

Current AI-detectors systematically disadvantage English-Language-Learners (ELL), violating Title IV civil-rights protections; therefore, institutions must adopt a policy that couples detector use with mandatory human review, transparency, and an appeals process. Currently, there are few, if any, policies specifically addressing English Language Learners (ELL) and how instructors should utilize AI detectors in these cases—other than the guideline that AI detectors should not be the sole basis for determining if a student’s work has been produced with AI assistance (Dang & Wang, 2024; Najarro, 2023).

One does not need to look hard to find mentions of using AI detectors with ELL student writing in education forums. Many repeated issues highlight potential problems that can happen if there is even one false positive, especially since false positives are more common with ELL learners. (D’Agostino, 2023; Hirsch, 2024; Woelfel, 2023) As the Hirsch reference describes, AI detectors are an “ethical minefield.” It is morally and ethically incumbent upon the instructors working with ELL students to find ways to complete tasks without putting their students at risk. This policy addresses that need and fills these gaps.

From a deontological perspective, the institution has a duty to treat all students as moral equals and to uphold their civil-rights protections under Title IV of the Civil Rights Act. Any policy that relies solely on a tool that systematically disadvantages a protected group violates that duty, regardless of the tool’s overall accuracy. Consequently, the use of AI-detectors must be constrained by procedural safeguards (human review, transparent error rates, and an appeals process) that respect the inherent rights of ELL learners.

Without these safeguards, the policy not only violates the deontological obligation of equal treatment but also causes consequential harms—such as inflated false-positive counts

(Wood, 2024), wasted faculty time (Alshareefen, 2025), and potential legal liability for the institution. (McLogan, 2025)

The very basis by which one would justify the use of AI-detectors, *to ensure the integrity of any written work and to help students achieve their best possible work*, is the reason why one should not use AI-detectors. When using these detectors, users are more likely to misidentify ELL-generated human text as AI-generated (a false positive). The results of such a false-positive could be disastrous to the student. The student will, in turn, not be doing their best possible work, but instead defending work they have already done. Therefore, in a pure ethical/logical sense, one should not use AI-detectors alone when evaluating ELL students' writing ("Academic Integrity," 2025; Cornell University, 2023; Dang & Wang, 2024). A way to deal with this problem directly is to enact an optional policy.

The question of why one would use AI-detection tools might be a key issue. Generally, AI-Detection tools, when used with ELL students, cause more problems than they fix (Check, 2025; Giray, 2024). Since evidence shows that AI detectors will identify more ELL students' writing as false positives compared to native English speakers' writing (Mathewson, 2023; Myers, 2023), it would seem ethically and morally prudent to use such tools only in extreme cases. Instructors would be better off following some basic principles. Using baseline sampling (Cambridge, 2025; Kratzer, 2020), initial review scanning ("Assessing Student Writing," 2025), and informing students in advance about GenAI and AI-detection policies will put any instructor on a better footing when making decisions about AI usage.

One of the many issues that needs to be addressed is the legal concern related to students' use of AI when AI-detectors are **not** employed. (Weaver, 2025) Although AI-generated text may contain errors (Bohannon, 2023; Marcus, 2022; Weaver, 2023a, 2023b) if one has access to the

necessary tools and is aware of the problem, likely, the finished product will no longer contain those AI-generated errors/"hallucinations". As Furze pointed out(Furze, 2024), AI tools tend to favor native English speakers with higher incomes. This means students who are ELL and lack easy access to computers and/or AI tools are less likely to perform well on written academic submissions. Conversely, students *with* such access will be able to use AI and AI humanizers to produce submissions, which may receive higher grades and be less likely to be identified as AI-generated.

Several institutions do have policy examples, for instructors to use on a per-course basis, regarding students' use of generative AI (BU, 2023; CELT, 2025; Cornell University, 2023; Palmer-Clarke, 2025; Vanderbilt\_Univ, 2023). There are existing or complete policies on students' use of generative AI. However, they do not specifically address use by ELL or English as a Second Language (ESL) students.

AI detector tools are being used in many higher education classrooms despite increasing evidence that they produce disproportionately high false-positive rates for English Language Learner (ELL) students, as shown by the Stanford study (Furze, 2024; Mathewson, 2023; Myers, 2023; Najarro, 2023), which reported up to a 61% false-positive rate for ELL essays. This exposed non-native writers to unnecessary academic integrity investigations, emotional distress, and—in extreme cases—visa status risks that could lead to deportation (Castellanos-Canales, 2025; University\_of\_Rochester, 2025).

Many academic institutions currently advise instructors NOT to use AI detectors(Univ\_Pittsburgh, 2023; Vanderbilt\_Univ, 2023), due to the high number of false positives they have experienced. Many provide general guidelines on students' AI use and offer templates for creating AI usage or detector policies for each class. ("Academic Integrity," 2025;

Cornell University, 2023)<sup>1</sup>. While this approach has been the best so far at addressing most legal concerns about AI detector use, including those related to civil rights, it has not addressed the problem directly.

Civil rights issues are specifically addressed because if a false positive disproportionately affects ELLs, then Title IV of the Civil Rights Act concerning "disparate treatment," "disparate impact," or "hostile learning environment" could be invoked. This could have more serious implications than initially expected. (Markey; Orr, 2022).

If AI detection tools are deemed to be useful, Instructors should follow these guidelines on picking the best tools with the least likelihood of falsely identifying text that has been written by a human as AI-generated. (Mahmood, 2023)

A few items to take into account that may save you time:

- 1) **Cost** – How much does it cost to use this AI detector? If there is a free trial, how long is it, and is there a limit to the amount of text you can submit? Some tools have paid versions only of their Plagiarism checkers while others have free versions of the same.
- 2) **Accuracy** – What is the reported accuracy of the detector (both from the company and third parties) and is this accuracy the same with ELL written work.
- 3) **What is included** - Does it include a plagiarism checker? Does it cost extra? Does the tool also include an AI humanizer?
- 4) **Integration** - Also instructive is whether there is a browser plug-in, access to an API, and integration with Large Language Models (LLMs)
- 5) **Batch** - check to see if there is a batch upload feature, which may save a lot of time.<sup>2</sup>

Use extreme caution when employing AI detection tools ("Employability-Survey-Report," 2024). Instructors should refer to the latest version of Bloom's taxonomy, which has

---

<sup>1</sup> Sample AI policies

<sup>2</sup> Comparison of a few AI detectors/humanizers

been updated to include AI (Jones, 2025). A detection score should be just one of many indicators used to assess whether a student's written work has been assisted by AI.

