
- Pons, A., 2023. Generative AI in the classroom: From hype to reality.
- Stewart, B., 2023. GenAI on GenAI: Two Prompts for a Position Paper on What Educators Need to Know. *Irish Journal of Technology Enhanced Learning*, 7(2), pp.32-41
- Tate, T., Doroudi, S., Ritchie, D., Xu, Y. and Warschauer, M., 2023. Educational research and AI-generated writing: Confronting the coming tsunami. *EdArXiv. January*, 10.
- Van Dis, E.A., Bollen, J., Zuidema, W., Van Rooij, R. and Bockting, C.L., 2023. ChatGPT: five priorities for research. *Nature*, 614(7947), pp.224-226.
- Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., Šigut, P. and Waddington, L., 2023. Testing of detection tools for AI-generated text. *International Journal for Educational Integrity*, 19(1), p.26.
- Whitham, R., Stockton, G., Richards, D., Lindley, J., Jacobs, N. and Coulton, P., 2023. Implications of Generative AI on Learning and Assessment in Higher Education and Design Research Practice.
- Zhai, X., 2022. ChatGPT user experience: Implications for education. *Available at SSRN 4312418*.