

ДОНИКА ГАВОВА



StrikePlagiarism.com

ORIGINALITY IS A VALUE

# Етични насоки при прилагане на ИИ в учебния процес

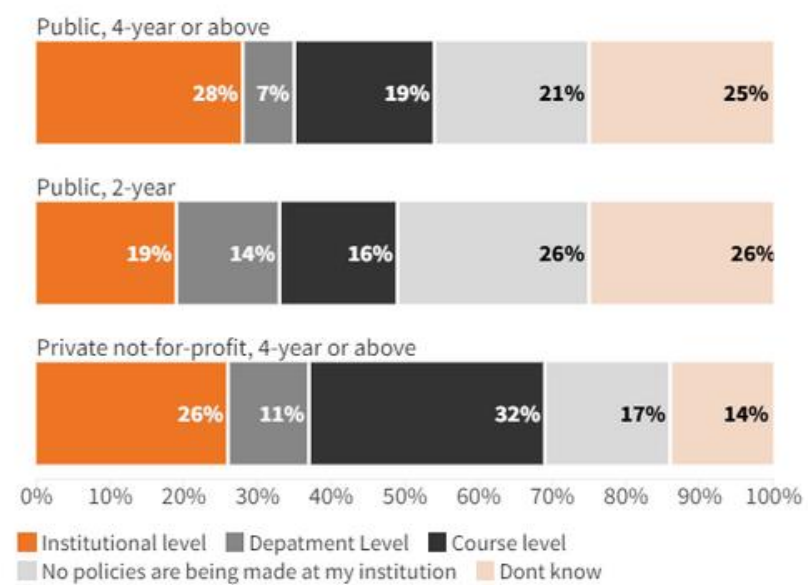


# СТАР ПРОБЛЕМ - НОВИ ИЗМЕРЕНИЯ

- ▶ “Copy and paste” е популярен метод в изготвянето на писмени работи
- ▶ „Ghost writer” и продажбата на завършени писмени работи по интернет
- ▶ **„Коучинг”** срещу заплащане
- ▶ Преводно плагиатство
- ▶ ИИ плагиатство – нова форма на измама в училищата и университетите

### AI policy-making

Where are policies regarding the use of generative AI writing tools like ChatGPT primarily being made at your institution?



Source: Fall 2023 Faculty Pulse Survey, Tyton Partners analysis

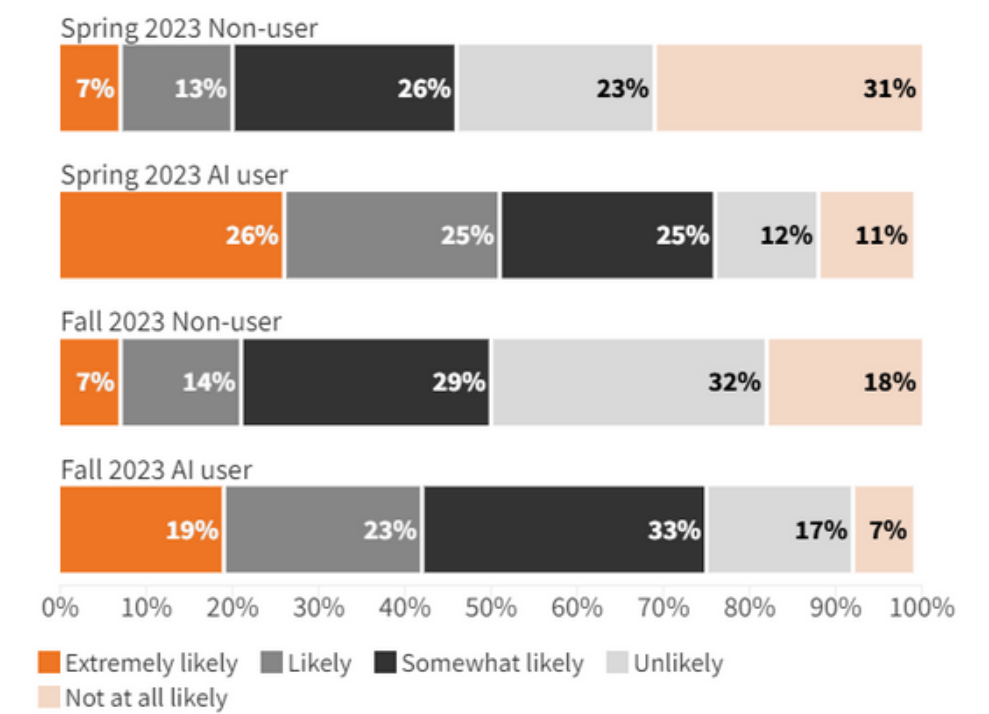
About a third (32 percent) of private, four-year institutions rely on instructors to set their own policy, although more than a quarter (26 percent) have institutionwide policies. Public four-year institutions are split on policies: 28 percent say policies at the institution have been set, 25 percent say they do not know if there are any policies and 21 percent said no policies were made at their institution.

