

AI detectors – a statement on the use of AI detection software in assessment at USW

Since the widespread emergence of Generative AI tools in 2023, several companies have attempted to bring to market AI detectors. These tools attempt to check submitted work for AI generated content. (2)

USW has chosen not to implement any AI detection tools at present. This is due to multiple unresolved issues with AI detector tools that make them currently unsuitable for implementation.

Studies into AI detection software have shown that AI detectors:

- Are unable to distinguish between AI generated and human-generated content without an unacceptably large number of false positives and false negatives. (1, 3, 4, 7)
- Are not transparent in how they determine a submission is AI generated and fail to provide proof of AI generation when queried. (1, 4, 5)

This means there is a significant risk of incorrect accusations of cheating being made against students, which could cause irreparable harm.

Additionally, AI detectors:

- Frequently flag more simple language as being AI generated when it is not. Not only does this represent a bias against students whose first language is not English, it also runs counter to our aspiration as a Higher Education provider to reduce barriers to access by encouraging writing in language that is more direct and accessible. (1, 5, 6)
- Can have the unintended effect of prompting students who have not used AI to use AI. For example, a student who writes directly and whose work is flagged as potentially written by AI, may then seek to use AI to increase the complexity of their language so as not to get flagged in future.
- Can flag software that is used genuinely by students as suspicious, for example, Grammarly. (2,9)
- Are unable to keep up with the rapid development of AI and software designed to subvert AI detection, for example, [Undetectable AI](#). (1,10)

At present no tool has addressed these concerns and our position is reflected across the sector (8).

The USW AI in Curriculum and Assessment Working Group will review the available AI detector tools on a termly basis and recommend any changes to the institution's position.

USW takes assessment and academic integrity very seriously providing robust, authentic, meaningful and fair assessments, whilst feedback is core to our organisational values and learning and teaching strategy. The University's position statement on AI can be viewed under [Regulations for Taught Courses](#).

We encourage any colleagues who may be concerned about the use of AI by students to engage with the guidance provided on [the AI Hub](#). Student guidance on the use of AI and academic integrity can be found on [LibGuides](#).

References:

1. Coley, M (2023, August 16). Guidance on AI Detection and Why We're Disabling Turnitin's AI Detector. <https://www.vanderbilt.edu/brightspace/2023/08/16/guidance-on-ai-detection-and-why-were-disabling-turnitins-ai-detector/>. Accessed on 31.01.24.
2. Bailey, J (2023, December 14). Should Students Stop Using Grammarly? <https://www.plagiarismtoday.com/2023/12/14/should-students-stop-using-grammarly/>. Accessed on 01.02.24.
3. Holston, B. (2023, September 7). Academic Writing Has Completely Changed: Turnitin Forges Ahead. <https://cwcaaccr.com/2023/09/07/chatgpt-has-changed-academic-writing-turnitin-forges-ahead/>. Accessed on 01.02.24.
4. Klee, M. (2023, June 6). She Was Falsely Accused of Cheating With AI — And She Won't Be the Last. Accessed on 31.01.24 <https://www.rollingstone.com/culture/culture-features/student-accused-ai-cheating-turnitin-1234747351>
5. Myers, A. (2023, May 15). AI-Detectors Biased Against Non-Native English Writers. Stanford University Institute for Human-Centered Artificial Intelligence. <https://hai.stanford.edu/news/ai-detectors-biased-against-non-native-english-writers> Accessed on 01.02.24.
6. Sample, I. (2023, July 10). Programs to Detect AI Discriminate Against Non-Native English Speakers, Shows Study. <https://www.theguardian.com/technology/2023/jul/10/programs-to-detect-ai-discriminate-against-non-native-english-speakers-shows-study>. Accessed on 01.02.24.
7. Staiman, A (2023, September 14). Publishers, Don't Use AI Detection Tools! <https://scholarlykitchen.sspnet.org/2023/09/14/publishers-dont-use-ai-detection-tools/>. Accessed on 01/02/24.
8. Staton, B (2023, April 3). Universities express doubt over tool to detect AI-powered plagiarism. <https://www.ft.com/content/d872d65d-dfd0-40b3-8db9-a17fea20c60c>. Accessed on 13.02.24.
9. Topinka, R (2024, February 13). The software says my student cheated using AI. They say they're innocent. Who do I believe? [The software says my student cheated using AI. They say they're innocent. Who do I believe? | Robert Topinka | The Guardian](https://www.theguardian.com/technology/2024/feb/13/the-software-says-my-student-cheated-using-ai-they-say-theyre-innocent-who-do-i-believe). Accessed on 14.02.24.
10. Undetectable AI <https://undetectable.ai/>. Accessed on 05/02/24.

Guidance links:

Regulations for Taught Courses: [Regulations for Taught Courses | University of South Wales](#)

USW AI Hub: <https://celt.southwales.ac.uk/policy/artificial-intelligence-ai-in-curriculum-and-assessment-hub/>

LibGuides: [Introduction - Artificial Intelligence at USW - LibGuides at University of South Wales](#)