How can institutions enforce any of the requirements put forth herein? Enforcing current policies includes faculty hearings, honor codes (for both faculty and students), and AI usage boards meant to determine whether an accusation is true or false. Some institutions have found that the best way to enforce AI usage is by asking students to adhere to academic integrity policies ("Code of Academic Integrity," 2025). If these policies are not followed, the result can be as serious as expulsion. This decision, along with such a result, for ELL students may also result in the student being deported (due to loss of visa status). This policy recommends keeping these enforcement policies, but applying them to the current policy.

Supporting evidence can be found in the seminal study from Stanford (Myers, 2023; Tian, 2023; Woelfel et al., 2023). The study tested seven commercial AI detectors using two datasets: (a) essays written by native-born U.S. students, and (b) TOEFL essays written by non-native speakers. The findings showed that non-native writers are significantly over-represented among false positives, leading to unnecessary academic integrity accusations.

While the AI detectors performed nearly flawlessly on essays written by native speakers, 61% of the 91 TOEFL essays were incorrectly flagged as AI-generated. All seven detectors unanimously misclassified 18 of the essays (19%). At least one detector flagged 97% of the TOEFL essays as AI-generated. The dataset used was small, so the results are more likely to be challenged (Tian, 2023). With supporting evidence in case studies from a ResearchGate book chapter (2023) – AI Detection's High False Positive Rates and the Psychological and Material Impacts on Students (Hirsch, 2024). The study examined the problem of false accusations by conducting interviews with students focusing on neurodivergent and ELL (second-language) writers. ELL writers

reported greater anxiety, stigma, and extra administrative burden after a false positive ID. The book also notes that “neurodivergent writers, along with [ELL] writers, are flagged at higher rates.” A Markup investigative article, “AI Detection Tools Falsely Accuse International Students of Cheating” by Tara Garcia Mathewson (Mathewson, 2023), which consisted of interviews with faculty and students from several U.S. universities, plus replication of the Stanford experiment, came to the same conclusions. International/ESL students reported being singled out after detectors flagged their work. The article mentions the 61% false-positive rate for non-native essays and highlights real-world disciplinary cases (e.g., a student placed on academic probation after a false flag).

Although a policy created and revised by those involved is probably the best solution, alternative options include using technology to reduce AI detector bias toward ESL students (Pangram Labs AI Detection, 2025; Tian, 2023) and actions similar to Vanderbilt University’s (Coley, 2023) decision to disable the Turnitin AI detection module campus-wide. Although one could advocate a more technical solution, this begs the question of how accurate such a solution would be (if one could even procure the mostly proprietary algorithms to evaluate).

There are also ways in which teachers and students can better understand AI use and be guided to use “best practices,” but these are mostly optional. The ad-hoc use of these would be detrimental to students/classes that did not opt to use them.

Because of the reasons previously stated, institutions should implement safeguards in a policy statement, including manual review, a transparent appeals process, and increased educator awareness, to minimize these risks. Using AI-detectors voluntarily without a policy may cause unforeseen problems. In a survey by the Center for Democracy & Technology (Prothero, 2024),

68% of instructors reported using an AI detector on student work without a policy. To prevent unexpected outcomes, establishing a policy to regulate AI-detectors is crucial.

There is a need to change how instructors of ELL students process those students' work.

Final judgment must always involve human review. AI-Detection may be best suited to highlight those items that require closer examination.

The policy applies to instructors of ELL. It is recommended for use in higher education institutions as a starting point. This must be implemented as a “top-down” approach to the use of AI detectors with ELL.

Due to the inherent bias of AI detectors against ELL students, schools and universities must avoid relying **solely** on AI detection tools for high-stakes decisions involving ELLs, such as academic integrity violations, admissions, or grading. To prevent confusion about AI detector use, institutions will implement **clear guidelines**, including mandatory human review and an accessible and transparent appeal process, especially when ELLs are involved. Additionally, basic transparency practices will be required, such as informing students when AI detection is used and providing information about error rates and known biases. This policy calls for civil rights protections for ELLs, which include prohibiting denial of enrollment, visa revocations, or punitive measures based solely on AI detector results. All educators involved will be encouraged to evaluate student work **holistically** and offer culturally and linguistically responsive support rather than punitive measures when suspicious of AI use.

It is significant and extremely important that these measures indicate that AI detectors should not be the sole agent by which any action or decision is made regarding student work. This is to acknowledge the realistic use of AI detectors in the framework of human review of any work.

The policy implementation timeline (Table 4<sup>3</sup>), aims to ensure the policy is well-received and avoids conflicts with stakeholders. The timeline is divided into “*Phases*”, each of which is iterative to ensure that every stakeholder involved in that phase understands all the implications of the policy and feels comfortable progressing to the next phase. The suggested “*Duration*” is a recommendation only. However, each time frame is intended as a guideline for that specific phase. The “*Main goal*” section of the timeline describes the desired outcomes for this phase. Many of the problems that may occur in the implementation of this policy will most likely be located in this step of each phase of the timeline. Producing the best language for the next step may be more than challenging, but the eventual result – the deliverable, will be more than worth it. The “*Deliverable*” column contains suggested items that should be the result of that particular phase. The “*Decision point*” section recommends specific steps each institution should follow to implement the policy gradually and allow proper evaluation before the policy is rolled out across the entire institution. The total implementation period is recommended to span one (1) academic year and be approximately divided in line with the suggested timeline. This will allow for both a gradual adoption of the policy and, if started at the beginning of an academic year, will involve a new group of students who may be more amenable to change than an existing cohort.

