



Guilty Until Proven Human

The Numbers Don't Lie.

The Detectors Do.

The data is damning. The consequences are real. And until now, students had no defence.

Company

False Flag Fixer

Website

dontturnitinyet.com

Tool Website

falseflagfixer.com

Launched

May 2026

Founder

Zakkiyah Daniels
Lawyer/Former English College Lecturer

Mission

To ensure no student is punished for writing their own work.

Key Stats

75M

Students flagged every month*

Extrapolated from Turnitin's 1.8B annual scan volume & independent false positive rates

[Based On Turnitin Scan Volume](#)

97.8%

One detector flagged almost every non-native essay as AI

[Stanford University Study](#)

61%

Non-native speaker essays wrongly flagged as AI

[Stanford University Study](#)

50%

Detector accuracy. Literally a coin flip.

[University of Maryland Research](#)

10+

Major universities that have disabled AI detection entirely

[Incl. MIT, Yale, Vanderbilt](#)

According to the data above, 2.5 million students are falsely flagged every single day. They wrote their own work. They followed the rules, and an algorithm decided they were cheaters. No warning. No explanation. Just a score, and a cloud over their academic future. AI detection tools were built to catch dishonesty. Instead they are punishing excellence.

This is not a glitch. It is a systemic failure. And for some students, Black students, international students, neurodivergent students, the risk is not just higher. It is devastating.

False Flag Fixer was built because no one else was building the defense.

I. False flags can ruin a student's life

A student submits their own essay. The algorithm flags it as AI-generated. They can appeal, but with what?

There is no technical proof. No standardised process. Just an algorithm's word against theirs. And the consequences while it plays out can be catastrophic: suspension, expulsion, a permanent academic record stain. All for writing too well.

Independent research from the [University of Maryland](#) and the [2025 HEPI Student Survey](#) confirms that major AI detectors carry false positive rates as high as 50% in certain academic contexts. That is not a margin of error. That is a coin flip deciding a student's future.

The "Founding Fathers" Paradox – The [U.S. Constitution was flagged](#) by a leading AI detector as "99% AI-Generated." Why? Because it is formal, structured, and precise. The exact qualities students are taught to strive for.

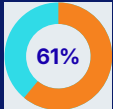
Excellence itself is now a red flag. ▶



ii. Linguistic redlining: the new digital divide

This is not just a student problem. It is a civil rights problem. False flags are not happening equally. The students most at risk are international students, Black students, and neurodivergent students.

There is a name for this: **Linguistic Redlining**.



Bias Against Non-Native Speakers

A landmark Stanford University study found that AI detectors misclassified over 61% of TOEFL essays written by non-native English speakers as "AI-generated." Writing well in your second language is now evidence of cheating.

3X

Risk for Diverse Voices

Black students and those writing in Standard Academic English are nearly three times more likely to be falsely accused. Detectors interpret formal, logical, structured writing, the hallmarks of excellence, as robotic.



The Neurodivergent Trap

Students with ADHD or Autism, who often write with highly structured and consistent patterns, are disproportionately flagged by algorithms that confuse logical consistency with machine output.

iii. The institutions that have already given up on AI detection

Major universities worldwide have disabled AI detection tools entirely, admitting publicly that the technology is too unreliable to use against students. If the tools do not work well enough for these institutions, why are students still being punished by them elsewhere?

University	Action	Date
MIT	Official guidance states: "AI Detectors Don't Work"	—
Yale University	Moved away from AI detection tools entirely	—
Vanderbilt University	Disabled Turnitin's AI detector. Cited false accusations and lack of transparency.	2023–24
University of Pittsburgh	"Not yet reliable enough to be deployed without substantial risk of false positives."	2023
Washington State University	Terminated Turnitin contract entirely. AI detection no longer used.	Feb 2026
University of Cape Town	"AI detectors are simply not reliable. There are no magic solutions."	Oct 2025
Curtin University	Disabled AI detection feature entirely.	Jan 2026
University of Waterloo	Turnitin AI detection tool discontinued.	Fall 2025
University of Arizona	Disabled AI detection due to reliability concerns and false positives.	Feb 2026

iv. The solution: False flag fixer

While the industry races to build better tools to catch students, False Flag Fixer is the first platform built to defend them. FFF uses AI to do what no other tool has done. It predicts exactly which words and phrases will trigger an AI detector before submission.

Students see their full risk score before they submit, not after the accusation. FFF identifies the exact trigger words and highlights them with reasons. Students swap flagged words with alternatives that preserve their voice. Real time scoring updates instantly.

The Appeals Kit — Free for All Users

Students facing accusations can download FFF's free Appeals Bundle. It's a practical toolkit including a checklist, step by step guidance, and template language to formally challenge a false flag.

The Appeals Report — Pro Users Only

Pro users get a personalised Appeals Report generated directly from their scan history. Drawing on Stanford and University of Maryland research, it gives students the exact evidence and language to fight back. No other tool offers this.

Why Now?

False Flag Fixer™ launched in April 2026 as the first tool of its kind built specifically to solve AI detection false positives. With AI detection now mandatory at thousands of institutions worldwide and no regulatory framework in place to protect students, the need has never been more urgent.

False Flag Fixer™ was built by a Bermudian lawyer and former English lecturer, working alone from Ho Chi Minh City, Vietnam with no coding background, no outside funding, and no development team. The entire product; the branding, the website, and the tool itself was built in under 5 months using AI as co-pilot.

A solo founder who used the very technology being weaponised against students to build their defence.



Founder Quote

"I built False Flag Fixer because I don't think any student should get accused of cheating for writing too well. As a lawyer and former English lecturer, I knew exactly how unjust that was — and I saw that no one was building the defense. So I did. A student should never have to prove their own innocence to an algorithm."

