Themes in the Declared Use of Generative Artificial Intelligence in Assessment

Joseph Maguire
School of Computing Science
University of Glasgow
Glasgow, United Kingdom
joseph.maguire@glasgow.ac.uk

Rosanne English
Department of Computer and Information Sciences
Strathclyde university
Glasgow, United Kingdom
rosanne.english@strath.ac.uk

Qi Cao
School of Computing Science
University of Glasgow
Glasgow, United Kingdom
Qi.Cao@glasgow.ac.uk

Chee Kiat Seow
School of Computing Science
University of Glasgow
Glasgow, United Kingdom
CheeKiat.Seow@glasgow.ac.uk

Abstract

Generative Artificial Intelligence use by students completing assessments has been an area of concern for academics. Some educators believe such use will undermine all assessment, while others think it has the potential to revolutionise assessments. This has resulted in some institutions and educators adopting various approaches to control the use of Generative Artificial Intelligence However, much of this is taking place without fully appreciating how students are already making use of such tools. In this paper a practice where an existing assessment is presented with the addition that students are not prevented from using Generative Artificial Intelligence but must declare and explain such use. These declarations and explanations are considered to better understand how students approached the assessment and how it could be refined in future.

CCS Concepts

• **Social and professional topics** → **Computing education**; • **Security and privacy** → *Human and societal aspects of security and privacy*; • **Human-centered computing** → Accessibility design and evaluation methods.

Keywords

cyber security, research-led teaching, artificial intelligence

ACM Reference Format:
Joseph Maguire, Rosanne English, Qi Cao, and Chee Kiat Seow. 2025. Themes in the Declared Use of Generative Artificial Intelligence in Assessment. In *Computing Education Practice (CEP '25), January 07, 2025, Durham, United Kingdom.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3702212.3702217

CEP '25, January 07, 2025, Durham, United Kingdom

ACM ISBN 979-8-4007-1172-5/25/01
https://doi.org/10.1145/3702212.3702217

1 Introduction

The prevalence of Generative Artificial Intelligence tools and services has grown in recent years. This has had an impact in higher education, where educators and institutions have scrambled to adjust teaching and assessment. Adjustments include the regulation of use through the creation of policies [14]. In Computing Science (CS), Generative Artificial Intelligence tools can readily generate working source code, particularly code which is less complex [12]. There is evidence students feel the use of such tools should be limited [9]. However there is limited insight into how students currently make use of these tools. This paper presents the adaptation of an assessment which requires students to declare the use of, and an explanation of how such tools were used. The declarations are analysed to establish themes which reflect use of Generative Artificial Intelligence in assessment in a single context.

The expectation is that by understanding how learners are making use of these tools in existing CS assessments that are not focused on programming, educators will be better able to make informed decisions around re-design of practice and assessment [13]. The contributions of the present paper are:

- Reported experience of requesting students to declare and explain use of Generative Artificial Intelligence in an existing CS assessment.
- Reported experience of themes that emerged from consideration of explanations in the context of a single course in a single institution.

The aim is that by disseminating the practice of considering declarations and explanations, educators may reflect and continually refine their assessment in response to the use of tools and services to tackle assessments.

2 Background

The challenge Generative Artificial Intelligence presents for CS assessment in higher education is evident [3, 17]. Programming exercises which were previously useful exercises for students to complete from scratch for assessment purposes can now be addressed *in-part* using Generative Artificial Intelligence services and tools [4]. Additionally, whilst not perfect, Generative Artificial Intelligence is also able to complete modeling tasks such as the generation of UML diagrams [3]. It seems inevitable that these

services and tools will continue to improve, and it will become increasingly challenging to continue with existing practice in higher education CS programmes. Stone outlines a scoping review into the impact and use of generative artificial intelligence in learning and assessing introductory programming [21]. Stone argues that the potential of generative artificial intelligence in excelling education is significant, if it can be harnessed effectively by educators and learners but that it requires more research.

JISC presents three general approaches to manage this shift: avoid, outrun or adapt [23]. If aiming to *avoid* the use of Generative Artificial Intelligence by students in an assessment one must return to assessments in controlled environments with no access to such resources. This is an appropriate choice where we must be certain students are able to demonstrate this knowledge and application thereof without access to external resources. This can be referred to as 'zero latency' knowledge, e.g. medical professionals in an emergency situation must triage effectively and rapidly without reviewing references. This option is arguably unrealistic for many assessments given resource constraints and assumes controlled environments can be consistently controlled to a level which completely stops such use. In practice, students may potentially find ways around such restrictions. Additionally, given many organisations incorporate Generative Artificial Intelligence tools, it is not unreasonable to expect students will be asked to work with them in future roles.