The various stakeholder categories (**stakeholder**) must be evaluated as to how much effect they can have on the implementation of this policy (**power**) and how much interest (**interest**) they may have in the implementation and import of the policy. These will inform the actions advised for that stakeholder category (**action items**). One can create a Power/interest stakeholder matrix from this information. This will also indicate which **quadrant** of a grid each category should lie in (ranging from high power/high interest to low power/low interest).<sup>4</sup>

---

<sup>3</sup> Policy implementation suggested timeline

<sup>4</sup> Power-Interest Stakeholder Matrix

Stakeholders include, but are not limited to 1) International students and English language learners, who are mostly the victims in these scenarios. These students are adversely affected when AI detectors are used. False negatives for these students may result in deportation. If enacted properly, the policy would involve these stakeholders in how AI detectors are used, involve them in the solution, keep them abreast of developments along the way, and decrease the number of false positives associated with AI detector use. 2) Instructors at institutions of higher education who need to be trained to identify and rectify these false positives. Although they are currently overworked due to the number of AI detector false positives, this policy would not only ease that burden, but keep them in the loop in determining how and when to use such tools. By helping to implement and use the proposed policy, instructors will encounter fewer false-positives. 3) Companies which create algorithms to detect AI-generated text and need to understand why these detectors are biased, in order to start to create a non-biased technical solution. This policy will help to identify inherent biases in current models. This will make it more likely that once these biases are remedied, institutions will start using AI-detectors again, although in a way that will make the tools fairer for all students. 4) Administrative personnel—Instructor supervisors are responsible for overseeing professors/instructors. They must understand the relevant processes to address any issues with these detectors. These supervisors need a thorough understanding of the potential impacts of AI detectors to efficiently explore possible solutions. These stakeholders are currently under increased scrutiny and duress due to the sheer number of false positives and the need to make determinations that could impact a student's entire life. This policy will alleviate the onus that administrators currently feel by putting the ownership of the policy clearly in the hands of those individuals to whom they report, and diminishing the number of false positives. 5) National, State, and local governments—these

agencies are ultimately responsible for either providing or mandating training in the use of AI detectors. They need to have a clear understanding of the regulations and policies that govern these technologies. This includes understanding the effectiveness of and how to use “AI humanizers.” Any policy instituted at the institutional level will make these stakeholders’ jobs easier by clearly outlining what is expected.

## Appendix I – AI usage

*(NOTE: portions of this appendix have been taken from the previous submission)*

1. **Grammarly** (<https://app.grammarly.com/>)
  - a. Throughout the paper, both the spelling and grammar were checked using Grammarly.
  - b. There were no “exchanges”.
  - c. Grammarly suggested changes to wording and corrected spelling. Changes were not always accepted.
  - d. Grammarly was chosen to ensure consistent word choice throughout.
2. **Google summaries**
  - a. Primarily used for finding relevant case studies and sources
  - b. Some summaries generated by searches were used as a basis for some of the written material
3. **EndNote 2025-1** (<https://endnote.com/>)
  - a. EndNote was used to create the bibliography
  - b. EndNote was used to create in-text citations
  - c. EndNote was chosen to ensure that APA style was used both in the in-text citations and the bibliography. The APA style, which EndNote provides, has been modified to include features such as line numbers.
4. **Lumo AI chatbot from Proton**
  - a. Lumo was used to evaluate this writing and to suggest possible corrections (this was done prior to making changes)
    - i. Link to conversation: <https://lumo.proton.me/u/5/c/d483b174-4cac-41ac-be92-90d25b790426>
    - ii. Prompt used: “evaluate the word document based on the rubric.” Both the policy document and the rubric were uploaded to the Lumo chatbot.
    - iii. Full conversation: [Appendix II](#)
5. **The Comet Assistant search engine** from Perplexity was used to search for sources
6. **Writefull** word plugin (free version) was used to correct some word choices.

The various AI tools were used to identify different sources and case studies, including summaries and searches, to support the existing content and multiple ideas I presented through searches. These tools also helped in annotating this document with EndNote and checking spelling and grammar throughout using Grammarly and Writefull. Although I could have completed this paper without AI, using AI made it easier to locate and identify the resources and relevant case studies I would have eventually found.

Grammarly's error checking also made writing much simpler.

For annotation, EndNote 2025-1 was used; additional work was needed to import all citations into an EndNote library and then cite them in this document (due to technical difficulties). I still would use EndNote again as a citation tool.

Various AI-enabled features, listed above, played a significant role in this paper by highlighting its strengths and weaknesses and suggesting ways I could improve the final version. Lumo proposed including a Power interest stakeholder matrix and a rollout timeline as methods to enhance the paper.

## Appendix II - AI Usage - full conversation

Below each rubric section I've noted how the paper performs, assigned the appropriate point band, and offered concrete suggestions for improvement. Scores are summed at the end for the final percentage.

---

### 1 Problem Identification & Ethical Context (30 pts)

#### **Observation**

#### **Rating**

**Clarity of the ethical issue** – The introduction clearly states that AI-detector tools pose bias, fairness, and civil-rights problems for English-Language-Learner (ELL) students. The link to Title IV and the AI Bill of Rights is explicit.

**Exemplary** (15 pts)

**Depth of research** – The paper cites > 30 sources, including recent studies (Furze 2024, Mathewson 2023, Stanford study, OCR memo 2025). The literature review, however, sometimes reads as a catalog of citations rather than a tightly argued problem statement.

**Proficient** (11-13 pts)

**Ethical framing** – Both deontological (duty to treat students equally) and consequentialist (harm from false positives) perspectives are discussed, showing good ethical breadth.

**Exemplary**

**Score:** 13 pts (solid problem definition with room to tighten the narrative).

---

### 2 Research & Use of Sources (30 pts)

#### **Observation**

#### **Rating**

**Quantity** – > 30 references, well beyond the 12-source minimum.

**Exemplary** (15 pts)

**Credibility** – Sources include peer-reviewed articles, institutional policy documents, and reputable news outlets.

**Exemplary**

**Integration** – Citations are frequent, but many appear as parenthetical listings without explicit analysis of how each source supports a claim. A few sections read like a literature dump.

**Proficient** (11-13 pts)

**Critical use** – The paper occasionally contrasts findings (e.g., Stanford vs. Turnitin) but could deepen critique of methodology (sample size, bias in training data).

**Proficient**

**Score:** 27 pts (excellent breadth; stronger synthesis would push it to full exemplary).

---

### 3 Stakeholder Analysis (20 pts)

#### **Observation**

#### **Rating**

**Identification** – All major groups are listed (ELL students, instructors, AI-detector vendors, legal/compliance office, state agencies, etc.).