Meanwhile, public two-year institutions have the lowest rate of policy adoption: 52 percent said they do not have a policy or do not know if they have a policy.

Проучване, проведено от Tyton Partners, потвърждава положителната динамика на използването на изкуствен интелект от ученици и студенти в САЩ. Повече от 70% от представителите на тези групи твърдят, че активно използват изкуствен интелект.

### Using AI tools

Student likelihood to continue to use GenAI writing tools despite restrictions



Source: Fall 2023 Faculty Pulse Survey, Tyton Partners analysis

Повечето училища и университети в САЩ и ЕС прилагат правила и процедури за използването на **изкуствен интелект**.



# ИИ В УЧЕБНИЯ ПРОЦЕС

## ПРОБЛЕМИ:

- Може да предостави фалшива, невярна информация.
- Тя не позволява да се развият уменията за търсене на информация, сравняване.
- Няма достъп до най-актуалната информация.
  - **Формира пристрастяване**
  - **Води до редица етични проблеми**

## ВЪЗМОЖНОСТИ:

- Спестява време за търсене на информация.
- Спестява време за анализ.
- Дава подробен отговор от различни източници.

# ЩО Е ТО? ЗЛОУПОТРЕБА С ИИ?

Примерите за неправилна употреба на ИИ включват, но не се ограничават до следното:

**Копиране или перифразирание** на части от съдържанието, генерирано от изкуствен интелект, така че работата вече не е **собствена работа на ученика**.

Копиране или перифразирание на цели отговори, генерирани от ИИ, **без позоваване на ИИ**.

Използване на изкуствен интелект за извършване на части от работата, така че работата да не отразява **независимост при анализа, оценката или изчисленията**.

**Непризнаване на използването на инструменти на ИИ**, когато те са били използвани като източник за информация.

**Подправянето на линкове** се случва, когато текст, генериран от изкуствен интелект, се представя за текст на автора и се предоставя фалшив линк.

**Учениците трябва да демонстрират собствените си знания, умения и разбиране, необходими за тази съответната задача.**

**Нарушаване на процедурите и правилата, приети от училището, относно правилното и неправилното използване на инструментите на ИИ.**

# ПОТЕНЦИАЛНИ ПОКАЗАТЕЛИ ЗА ИЗПОЛЗВАНЕ НА ИИ:

- »» Използване по подразбиране на **език или лексика**, които не съответстват на нивото на уменията на ученика.
- »» Без **преки цитати и/или препратки**, когато това е необходимо или очаквано.
- »» Включване на **връзки, които не могат да бъдат намерени или проверени** (някои инструменти за изкуствен интелект предоставят фалшиви линкове към книги или статии от истински автори).
- »» **Липса на препратки** към събития, настъпили след определена дата (която отразява времето на създаване на източника на данни на инструмента за изкуствен интелект), което може да е показателно за някои дисциплини.
- »» Случаи на **неправилно/непоследователно използване на първо и трето лице**.
- »» **Използването на много подточки**. Някои (не всички) асистенти използват форматиране на текста в подточки, за да комуникират по сбит и ясен начин.
- »» **Липса на графики/таблици с данни/визуални средства** там, където те обикновено се очакват.
- »» Съдържанието е общо по своя характер (липсва личен опит) и **не е специфично за даден ученик** или за конкретна задача, както се изисква или очаква.
- »» **Учениците по невнимание включват предупреждения или уговорки**, направени от изкуствения интелект, за да подчертаят ограниченията на неговите възможности или хипотетичния характер на неговите резултати.
- »» Включване на очевидно **нелогични** или очевидно неверни твърдения в иначе подходящо съдържание.
- »» Прекалено **многословни или хиперболистични** изказвания, които може да не съответстват на обичайния стил на ученика, са неприемливи.