To *outrun* Generative Artificial Intelligence requires assessments to be re-designed to ensure that such systems cannot competently complete them. In this approach, one may look to incorporate elements which are less easily replicated by Generative Artificial Intelligence, such as embedding personal reflection, assessment of the process, and incorporating real world contexts and higher order skills [1, 11]. For example, Williams evaluated the performance across university biomedical essay type questions using three different Generative Artificial Intelligence tools (ChatGPT [18], Bard [7], and Bing[15]) and using different levels of study (1st year, 2nd year, 3rd year, postgraduate). The authors found that accuracy was generally high, with ChatGPT performing best across all tools, but performing least well at postgraduate level. However, Bing and Bard performed better at postgraduate level, indicating some variability between tools [24].

The final option proposed by JISC is to *adapt* where educators accept Generative Artificial Intelligence and incorporate this into assessments, e.g. using Generative Artificial Intelligence to generate a security policy and critiquing the output. Generative Artificial Intelligence has the potential to stimulate student learning [2, 10]. Another approach could be to ask students to declare and explain their use of Generative Artificial Intelligence and focus marking criteria on the process. For example, a programming project which does not assess the code itself, but instead the process of generating the code. Tang et al. argues that being transparent, declaring the use of Generative Artificial Intelligence is crucial in maintaining the credibility of a discipline and its instruments [22]. Similarly, Overono and Ditta argues that requiring students to declare use of such services will also prepare them to not only communicate ethically about their use, but also serve as an opportunity to prompt articulation of how output has been evaluated [19]. Consequently, there could be considerable value to educators and learners in requiring them to declare and explain use of Generative Artificial Intelligence.

Beyond the aforementioned approaches, there are other factors to consider. If educators embrace Generative Artificial Intelligence and expect it to be used by students, educators and institutions may unintentionally introduce inequalities. For example, better Generative Artificial Intelligence options are presently available for those able and willing to pay. This model could result in a stronger assessment submission, but is unlikely to be an option which all students are able to use [25]. Additionally, where students work as part of a team in an assessment which prohibits use of Generative Artificial Intelligence, we may face additional challenges of some team members using it without the knowledge of their teammates. It is clear there is much to consider when reflecting on assessment in the context of Generative Artificial Intelligence, but at present there is limited information on how students use such tools and services with existing assessments.

To better understand how students use generative Generative Artificial Intelligence in an existing assessment, this paper presents the practice of adjusting an existing assessment to require students to declare and explain their use of Generative Artificial Intelligence tools, provides a reflection on the reported use by students, and considers how this might inform assessment design in future.

3 Context

The context for the assessment was a course delivered in the final year of a three-year undergraduate program in computing science. The undergraduate program is delivered on an ongoing basis by the University of Glasgow as part of a transnational education partnership with the Singapore Institute of Technology. The course was focused on Forensics and considered theoretical and applied aspects of the topic. The course is delivered intensively in a block mode with assessments due on the final day of the block. A total of more than 140 undergraduate students were enrolled in the course. The students are typically from Singapore and the surrounding region. It is expected that many of the students speak multiple languages, with English as their first language, but this will not be the case for all students. The degree programme and comprising courses are delivered in English.

4 Assessment Design

The coursework comprised of two team assessments. Students were required to self-organised into an assessment team that comprised of no less than 3 members and no more than 4 members to complete both assessments.

The first assessment required teams to create a digital investigation process tailored for a fictional healthcare organisation that had suffered multiple data breaches. The teams were also expected to address other key challenges within the context including increasing data volume with investigation, privacy and legal issues as well as unpredictable and emerging threats due to rapidly changing technologies and patient behavior. In development of the processes, teams were required to take into account existing processes, case studies, legal concerns, and future challenges. The final digital investigation process must be justified with research and presented in a written report for a non-technical audience.

The second assessment required teams to evaluate a previously conducted investigation at a fictional organisation. The investigation considered the actions of a former employee that utilised steganography to exfiltrate data out of the organisation to evade security procedures. Teams had to critique the actions of the investigators as well as consider adherence to guidelines, best practice as well as ethical and legal concerns. This was then assessed through a 10-minute video presentation.