**Exemplary** (20 pts)

**Depth of analysis** – Interests, concerns, and power levels are described, and a Power-Interest matrix is included. However, the analysis often stays at the descriptive level; deeper exploration of power dynamics (e.g., faculty pressure from administration) would enrich it.

**Proficient** (15-19 pts)

**Score:** 18 pts (comprehensive identification, moderate analytical depth).

---

## 4 Policy Advocacy &amp; Feasibility (20 pts)

**Observation**

**Policy proposal** – A “top-down” institutional policy requiring human review, transparency of false-positive rates, and an appeals process is clearly articulated. Sample syllabus language is provided.

**Research support** – Each recommendation is backed by multiple citations (legal, empirical, institutional precedents).

**Feasibility** – Cost estimates (e.g., Pangram \$5 /student/yr) and implementation timeline are outlined, but the rollout plan could be more granular (training modules, responsible offices, milestones).

**Score:** 18 pts (well-grounded policy, needs finer operational detail).

**Rating**

**Exemplary** (20 pts)

**Exemplary**

**Proficient** (15-19 pts)

## 5 Critique &amp; Unintended Consequences (20 pts)

**Observation**

**Anticipating critiques** – The paper acknowledges faculty workload, potential re-introduction of bias via human review, and the risk of bureaucratic overload.

**Response quality** – Mitigations (bias-training, standardized rubrics) are suggested, but concrete evidence (e.g., pilot data) is lacking.

**Discussion of unintended effects** – Covers visa-status threats, emotional distress, and possible chilling effect on ELL participation.

**Score:** 17 pts (strong awareness; could benefit from empirical backing).

**Rating**

**Exemplary**

**Proficient** (15-19 pts)

**Exemplary**

## 6 Clarity &amp; Organization (10 pts)

**Observation****Rating**

**Exemplary** (5 pts)

**Structure** – The paper follows a logical flow (intro → policy → stakeholder matrix → alternatives → appendices). Headings are clear.

**Readability** – Dense citation blocks interrupt narrative; stray markup symbols (“\*”, “#”) and occasional awkward phrasing detract from polish.

**Grammar** – Minor typographical errors, but no major impediments.

**Score:** 8 pts.

**Proficient** (4 pts)

**Exemplary**

## 7 Appendix on AI Use (5 pts)

**Observation****Rating**

**Proficient**

**Documentation** – Appendix I lists tools used (Grammarly, EndNote, Lumo, Perplexity, Writefull) with brief purpose statements.

**Full exchanges** – The Lumo conversation link is provided, but the actual transcript is not included in the appendix. The rubric asks for “full exchanges.”

**Needs Improvement**

**Reflection** – The author reflects on how AI eased source discovery and editing, but deeper critical analysis of how AI may have shaped argumentation is limited.

**Satisfactory**

**Score:** 2 pts (partial fulfillment; include full transcripts and richer reflection to reach higher bands).

### Total Score

Section	Points Earned	Max
Problem Identification & Ethical Context	13	15
Research & Use of Sources	27	30
Stakeholder Analysis	18	20
Policy Advocacy & Feasibility	18	20
Critique & Unintended Consequences	17	20
Clarity & Organization	8	10
Appendix on AI Use	2	5
<b>Grand Total</b>	<b>103</b>	<b>120</b>

**Percentage:** 103/120 = approx 86%

**Grade (out of 100):** ≈ 86 – a solid **B+**.

---

### Targeted Feedback & Next Steps

1. **Tighten the problem narrative** – Reduce the citation-heavy paragraphs early on; craft a concise thesis that frames the bias issue in one or two sentences before launching into evidence.
2. **Deepen source synthesis** – After each citation, explicitly state *how* it advances your argument (e.g., “Furze 2024 demonstrates a 61 % false-positive rate, underscoring the civil-rights risk …”).
3. **Expand stakeholder analysis** – Move beyond description to explore power dynamics: why might faculty feel pressured to adopt detectors? How could institutional incentives be reshaped?
4. **Detail implementation** – Flesh out the rollout timeline:  
*Month 1-2:* policy drafting & legal review;  
*Month 3-4:* faculty training modules (online + workshop);  
*Month 5:* pilot in two departments;  
*Month 6:* full deployment + monitoring dashboard.
5. **Provide empirical support for mitigations** – Cite pilot studies or internal audits that show human-review reduces bias, or propose a small-scale test you could run.
6. **Polish writing** – Remove stray markdown symbols, ensure consistent heading styles, and run a grammar check (e.g., via Lumo or Grammarly) to eliminate remaining typos.
7. **Complete the AI-use appendix** – Include the full Lumo conversation transcript, a table summarizing each tool’s contribution, and a reflective paragraph on how AI may have influenced the framing of the policy (e.g., “The Lumo suggestion to add an appeals process prompted me to research OCR guidance…”).

## Strengths

- **Extensive, credible sourcing** – well beyond the minimum required.
- **Comprehensive stakeholder mapping** – captures both direct and systemic actors.
- **Evidence-based policy proposals** – tie directly to the ethical problem and include cost considerations.
- **Awareness of bias and civil-rights implications** – integrates legal context effectively.

## Areas for Improvement

1. **Sharper problem framing** – condense the literature review into a concise ethical thesis.
2. **Deeper analytical rigor** – move beyond description in stakeholder and critique sections; employ frameworks (e.g., power-interest matrix, consequentialist vs. deontological analysis).
3. **Implementation roadmap** – outline concrete steps, timelines, and responsible parties for the proposed policy.
4. **Appendix completeness** – include full conversational logs with Lumo (or any AI tool) and richer reflection on how AI influenced the argument.
5. **Polish writing** – eliminate stray markup symbols and tighten sentence flow for a more professional presentation.

