Faculty Guide
Fairly Investigating Misuse of GenAI

The following guide provides direction to faculty on:

1. How to fairly detect misuse of AI-generated text
2. What to do if you detect misuse of AI-generated text in a learner's written submission

How to Fairly Detect AI-Generated Text

GenAI Detector Tools - to use or not to use?

Use of AI-generated text is presently very hard to identify. Currently, no singular AI detection tool, including that developed by Turnitin, has been deemed statistically reliable or accurate in their detection capabilities. This is coupled with the challenge for detection tools to also keep pace with the rapid evolution of AI language models. Furthermore, current detection tools can:

- Generate false positives; incorrectly identify human-written text as AI-generated
- Mislabel non-native English writing as AI-generated
- Raise concerns about bias against non-native English writers

For these reasons, a multi-pronged detection approach is best to prevent the occurrence of false academic misconduct allegations causing unnecessary stress and respectively, devastation for learners (Eaton, 2023; Elkhataat, 2023; Hikkila, 2023; Liang et al., 2023; Webb, 2023).

Remember to communicate clear guidelines around the use of GenAI in weekly schedules and all assignment instructions. Visit LTS for sample statements and further guidance.

Multi-pronged Approach to Uncovering GenAI Misuse (written work)

1

Analyze Submission for Learner's Voice

Look for unlikely transitions in writing, artificial sounding sentences as well as plagiarism that appears to have been created by cutting and pasting.

2

Compare Submission Against Past Samples

- Is the writing style, tone, or grammar changed?
- Is there a sudden increase in the quality of work from a learner who has been struggling?

3

Inconsistent with Assignment Guidelines

An AI generated submission may not be able to follow the instructions e.g. reference specific data and sources.

4

Credibility, Quality of Content and Sources

- Use of non-existent and/or non-relevant sources
- Inaccurate content

5

Use AI Detectors as a Conversation Tool

For responsible AI detection software usage and report interpretation, please visit LTS.

If using GenAI Detector tools, use ethically. Guiding Principles:

- Communicate usage at start of term for transparency and to avoid data privacy breaches and deceptive assessment accusations.

- Deceptive assessment is a breach of academic integrity on the part of the educator. Modelling ethical behaviour fosters learners acting with integrity.

- Inform learners of the use of detection tools (e.g., in weekly schedule, applicable assessment instructions, etc.), how the output will be used, and which detector(s) will be used.

- A single result from an AI-writing detection tool is not evidence to support academic misconduct allegations. It must never be used as the only evidence in an academic misconduct allegation. Instead, use the AI detection tool results to support your conversation with learners.

(Eaton, 2023; Gallant, 2023; Packback, 2023)

What To Do When Misuse of GenAI Text is Detected
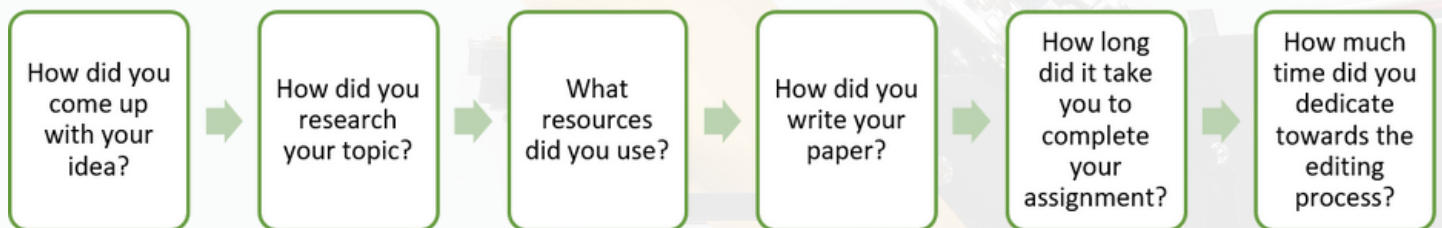
6   **Request a Meeting**

Request a meeting with the learner (virtually or in-person) to discuss your concerns as close to the time of observation as realistically possible. Avoid making an accusation, but rather have a conversation and focus on sharing your observations. Access to a sample email can be downloaded here.

7   **Meeting with Learner**

During your meeting with the learner, review the observations you made in Steps 1 through 5 above, while being careful not to make an accusation. Explain to the learner that there are concerns with their submission (e.g. fake sources, non-relevant sources, inconsistent with previous writing samples, inaccuracies in content, etc.).

Following your review of your observations, ask the learner to explain their writing process using some of the prompts below.

| How did you come up with your idea? | → | How did you research your topic? | → | What resources did you use? | → | How did you write your paper? | → | How long did it take you to complete your assignment? | → | How much time did you dedicate towards the editing process? |

Probe for general understanding of the concepts they presented. In some cases, you may ask the learner to re-write a shorter assignment in class (under your supervision) or resubmit the original assignment outside of class time, if you feel the conversation has led to a valuable learning opportunity.

Despite the multi-pronged approach offered in this guide, it is ultimately challenging to prove with full certainty that any learner has misused AI-generated text. Therefore, if time permits, and the circumstances are warranted, resubmission with advanced notice and within a suitable timeframe, may be an ideal option to consider.

If the learner cannot explain their assignment or their choices, or admits to misusing GenAI, document this and add it to the integrity violation allegation.

8 Report Incident to the AIO

If you continue to suspect misuse of GenAI content following a meeting with the learner, or the learner cannot explain their assignment or their choices, or admits to misusing GenAI, document this and complete the Academic Integrity Incident Report Form and attach all supporting documentation, including:

- Assignment instructions and rubrics
- Learner's submission
- Learner's previous writing sample (if available)
- All other evidence reviewed with the learner

Remember to notify the learner that you are submitting an Academic Integrity Incident Report to the Academic Integrity Office (AIO) prior to submitting the form. For assistance with filing a report, please contact the AIO at aio@algonquincollege.com.

References

AI in Education. (2023). *Won't My Students Use AI to Cheat?*
https://www.aiforeducation.io/blog/deg5axlmx45rdwz7k7yxvj1tv4mtw2

Eaton, S. (2023). *The Use of AI-Detection Tools in the Assessment of Student Work*
https://drsaraheaton.wordpress.com/2023/05/06/the-use-of-ai-detection-tools-in-the- assessment-of-student-work/

Elkhataat, A.M. Evaluating the authenticity of ChatGPT responses: a study on text-matching capabilities. *Int J Educ Integr* **19**, 15 (2023). https://doi.org/10.1007/s40979-023-00137- 0

Gallant, T. (2023). *GenAI, Cheating and Reporting to the AI Office*
https://adminrecords.ucsd.edu/Notices/2023/2023-5-17-1.html

Heikkila, M. (2023). *How to spot AI-generated text*
https://www.technologyreview.com/2022/12/19/1065596/how-to-spot-ai-generated-text/

Lian, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). *GPT detectors are biased against non-native English writers.* https://www.cell.com/patterns/fulltext/S2666-3899(23)00130-7

Packback. (2023). Professor Guide to AI Text Detection.
https://www.packback.co/pedagogy/professor-guide-to-ai-text-detection/

Webb, M. (2023). *AI Detection - Latest Recommendations.*
https://nationalcentreforai.jiscinvolve.org/wp/2023/09/18/ai-detection-latest-recommendations/

AC | Academic Integrity Office