5 Declaration and Explanation of Generative Artificial Intelligence

Teams were required to declare and explain any use of Generative Artificial Intelligence in both coursework assignments. The declaration had to cover both assignments and had to be included in the appendix of the first assessment submission. The declaration approach broadly followed the approach recommended and used by Monash University [16]. Monash University provide guidelines to students that support them in being transparent about their use of Generative Artificial Intelligence. It requires students to explain the solutions employed as well as the rationale for employing them. It also requires students to document their interactions, in terms of the iterations as well as prompts used. Students are also required to describe the output received from tools and the changes they made to their work based on the feedback received from the tools they employed. In the present context, a similar approach was adopted. Teams were advised for each interaction with a Generative Artificial Intelligence solution they should (a) declare the system, (b) provide a rationale or explanation for using it, (c) the instruction issued to it and then the (d) output received. The class was advised the approach was modeled on that adopted at Monash University and provided a link to resources at the institution.

6 Lessons Learned

The more than 140 students formed 37 teams, comprising of either 3 or 4 members. There was 74 submissions in total, 2 submissions from each team for each assessment. 29 teams declared and explained use of Generative Artificial Intelligence for their coursework. 8 teams did not declare use of Generative Artificial Intelligence and the assumption is that those 8 teams (approximately 32 students) did not use Generative Artificial Intelligence in attempting the assignments.

The Generative Artificial Intelligence tools declared in descending order of popularity were: ChatGPT[18], Grammarly[8], Phind[20] Google Gemini[6] and Github Co-Pilot [5]. ChatGPT was the most popular service among teams, being declared by most of them. A single team reported use of Github Co-Pilot and this was to produce TeX source for their written submission. It should be noted their was no specific template for the written submissions and teams were permitted to submit Microsoft Word or PDF format. It is interesting to note that teams did not use one tool exclusively as some teams reported using multiple tools indicating students may have identified some tools are more useful for a specific task than others.

The details of use and the prompts provided were reviewed to identify common themes. A total of 4 themes were identified: *Written Quality*, *Research Assistance and Knowledge Building*, *Legal, Privacy, and Ethical Concerns* as well as *Presentation and Script Creation*. These 4 themes are considered in turn in the following sections. Additionally, we also include a further section, *Assessment and Feedback*, which does not represent a theme but was of particular interest and was reported by a single team.

6.0.1 Written Quality. . A number of teams reported using tools to improve the written quality of their submission, a prompt reported by P3, "*Improve the written language and readability*" is representative of many prompts reported by teams. Teams also reported using the tools to effectively produce parts of the written submission, P21 prompt: *"Help me to write these pointers into paragraphs"*.

Teams also declared more explicit content creation of written reports, P4 declared the prompt *"Summarise the detailed explanations of digital forensic investigation models into one short paragraph each"*, P8 declared: *"Summarise the healthcare sector's privacy, legal, and future challenges"*. These prompts are just some of the examples from teams that declared using tools to perform parts of the assessment.

Many teams declared prompts to tackle specific elements of the assessment, e.g. P21 declared *"Help me to improve the Ciardhuáin model section's paragraphs in the report"* and P19 stated *"Help me to improve the Casey model section's paragraphs in the report"*. Teams also used the tools to consider the accuracy of formatting and citation of written submissions. Teams were recommended to use the ACM format, but this was not a requirement. Any published and accepted citation format was permissible as long as it was used consistently and accurately. P17 declared the prompt was representative of most teams, *"Are my citations correctly cited in ACM format?"*.

6.0.2 Research Assistance and Knowledge Building. Teams also declared use of tools to try and understand and comprehend aspects of the problem they has been set. P1 declared the prompt *"Could you explain the use of digital investigation models in the healthcare industry?"* and *"What are some ways existing digital investigation models can be improved?"* as well as *"What makes an investigation forensically sound?"*. There was also examples of teams declare the use of tools to inspire and spark creatively, P10 declared the prompt *"Can you provide some ideas on how I can create a novel digital forensic investigation model?"*.

6.0.3 Legal, Privacy, and Ethical Concerns. Teams were also asked to consider legal, privacy and ethical concerns as part of the assignments. Teams declared use of tools to inspire tackling the requirement. For example P3 declared *"Ideas for privacy, legal, and future challenges in healthcare digital forensics"* and P6 declared *"Could you name me some privacy and legal considerations?"* whilst P13 stated *"Summarize privacy, legal, and future challenges in digital forensic investigations"*.