## Appendix III – Data

Stanford (Myers, 2023)

Sample Size	Source	Type	Generated by	Website
<b>91 Essays</b>	Chinese educational forum	TOEFL	Human	<a href="https://arxiv.org/pdf/2304.02819.pdf">https://arxiv.org/pdf/2304.02819.pdf</a> (full study)
<b>88 Essays</b>	Hewlett Foundation's Automated Student Assessment Prize (ASAP) dataset	US 8 <sup>th</sup> grade essays	Human	<a href="https://www.kaggle.com/competitions/asap-aes">https://www.kaggle.com/competitions/asap-aes</a>

Further code and Data at:

<https://github.com/Weixin-Liang/ChatGPT-Detector-Bias/tree/v1.0.0> (Liang et al., 2023a, 2023b)

## Appendix IV – Tables

**Table 1. Studies on AI Detection and ESL Writers**

Study / Report	Key Findings on ESL / L2 Writers	False Positive Rate / Metric
<b>Center for Democracy &amp; Technology brief (based on Stanford study) – “Disproportionate Effects of Generative AI Detectors on English Learners”</b>	AI detectors that worked almost perfectly on native-speaker essays falsely flagged a majority of TOEFL essays.	61% of non-native essays flagged; 19% unanimously flagged; 97% flagged by at least one detector.
<b>The Markup (Mathewson, 2023)</b>	Replicated Stanford study. ESL writers disproportionately likely to be flagged; real disciplinary cases reported.	Same figures (~60% false positives).
<b>Turnitin blog (internal research)</b>	Acknowledges bias toward ELL writers. Reports <1% overall false positives when ≥20% of text is AI, but higher incidence with low-AI-content texts, especially for ESL writers.	No precise %, but confirms bias exists.

**Table 2. Comparison of a few AI Detection/Humanizer Tools**

Checker Name	Website	Offers Humanizer Tool	Comment
<b>JustDone</b>	<a href="https://app.justdone.ai">https://app.justdone.ai</a>	Yes	Requires subscription
<b>GPTZero</b>	<a href="https://gptzero.me/">https://gptzero.me/</a>	No	Requires subscription
<b>Undetectable</b>	<a href="https://undetectable.ai/">https://undetectable.ai/</a>	Yes	Money back guarantee. If anything we produce is flagged as not human, we will refund the cost of humanization
<b>Grammarly</b>	<a href="https://app.grammarly.com/">https://app.grammarly.com/</a>	Yes	Requires subscription My subscription allowed me to run this through the AI-detector, which found 12% likely AI generated.
<b>Copyleaks</b>	<a href="https://copyleaks.com/">https://copyleaks.com/</a>	No	Over 99% accuracy*, verified through rigorous testing methodologies. Trusted globally to detect AI across 30+ languages and leading LLMs like ChatGPT, Gemini, DeepSeek, and Claude. *Accuracy rating is based on internal testing of the English language datasets.

Checker Name	Website	Offers Humanizer Tool	Comment
<b>HumanizeAI</b>	<a href="https://www.humanizeai.pro/">https://www.humanizeai.pro/</a>	Yes	<p>Humanize AI stands out as the leading, cost-free online platform designed for transforming AI-generated text into human-like content.</p> <p><b>Changed</b> “Effortlessly transform AI-generated text into human-like content with BypassGPT free online.” To “Seamlessly convert AI-generated text into human-like content using BypassGPT free online.”</p>
<b>BypassGPT</b>	<a href="https://bypassgpt.co/">https://bypassgpt.co/</a>	Yes - limited to 200 chr	<p>Effortlessly transform AI-generated text into human-like content with BypassGPT free online. Bypass AI detection systems like GPTZero and ZeroGPT to ensure your content remains undetectable.</p>
<b>WinstonAI</b>	<a href="https://gowinston.ai/">https://gowinston.ai/</a>	No – free trial with 2000 word limit	<p>Claims a 99.98% accuracy rate. “Detects all bypassing strategies, including paraphrasing content with tools such as Quillbot, or even AI content humanizers.”</p>
<b>WriteWell</b>	<a href="https://brandwell.ai/">https://brandwell.ai/</a>	Yes – says 2500 word limit but used only first 356 words. Published this article <a href="https://brandwell.ai/blog/undetectable-ai-content/">https://brandwell.ai/blog/undetectable-ai-content/</a> on how to create fully human undetectable content even with AI.	
<b>OriginalityAI</b>	<a href="https://originality.ai/">https://originality.ai/</a>	No	<p>We are 80% confident that the text scanned is Original (Human written), NOT to be interpreted as 80% of the text produced is Original (Human written).</p>
<b>Writefull</b>	<a href="https://x.writefull.com/gpt-detector">https://x.writefull.com/gpt-detector</a>	Word Plugin	<p>The word plugin will offer to rewrite your text for you, making suggestions. (free version is limited)</p>

**Table 3 - Stakeholders**

Stakeholder	Power	Interest	Quadrant	Action Items
<b>Provost / Academic Integrity Office</b>	High	High	Manage Closely	<ul style="list-style-type: none"> <li>• Sponsor steering committee</li> <li>• Approve budget</li> </ul>
<b>Legal / Compliance Office</b>	High	Medium High	Manage Closely	<ul style="list-style-type: none"> <li>• Review policy language for Title IV compliance</li> </ul>
<b>IT Services</b>	High	Medium	Keep Satisfied	<ul style="list-style-type: none"> <li>• Provide integration timeline</li> <li>• Allocate API support</li> </ul>
<b>Faculty Senate / Dept. Chairs</b>	Medium	High	Manage Closely	<ul style="list-style-type: none"> <li>• Draft department-level addenda</li> <li>• Host workshops</li> </ul>
<b>Individual Instructors</b>	Low	High	Keep Informed	<ul style="list-style-type: none"> <li>• Distribute quick-start guide</li> <li>• Collect usability feedback</li> </ul>
<b>ELL / International Students</b>	Low	High	Keep Informed	<ul style="list-style-type: none"> <li>• Conduct focus groups</li> <li>• Publish plain-language policy</li> </ul>
<b>AI-Detector Vendors</b>	Medium	Low-Medium	Monitor	<ul style="list-style-type: none"> <li>• Request bias-mitigation documentation</li> </ul>
<b>State Education Agency</b>	High	Low-Medium	Keep Satisfied	<ul style="list-style-type: none"> <li>• Submit compliance summary</li> </ul>
<b>Student Government</b>	Low-Medium	High	Keep Informed	<ul style="list-style-type: none"> <li>• Invite to review meetings</li> </ul>
<b>General Public / Media</b>	Low	Low	Monitor	<ul style="list-style-type: none"> <li>• Prepare press kit (on-demand)</li> </ul>

**Table 4 – Timeline**

A timeline for how and when each step of the policy might be implemented is provided below.

