

# Nashville State Community College

## AI Detection in Turnitin

How to use the AI writing detection tool in D2L as well as information on what AI detection tools are and what they do.

## Office of Online Learning

Last updated 05/18/2023

## What is AI Detection and What does it Do?

The job of an artificial intelligence (AI) detection tool is to determine whether content was generated by AI software or if it was created by a human. It does this by examining large online data sets to predict the possibility of words or phrases being utilized in a piece of content. AI detection tools, like Turnitin, do this by combining machine learning and natural language processing to identify those repeating patterns or phrases. In addition to repeating or predictable patterns, AI detection also looks for complexity, creativity, and lack of depth in language compared to that created by humans.

AI detection tools complete this process by using predicative text to generate content. As machines are more likely to use acritical patterns in their writing, they do not understand the meaning of words. Because of this, they require context to analyze text. This context comes from the AI examining the previously generated content and determining what the next word should be in relation to what is already there. If the content is found to follow predictable patterns, the AI detection tool will mark the content as being AI-generated.

## AI Detection and Turnitin

The Similarity Report in Turnitin now includes a new AI writing indication feature. This AI detection tool in Turnitin gives an overall percentage of the documents that may be AI-generated. While the AI writing indication tool can be found in the Similarity report, the AI writing detection percentage does not affect the Similarity Score. Turnitin's AI writing indication tool can detect text generated by GPT-3, GPT-3.5, and other variants including ChatGPT. They have also done testing with GPT-4 (ChatGPT-Plus) and are able to detect AI-generated content from it most of the time. Please note that as this new technology is evolving, it is not always perfect so please use your best judgement when assessing submitted work with this feature.

## What do the Different Color Indicators Mean?

If the AI writing indicator is blue, with a percentage between 1-100, then the submission has been processed successfully.

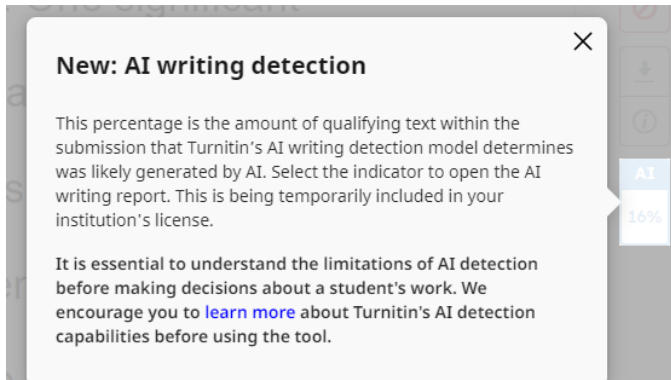
If the AI writing indicator is gray with no percentage or showing (--), then the AI writing detection indicator was unable to process the submission. There are several reasons this may happen.

- The submission did not meet the file requirements.
- The submission was made before the AI writing detection indicator was released in Turnitin.

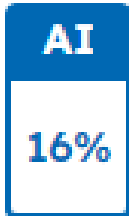
If the AI writing indicator reads Error (!), this means that it was Turnitin and not the AI writing detection indicator that failed to process the submission.

## Using the AI writing Indication Tool in Turnitin

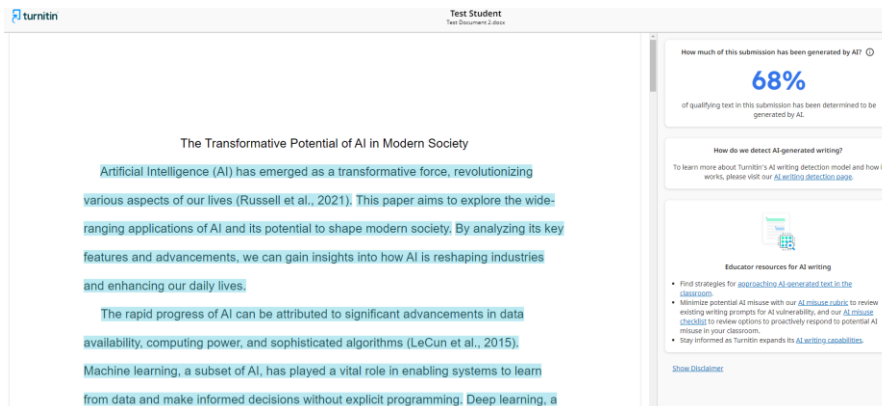
1. To view this new AI writing indication tool in Turnitin, go view the Similarity Report as usual after selecting an assignment. If you have not viewed a Similarity Report since the AI writing indication tool was added, the below message will appear.