6.0.4 Presentation and Script Creation. Some teams reported use of the tools to generate content, in particular for the presentation e.g. P18 declared *"Based on the information above, generate a 2-minute presentation script"*. Teams also used tools to improve the overall quality of the submitted artefacts with some teams using tools to refine the quality of slides e.g. P14 declared the prompt *"Help me improve the quality of slides, in terms of language and terms used"*.

6.0.5 Assessment and Feedback. Whilst not a theme or common prompt, one team reported the use of the tools to assess and provide feedback on the quality of the work which was of interest. P20 declared *"Is this section worthy of an A based on the rubrics given to you in the image attached?"*. The team stated this was not to assess the work, but to gain insight into how to improve the written quality of the submission.

7 Discussion and Conclusions

The present paper has presented a review of students declarations and explanations of how they used Generative Artificial Intelligence tools on assessments for a single course. We identified four broad themes in how students are currently using these tools and services for existing assessments. Whilst some may come as no great surprise, such as to improve written quality, others perhaps less so. It was evident that many teams used the tools as an assistant - software they could have a conversation with which would aid their understanding of some of the key components of the assessment such as privacy and ethical concerns in the given problem domain. It appears Generative Artificial Intelligence was a starting point for students to consider how to approach the assessment rather than a shortcut to a solution.

There are many limitations in interpreting the experience of the practice presented, notably that this was carried out in a single course which is run in a block mode of teaching in single institution. The important aspect of this particular reported experience is that the educators learned something about how students were tackling their assessment, that they would not have known previously. Interpretation of declarations and explanations supports the educators in refining the assessment, not to make it resilient to generative artificial intelligence tools and services, but to ensure that the assessment is assessing what the setter intended.

If we consider consumer-level generative artificial intelligence tools and services at present we could liken it to use of calculator in an assessment. There are some assessments, specifically exams, where no calculators are permitted and then there are other assessments where can make it explicit that no calculators are permitted or we except some learners may use them. As many of the same questions are relevant in such an analogy, for example some students can afford more sophisticated calculators, some know how to use them more effectively etc.

For the time being, what most educators in computing science outside the realm of programming may want to consider asking learners to declare and explain use of such services and tools. The benefit for the learner is they can actually reflect and evaluate the output they are considering and what value they as individuals are bringing to the assessment. Similarly, the educator can reflect on the value of the assessment in terms of how it is motivating the learner to achieve the intended aims of the assessment.

References

[1] Matt Bower, Jodie Torrington, Jennifer W. M. Lai, Peter Petocz, and Mark Alfano. 2024. How should we change teaching and assessment in response to increasingly powerful generative Artificial Intelligence? Outcomes of the ChatGPT teacher survey. *Education and Information Technologies* (Jan. 2024). https://doi.org/10.1007/s10639-023-12405-0

[2] Ritvik Budhiraja, Ishika Joshi, Jagat Sesh Challa, Harshal D Akolekar, and Dhruv Kumar. 2024. "It's not like Jarvis, but it's pretty close!"-Examining ChatGPT's Usage among Undergraduate Students in Computer Science. In *Proceedings of the 26th Australasian Computing Education Conference*. 124–133.

[3] Javier Cámara, Javier Troya, Lola Burgueño, and Antonio Vallecillo. 2023. On the assessment of generative AI in modeling tasks: an experience report with ChatGPT and UML. *Software and Systems Modeling* 22, 3 (May 2023), 781–793. https://doi.org/10.1007/s10270-023-01105-5

[4] Marian Daun and Jennifer Brings. 2023. How ChatGPT Will Change Software Engineering Education. In *Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education V. 1 (ITiCSE 2023)*. Association for Computing Machinery, New York, NY, USA, 110–116. https://doi.org/10.1145/3587102.3588815

[5] GitHub. [n. d.]. GitHub Copilot. https://github.com/features/copilot. Accessed: 2024-10-26.

[6] Google. [n. d.]. Gemini AI. https://gemini.google.com/app. Accessed: 2024-10-26.

[7] Google. [n. d.]. Google Bard. https://bard.google.com. Accessed: 2024-10-26.

[8] Grammarly Inc. [n. d.]. Grammarly. https://www.grammarly.com. Accessed: 2024-10-26.