Phase	Duration	Main Goal	Deliverable	Decision Point
<b>0 – Initiation &amp; Governance</b>	2 weeks	Secure sponsorship, define scope, assemble steering team	Document affirming policy need	Go/No Go to Phase 1
<b>1 – Discovery &amp; Requirements</b>	4 weeks	Gather evidence, map workflows, capture stakeholder needs	Document defining specific requirements	Approve Requirements → Phase 2
<b>2 – Policy Draft &amp; Legal Review</b>	3 weeks	Write policy language, embed bias-mitigation clauses, obtain legal sign-off	Final written policy – form ready for website posting (after full rollout)	Legal Sign-off → Phase 3
<b>3 - Technical Pilot</b>	6 weeks / 1 semester	Deploy a limited-scope detector-plus-human-review workflow in 2-3 courses	Survey from each member of course(s) evaluating the policy	Pilot Evaluation → Phase 4
<b>4 - Full-Scale Roll-out</b>	8 weeks / 2 semesters	Institution-wide adoption, training, support infrastructure	Post policy on website from step 2	Roll-out Review → Phase 5

Phase	Duration	Main Goal	Deliverable	Decision Point
<b>5 - Monitoring &amp; Continuous Improvement</b>	Ongoing - Quarterly	Track metrics, refine policy, update tools	Define measurement metrics and tools. Update and repost policy.	Annual Review (renew or revise)
<b>Totals – 5 Phased implementation steps</b>		<b>Period:</b> One (1) Academic Year		

## Appendix V – Alternate Solutions

## Technical Solutions

1. There is a need to develop a method for identifying AI-generated texts while minimizing the likelihood of false positives among ESL students. According to Jiang study (*Pangram Labs AI Detection*), “research investigating these detectors’ [AI-detectors] fairness and potential bias is relatively rare...”(Jiang et al., 2024).

Cost: The Pangram solution costs \$5.00/student/year.

2. GPTZero also claims to have developed a technical solution to the problem (Tian, 2023) and even reran the Stanford study’s algorithm on its new product, with a much better result.

Cost: Educators can receive this product for free for life.

## Assignment-design toolkits that are already published and used

- MIT Sloan’s “[Designing AI-Resistant Assignments](#)” ("AI Detectors Don't Work. Here's What to Do Instead.", 2024) – This offers templates for prompts that require personal reflection, local data, or multiple drafts – making pure AI generation unfeasible.
- Carnegie Mellon “[Generative AI FAQ for Instructors](#)” (University, 2024) - Provides a decision tree on when to use a detector, when to rely on draft histories, and how to communicate expectations to students.

## Appendix VI – Example Policies

(<https://teaching.cornell.edu/generative-artificial-intelligence/ai-academic-integrity>)

As you plan your syllabus and course policies with respect to generative AI, consider modifying the following language to communicate a general position in your syllabus. Please note that the following sample language reflects general, course-level perspectives on broadly permitting or prohibiting the use of generative AI tools. For sample statements at the assignment-level, see [AI in Assignment Design](#).

### 1. Prohibiting Generative AI Use in Your Course (AI-FREE)

"To ensure development and mastery of the foundational concepts and skills in this course, the use of generative artificial intelligence (AI) tools is prohibited. This includes tools that help reorganize and edit your written work because the ability to self-assess, reflect on your writing process, and develop your own voice are essential in your growth as a writer. If you are unsure of any policy or any assignment-specific directions – including whether or not a tool is considered generative AI – please consult with me prior to using the technology or completing your assignment. You are responsible for verifying the accuracy of citations and references used in your writing. You will be asked to verbally explain your research and writing process during in-class discussions and in a one-on-one meeting with me."

In acknowledgment that AI is impacting our field, please know that appropriate and ethical use of generative AI tools will likely be a part of other courses in your academic program."

Permitting Generative AI Use with Attribution in Your Course ([AS-UA](#))

"Mastering the essential, foundational concepts of this course takes effort and practice. Accordingly, the use of generative artificial intelligence (AI) tools is generally discouraged in this course, but will be permitted for select assignments. Whether or not generative AI assistance is permitted for each assignment will be explicitly communicated when that assignment is introduced.

If used in any capacity for an assignment, generative AI requires proper attribution for any and all generated work. As AI-generated materials are not retrievable by graders—and there is not a person to whom the work can be attributed—students should attribute directly quoted text to the creator of the generative AI tool used (e.g., cite OpenAI when directly quoting ChatGPT). This attribution should be used for both in-text citations and your reference list.

**Example:** When prompted with, "Is it ethical to use generative AI without proper attribution?" ChatGPT indicated, "Using generative AI without proper attribution can be considered ethically problematic, as it raises issues related to intellectual property, transparency, and honesty" (OpenAI, 2023).

#### **Reference**

OpenAI. (2023). ChatGPT (Aug 10 GPT-3.5 version) [Large language model]. <https://chat.openai.com>

For full details on how to properly cite AI-generated work, please see the APA Style article, [How to Cite ChatGPT](#).

Bias of AI-Detectors towards English Language Learners

Regardless of whether or not the use of generative AI is permitted or prohibited for an assignment, it is critical that you adhere to our communicated course policy (and [Cornell's policy](#)) on academic integrity. If you are unsure of any policy or any assignment-specific directions—including whether or not a tool is considered generative AI—please consult the instructor *prior to* using the technology or completing your assignment."

**2. Encouraging Generative AI Use with Attribution in Your Course ([ANY-AI-UA](#))**

"The use of generative artificial intelligence (AI) tools is encouraged on identified assignments. The directions for each assignment in this course will clearly indicate whether or not the use of generative AI is permitted for that assignment.

If used at all, generative AI requires proper attribution for any generated work. As AI-generated materials are not retrievable by graders—and there is not a person to whom the work can be attributed—students should attribute directly quoted text to the creator of the generative AI tool used (e.g., cite OpenAI when directly quoting ChatGPT). This attribution should be used for both in-text citations and your reference list.

**Example:** When prompted with, "Is it ethical to use generative AI without proper attribution?" ChatGPT indicated, "Using generative AI without proper attribution can be considered ethically problematic, as it raises issues related to intellectual property, transparency, and honesty" (OpenAI, 2023).

