

Falsely Flagged by an AI Detector?

What the research actually says before you panic

Dr. Tolulope Ajidahun

Licensed Medical Doctor with hands-on experience evaluating AI writing and detection tools.

An AI-detection score is not proof of misconduct. It is a probability estimate from a tool with a documented error rate, and that error rate is not evenly distributed across students. This guide covers what the published research actually shows, before any decision about an appeal gets made.

What the Research Actually Says

- Turnitin states its AI-writing detector maintains a false-positive rate under 1% on text flagged above its 20% reporting threshold, and its own guidance advises that a score “should not be the sole basis for adverse action” against a student.
- A Stanford-led study of seven widely used AI detectors found an average false-positive rate of 61.3% on essays written by non-native English speakers, meaning normal human writing was misclassified as AI-generated at a rate far above the vendor's published figure.
- Vanderbilt University disabled Turnitin's AI detector across its own campus in 2023 after internal testing, and its guidance to instructors states plainly that “AI detection software is not an effective tool” for disciplinary decisions.

Source: [Turnitin, Understanding False Positives in Our AI Writing Detection Capabilities](#)

Source: [University of San Diego Legal Research Center, Generative AI Detection Tools guide](#)

Source: [Vanderbilt University, Guidance on AI Detection and Why We're Disabling Turnitin's AI Detector](#)

The gap between the vendor's marketing number and independent testing is exactly the space an appeal lives in. Most institutions require more than a single algorithmic score to support a misconduct finding. That means there is a process, and a process can be engaged properly.

Where I Fit

I'm not a lawyer, and nothing here is legal advice. I work with AI writing and detection tools directly and understand how they actually score text: what triggers a false flag, what evidence holds up when a committee reviews a case, and how to present a defense so it gets read as carefully as the algorithm's score was.

The full kit builds on this research with three ready-to-use documents: a formal evidence request, a writing-process reconstruction guide, and an appeal letter template built around what committees actually respond to.

§ See the AI Detection Defense Kit for pricing and templates, or write in directly with your situation.

Full kit and pricing available on request at contact@drtolu.com.