[9] Heather Johnston, Rebecca F Wells, Elizabeth M Shanks, Timothy Boey, and Bryony N Parsons. 2024. Student perspectives on the use of generative artificial intelligence technologies in higher education. *Int. J. Educ. Integr.* 20, 1 (Feb. 2024).

[10] Martin Jonsson and Jakob Tholander. 2022. Cracking the code: Co-coding with AI in creative programming education. In *Proceedings of the 14th Conference on Creativity and Cognition*. 5–14.

[11] Enkelejda Kasneci, Kathrin Sessler, Stefan Küchemann, Maria Bannert, Daryna Dementieva, Frank Fischer, Urs Gasser, Georg Groh, Stephan Günnemann, Eyke Hüllermeier, Stephan Krusche, Gitta Kutyniok, Tilman Michaeli, Claudia Nerdel, Jürgen Pfeffer, Oleksandra Poquet, Michael Sailer, Albrecht Schmidt, Tina Seidel, Matthias Stadler, Jochen Weller, Jochen Kuhn, and Gjergji Kasneci. 2023. ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences* 103 (April 2023), 102274. https://doi.org/10.1016/j.lindif.2023.102274

[12] Daniel Lee, Matthew Arnold, Amit Srivastava, Katrina Plastow, Peter Strelan, Florian Ploeckl, Dimitra Lekkas, and Edward Palmer. 2024. The impact of generative AI on higher education learning and teaching: A study of educators' perspectives. *Computers and Education: Artificial Intelligence* 6 (2024), 100221. https://doi.org/10.1016/j.caeai.2024.100221

[13] Jiahui Luo. 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* 49, 5 (2024), 651–664. https://doi.org/10.1080/02602938.2024.2309963 arXiv:https://doi.org/10.1080/02602938.2024.2309963

[14] Nora McDonald, Aditya Johri, Areej Ali, and Aayushi Hingle. 2024. Generative artificial intelligence in higher education: Evidence from an analysis of institutional policies and guidelines. *arXiv preprint arXiv:2402.01659* (2024).

[15] Microsoft. [n. d.]. Microsoft Copilot. https://www.microsoft.com/en-us/microsoft-365/copilot. Accessed: 2024-10-26.

[16] Monash University. 2023. Acknowledging the Use of Generative Artificial Intelligence. https://www.monash.edu/student-academic-success/build-digital-capabilities/create-online/acknowledging-the-use-of-generative-artificial-intelligence Accessed: 2024-09-22.

[17] Bayode Ogunleye, Kudirat Ibilola Zakariyyah, Oluwaseun Ajao, Olakunle Olayinka, and Hemlata Sharma. 2024. A Systematic Review of Generative AI for Teaching and Learning Practice. *Education Sciences* 14, 6 (June 2024), 636. https://doi.org/10.3390/educsci14060636

[18] OpenAI. [n. d.]. ChatGPT. https://chatgpt.com. Accessed: 2024-10-26.

[19] Acacia L Overono and Annie S Ditta. 2024. The use of AI disclosure statements in teaching: developing skills for psychologists of the future. *Teaching of Psychology* (2024), 00986283241275664.

[20] Phind, Inc. [n. d.]. Phind. https://www.phind.com. Accessed: 2024-10-26.

[21] Irene Stone. 2024. Exploring Human-Centered Approaches in Generative AI and Introductory Programming Research: A Scoping Review. In *Proceedings of the 2024 Conference on United Kingdom Ireland Computing Education Research (UKICER '24)*. ACM, Manchester, UK. Accessed: 2024-10-26.

[22] Arthur Tang, Kin-Kit Li, Kin On Kwok, Liujiao Cao, Stanley Luong, and Wilson Tam. 2024. The importance of transparency: Declaring the use of generative artificial intelligence (AI) in academic writing. *Journal of nursing scholarship* 56, 2 (2024), 314–318.

[23] M. Webb. [n. d.]. Generative AI - A Primer, Jisc. https://www.jisc.ac.uk/reports/generative-ai-a-primer

[24] Andrew Williams. 2024. Comparison of generative AI performance on undergraduate and postgraduate written assessments in the biomedical sciences. *International Journal of Educational Technology in Higher Education* 21, 1 (Sept. 2024). https://doi.org/10.1186/s41239-024-00485-y

[25] Ben Williamson and Rebecca Eynon. 2020. Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology* 45, 3 (July 2020), 223–235. https://doi.org/10.1080/17439884.2020.1798995