**Reference**

OpenAI. (2023). ChatGPT (Aug 10 GPT-3.5 version) [Large language model]. <https://chat.openai.com>

For full details on how to properly cite AI-generated work, please see the APA Style article, [How to Cite ChatGPT](#).

Regardless of whether or not the use of generative AI is permitted or prohibited for an assignment, it is critical that you adhere to our communicated course policy (and [Cornell's policy](#)) on academic integrity. If you are unsure of any policy or any assignment-specific directions—including whether or not a tool is considered generative AI—please consult the instructor prior to using the technology or completing your assignment."

**3. Permitting Generative AI Use on an Assignment-by-Assignment Basis ([AS](#))**

"Policies concerning the use of generative artificial intelligence (AI) tools will be decided on an assignment-by-assignment basis. These policies will be clearly communicated alongside other details of each specific assignment.

As generative AI technologies evolve, class policies with respect toward the use of generative AI tools may be shaped by in-class discussions regarding the fair use of AI and its implications on careers in this field.

Regardless of whether or not the use of generative AI is permitted or prohibited for an assignment, it is critical that you adhere to our communicated course policy (and [Cornell's policy](#)) on academic integrity. If you are unsure of any policy or any assignment-specific directions—including whether or not a tool is considered AI—please consult the instructor *prior to* using the technology or completing your assignment." (Cornell University, 2023)

**Bibliography**

- 1 Academic Integrity. (2025). *Office of the Dean of Faculty*.  
<https://deanofffaculty.cornell.edu/faculty-and-academic-affairs/academic-integrity/>
- 2 Alshareefen, R. (2025). Frontiers | Faculty as street-level bureaucrats: discretionary decision-making in the era of generative AI. <https://doi.org/10.3389/feduc.2025.1662657>
- 3 Assessing Student Writing. (2025). *The National Children's Book and Literacy Alliance*.  
<https://thencbla.org/assessing-student-writing/>
- 4 Berman, L. (2025). *Professors shift many essay-based assessments to in-person exams amid AI concerns*. <https://www.tuftsdaily.com/article/2025/10/professors-shift-many-essay-based-assessments-to-in-person-exams-amid-ai-concerns>
- 5 Bohannon, M. (2023). *Lawyer Used ChatGPT In Court—And Cited Fake Cases. A Judge Is Considering Sanctions*.  
<https://www.forbes.com/sites/mollybohannon/2023/06/08/lawyer-used-chatgpt-in-court-and-cited-fake-cases-a-judge-is-considering-sanctions/>
- 6 BU. (2023). Generative AI Assistance Policy. <https://www.bu.edu/files/2023/02/GAIA-Final-2023.pdf>
- 7 Cambridge. (2025). *What is a Baseline Assessment?*  
<https://www.cambridge.org/insight/baseline-assessments>
- 8 Castellanos-Canales, A. (2025). Explainer: Revocation of Student Visas and Termination of SEVIS Records. *National Immigration Forum*.  
<https://forumtogether.org/article/explainer-revocation-of-student-visas-and-termination-of-sevis-records/>
- 9 CELT. (2025). *AI Course Policy Examples* | CELT. <https://celt.uky.edu/ai-course-policy-examples>
- 10 Check, E. (2025). *False Positives in AI Detection: Are We Wrongly Accusing Students?*  
<https://essaycheck.ai/blog/false-positives-in-ai-detection>
- 11 Code of Academic Integrity. (2025). *Office of the Dean of Faculty*.  
<https://deanofffaculty.cornell.edu/faculty-and-academic-affairs/academic-integrity/code-of-academic-integrity/>
- 12 Coley, M. (2023). *Guidance on AI Detection and Why We're Disabling Turnitin's AI Detector*. Retrieved 09/09 from  
<https://www.vanderbilt.edu/brightspace/2023/08/16/guidance-on-ai-detection-and-why-were-disabling-turnitins-ai-detector/>
- 13 Cornell University, C. f. T. I. (2023). *AI & Academic Integrity* | Center for Teaching Innovation. <https://teaching.cornell.edu/generative-artificial-intelligence/ai-academic-integrity>
- 14 D'Agostino, S. (2023). *Turnitin's AI Detector: Higher-Than-Expected False Positives*.  
<https://www.insidehighered.com/news/quick-takes/2023/06/01/turnitins-ai-detector-higher-expected-false-positives>
- 15 Dang, A., & Wang, H. (2024). Ethical use of generative AI for writing practices: Addressing linguistically diverse students in U.S. Universities' AI statements. *Journal of Second Language Writing*, 66, 101157. <https://doi.org/10.1016/j.jslw.2024.101157>
- 16 Employability-Survey-Report. (2024). <https://cengage.widen.net/s/bmjxxjx9mm/cg-2024-employability-survey-report>