Zakkiyah Daniels, Founder, False Flag Fixer | Barrister-at-Law (non-practising) / Former English Lecturer

About False Flag Fixer

False Flag Fixer (falseflagfixer.com) is the world's first AI detection false positive pre-check tool, also known by its unforgettable domain, dontturnitinyet.com. Built to defend students who write their own work but get wrongly flagged by tools like Turnitin, GPTZero, and Originality.ai. False Flag Fixer uses AI to identify the exact trigger words that cause false positives, allows students to swap them while preserving their voice, and provides before-and-after risk scores as proof.

The message is simple: before you turn it in, run it through False Flag Fixer First.

False Flag Fixer was founded in 2026 by a Bermudian lawyer and former English lecturer.

**Not a humanizer. A detector for detectors. AI predicting AI.
That's FFF. The defense.**

Media Contct

Zakkiyah Daniels – Founder & Advocate
press@falseflagfixer.com

Location

Vietnam/Bermuda

Websites

dontturnitinyet.com
falseflagfixer.com

Verified Sources

All statistics and claims in this press kit are sourced from peer-reviewed research, institutional reports, and major publications.

Sources are listed in order of recency.

1

BREAKING — APRIL 2026

Opinion: AI detectors are hurting honest students. Schools should ban them.

The Washington Post, April 13, 2026

Published 5 days before this press kit. Confirms the urgency of the false positive crisis.
[washingtonpost.com/opinions/2026/04/13/ai-detectors-students](https://www.washingtonpost.com/opinions/2026/04/13/ai-detectors-students)

2

PEER-REVIEWED RESEARCH

GPT Detectors Are Biased Against Non-Native English Writers

Stanford University / Patterns Journal, July 2023 — Weixin Liang, James Zou et al.

Tested 7 AI detectors against TOEFL essays. Found 61% average false positive rate for non-native speakers; one detector flagged 97.8% of non-native essays as AI-generated.
hai.stanford.edu/news/ai-detectors-biased-against-non-native-english-writers

3

UNIVERSITY RESEARCH

Is AI-Generated Content Actually Detectable?

University of Maryland, Prof. Soheil Feizi, May 2023

Confirmed AI detectors drop to "the randomness of a coin flip" (~50% accuracy) when paraphrasing is applied. "Current detectors of AI are not reliable in practical scenarios."
research.umd.edu/news/ai-generated-content-actually-detectable

4

INSTITUTIONAL ACTION

Turnitin Is Out: WSU Terminates Contract with Turnitin

The Daily Evergreen / WSU Office of the Provost, February 2026

Washington State University terminated its Turnitin AI detection contract mid-semester, February 2026, citing high false positive rates and student distress.
dailyevergreen.com/195009/news/turnitin-is-out-wsu-terminates-contract-with-turnitin

5

INSTITUTIONAL ACTION

Curtin University to Disable Turnitin AI Detection Tool in 2026

EdTech Innovation Hub, September 2025

Curtin University disabled Turnitin's AI detection from January 1, 2026. "This change is about fostering trust and clarity within a modern academic culture."
<https://www.curtin.edu.au/news/oasis-news/update-on-turnitin-ai-detection-tool/>

6

OFFICIAL INSTITUTIONAL GUIDANCE

AI Detectors Don't Work. Here's What to Do Instead.

MIT Sloan Teaching & Learning Technologies

MIT's official faculty guidance page states directly that AI detectors do not work and advises alternative assessment strategies.
mitsloanedtech.mit.edu/ai/teach/ai-detectors-dont-work

7

TURNITIN OFFICIAL DISCLAIMER

Turnitin AI Writing Detection — Official Product Disclaimer

Turnitin.com (official product documentation)

"Our AI writing detection model may not always be accurate... it should not be used as the sole basis for adverse actions against a student." Turnitin also hides scores below 20% due to high false positive likelihood.
turnitin.com

8 MEDIA INVESTIGATION**We Tested a New ChatGPT-Detector for Teachers. It Flagged an Innocent Student.**

The Washington Post, April 2023 — Geoffrey A. Fowler

Washington Post investigation found false positive rates as high as 50% in real-world testing. Found Turnitin flagged genuine student work as AI-generated.

[washingtonpost.com/technology/2023/04/01/chatgpt-cheating-detection-turnitin](https://www.washingtonpost.com/technology/2023/04/01/chatgpt-cheating-detection-turnitin)

9 INSTITUTIONAL ACTION**Vanderbilt University Disables Turnitin AI Detector**

Inside Higher Ed, February 2024

Vanderbilt cited lack of transparency and potential harm from false accusations. Also confirmed: University of Texas at Austin and Northwestern University took similar action.

insidehighered.com/news/tech-innovation/artificial-intelligence/2024/02/09/professors-proceed-caution-using-ai

10 INSTITUTIONAL ACTION**University of Waterloo — Turnitin AI Detection Discontinued**

University of Waterloo official communications, Fall 2025

University of Waterloo confirmed Turnitin AI detection tool would no longer be available as of Fall 2025.

uwaterloo.ca

11 INSTITUTIONAL ACTION**University of Cape Town Scraps AI Detectors**

October 2025

"AI detectors are simply not reliable. There are no magic solutions." — UCT Director of Academic Integrity.

uct.ac.za

12 SCAN VOLUME DATA**Turnitin Scan Volume and False Positive Extrapolation**

Based on Turnitin's published figure of 1.8 billion essays scanned annually (approx. 150 million monthly) *At a conservative 30–50% false positive rate (per University of Maryland and Washington Post research), this produces an estimated 45–75 million false flags per month — approximately 2.5 million per day.*

turnitin.com

For additional research, data, or interview requests contact:
press@falseflagfixer.com