2. Below the other tools on the left-hand side of the Similarity Report, is the new AI writing indication tool, which will appear with a percentage.



3. If you click on the AI writing indication tool, it will pull up the assignment, and you will be able to see everything the tool has highlighted that might have been generated by an AI.

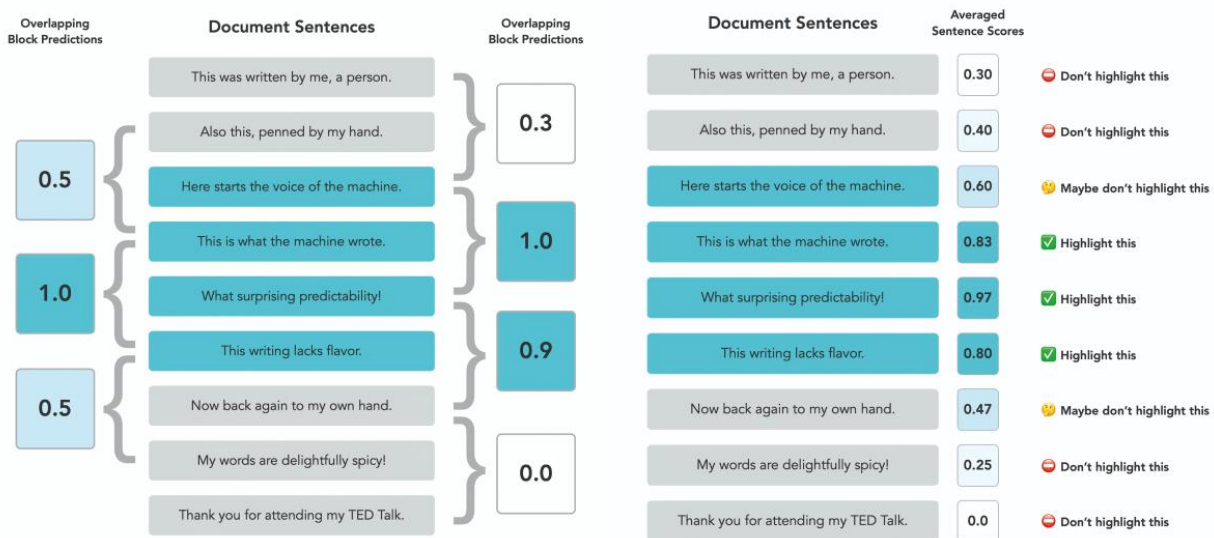


## How does Turnitin's AI Detection Work?

AI detection works in Turnitin by breaking submitted content into smaller chunks of text or sentences. Those segmented sentences are then overlapped to help capture each portion in context.

Next, those portions of text are run against the AI writing indicator in Turnitin. Each portion of text (or sentence) is given a score between 0 and 1 to determine if it was written by a human or if it is AI-generated. If a sentence or portion of text is determined to not be AI-generated it will earn a score of 0. If it is determined that an entire sentence was AI-generated, then it will earn a score of 1.

Based off the average scores earned by all segments of the original content, the AI writing indicator will create an overall prediction of how much text has been AI-generated. The AI writing indicator in Turnitin only flags text as being AI-generated if it is 98% certain it was written by an AI.



Example photo taken from [Turnitin's AI Writing Detection Capabilities – Frequently Asked Questions page](#) under the how does it work question.

## Educator Resources for AI Writing

1. To learn more about Turnitin's AI writing detection model and how it works, please visit their [AI writing detection page](#).
2. Find strategies for [approaching AI-generated text in the classroom](#).
3. Minimize potential AI misuse with our [AI misuse rubric](#) to review existing writing prompts for AI vulnerability, and our [AI misuse checklist](#) to review options to proactively respond to potential AI misuse in your classroom.
4. Stay informed as Turnitin expands its [AI writing capabilities](#).