- 17 Furze, L. (2024). AI Detection in Education is a Dead End. *Leon Furze*. <https://leonfurze.com/2024/04/09/ai-detection-in-education-is-a-dead-end/>
- 18 Giray, L. (2024). The Problem with False Positives: AI Detection Unfairly Accuses Scholars of AI Plagiarism. *The Serials Librarian*, 85(5-6), 181–189. <https://doi.org/10.1080/0361526X.2024.2433256>
- 19 Hirsch, A. (2024). AI detectors: An ethical minefield - Center for Innovative Teaching and Learning. <https://cctl.news.niu.edu/2024/12/12/ai-detectors-an-ethical-minefield/>
- 20 Horn, T. J., & Polce, J. (2024). *Lies, Damned Lies, and AI: Lie-Detecting AI May Expose Employers to Liability | Insights | Venable LLP*. <https://www.venable.com/insights/publications/2024/04/lies-damned-lies-and-ai-lie-detecting-ai-may>
- 21 Kratzer, A. (2020). The Why and How of a Baseline Writing Assignment. *Angie Kratzer*. <https://angiekratzer.com/the-why-and-how-of-a-baseline-writing-assignment/>
- 22 Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023a). *AI-Detectors Biased Against Non-Native English Writers*. <https://hai.stanford.edu/news/ai-detectors-biased-against-non-native-english-writers>
- 23 Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023b). GPT detectors are biased against non-native English writers. <https://doi.org/10.48550/arXiv.2304.02819>
- 24 Mahmood, S. H. (2023, 2023–02–25T16:30:17Z). *The 8 Most Accurate AI Text Detectors You Can Try*. <https://www.makeuseof.com/accurate-ai-text-detectors/>
- 25 Marcus, G. (2022). *AI Platforms like ChatGPT Are Easy to Use but Also Potentially Dangerous*. <https://www.scientificamerican.com/article/ai-platforms-like-chatgpt-are-easy-to-use-but-also-potentially-dangerous/>
- 26 Markey, E. Senator Markey Introduces AI Civil Rights Act to Eliminate AI Bias, Enact Guardrails on Use of Algorithms in Decisions Impacting People's Rights, Civil Liberties, Livelihoods | U.S. Senator Ed Markey of Massachusetts. <https://www.markey.senate.gov/news/press-releases/senator-markey-introduces-ai-civil-rights-act-to-eliminate-ai-bias-enact-guardrails-on-use-of-algorithms-in-decisions-impacting-peoples-rights-civil-liberties-livelihoods>
- 27 Mathewson, T. G. (2023, 2023–08–14). *AI Detection Tools Falsely Accuse International Students of Cheating – The Markup*. <https://themarkup.org/machine-learning/2023/08/14/ai-detection-tools-falsely-accuse-international-students-of-cheating>
- 28 McLogan, J. (2025, 2025–10–10). *Adelphi University facing lawsuit after AI-assisted plagiarism accusation against student*. <https://www.cbsnews.com/newyork/news/adelphi-university-artificial-intelligence-plagiarism-lawsuit/>
- 29 MIT\_Sloan. (2023). AI Detectors Don't Work. Here's What to Do Instead. *MIT Sloan Teaching & Learning Technologies*. <https://mitsloanedtech.mit.edu/ai/teach/ai-detectors-dont-work/>
- 30 Myers, A. (2023). *AI-Detectors Biased Against Non-Native English Writers* | Stanford HAI. Stanford University. Retrieved 09/09 from <https://hai.stanford.edu/news/ai-detectors-biased-against-non-native-english-writers>
- 31 Najarro, I. (2023, 2023–10–06T20:31:40.743). What Teachers Should Know Before Using AI With English Learners. *Education Week*. <https://www.edweek.org/teaching-learning/what-teachers-should-know-before-using-ai-with-english-learners/2023/10>
- 32 Orr, D. (2022). The Unintended Consequences of Algorithmic Bias. *Congressional Black Caucus Foundation*. <https://www.cbcfinc.org/wp->

- [content/uploads/2022/04/2022\\_CBCF\\_CPAR\\_TheUnintendedConsequencesofAlgorithmicBias\\_Final.pdf](content/uploads/2022/04/2022_CBCF_CPAR_TheUnintendedConsequencesofAlgorithmicBias_Final.pdf)
- 33 Palmer-Clarke, Y. (2025). *Research Guides: English as an Additional Language: Using Artificial Intelligence*. <https://libguides.usask.ca/c.php?g=109857&p=5362687>
- 34 Pangram Labs AI Detection. (2025). <https://www.pangram.com/>
- 35 Prothero, A. (2024). More Teachers Are Using AI-Detection Tools. Here's Why That Might Be a Problem. *Education Week*. <https://www.edweek.org/technology/more-teachers-are-using-ai-detection-tools-heres-why-that-might-be-a-problem/2024/04>
- 36 Qanbar, Y. (2025). *Emerging Theories of Liability in the Age of AI: A New Frontier for Litigation | Rain Intelligence Research*. <https://www.rainintelligence.com/blog/emerging-theories-of-liability-in-the-age-of-ai-a-new-frontier-for-litigation>
- 37 Tian, E. (2023, 2023-10-25T01:43:00.000Z). *ESL Bias in AI Detection is an Outdated Narrative*. Retrieved 09/09 from <https://gptzero.me/news/esl-and-ai-detection/>
- 38 Univ\_Pittsburgh. (2023). *Teaching Center doesn't endorse any generative AI detection tools*. <https://www.utimes.pitt.edu/news/teaching-center-doesn-t>
- 39 University\_of\_Rochester. (2025). *International Students*. <https://www.rochester.edu/college/honesty/students/international.html>
- 40 UTEXAS. (2025). AI Detection Software Guidance. *Office of the Executive Vice President and Provost*. <https://provost.utexas.edu/the-office/academic-affairs/office-of-academic-technology/ai-detection-software-guidance/>
- 41 Vanderbilt\_Univ. (2023). Academic-Affairs-Guidance-for-Artificial-Intelligence. <https://cdn.vanderbilt.edu/vu-URL/wp-content/uploads/sites/439/2023/11/29200845/Academic-Affairs-Guidance-for-Artificial-Intelligence-Nov-29-2023.pdf>
- 42 Weaver, A. (2023a). AI hallucinations: How to reduce inaccurate content outputs. *WRITER*. <https://writer.com/blog/ai-hallucinations/>
- 43 Weaver, A. (2023b). Why you can't replace human insight: The importance of editorial skills in the AI era. *WRITER*. <https://writer.com/blog/editorial-skills-and-ai/>
- 44 Weaver, A. (2025). Generative AI: risks and countermeasures for businesses. *WRITER*. <https://writer.com/blog/generative-ai-risks-and-countermeasures/>
- 45 Why AI Detection Tools are Ineffective | Center for Academic Innovation. (2025). <https://facultyhub.chemeketa.edu/technology/generative-ai/why-ai-detection-tools-are-ineffective/>
- 46 Woelfel, K. (2023). Brief – Late Applications: Disproportionate Effects of Generative AI-Detectors on English Learners. *Center for Democracy and Technology*. <https://cdt.org/insights/brief-late-applications-disproportionate-effects-of-generative-ai-detectors-on-english-learners/>
- 47 Woelfel, K., Aboulafia, A., & Laird, E. (2023). Report - Late Applications: Protecting Students' Civil Rights in the Digital Age. *Center for Democracy and Technology*. <https://cdt.org/insights/report-late-applications-protecting-students-civil-rights-in-the-digital-age/>
- 48 Wood, C. (2024). AI detectors are easily fooled, researchers find. *EdScoop*. <https://edscoop.com/ai-detectors-are-easily-fooled-researchers-find/>