## ЦИТИРАНЕ НА ТЕКСТ, **генериран от ИИ**

### ▶ Пример

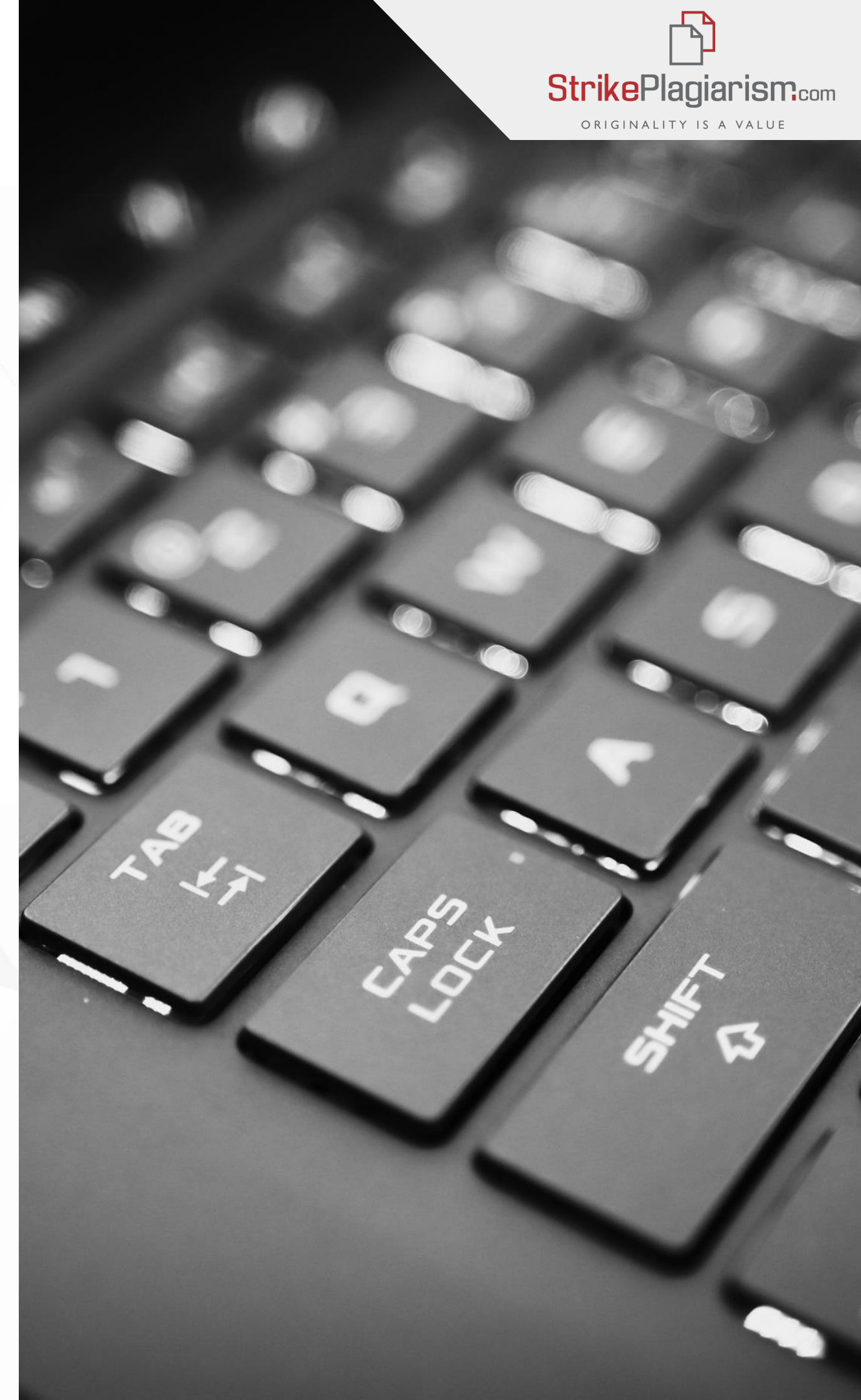
#### **ChatGPT 3.5 (<https://openai.com/blog/chatgpt/>), 25/01/2023.**

Ученикът трябва да запази копие от въпроса/въпросите и компютърно генерираното съдържание за справка и удостоверяване на автентичността във формат, който не може да бъде изтрит (например като снимка на екрана), за да бъде приложен към писмената работа.

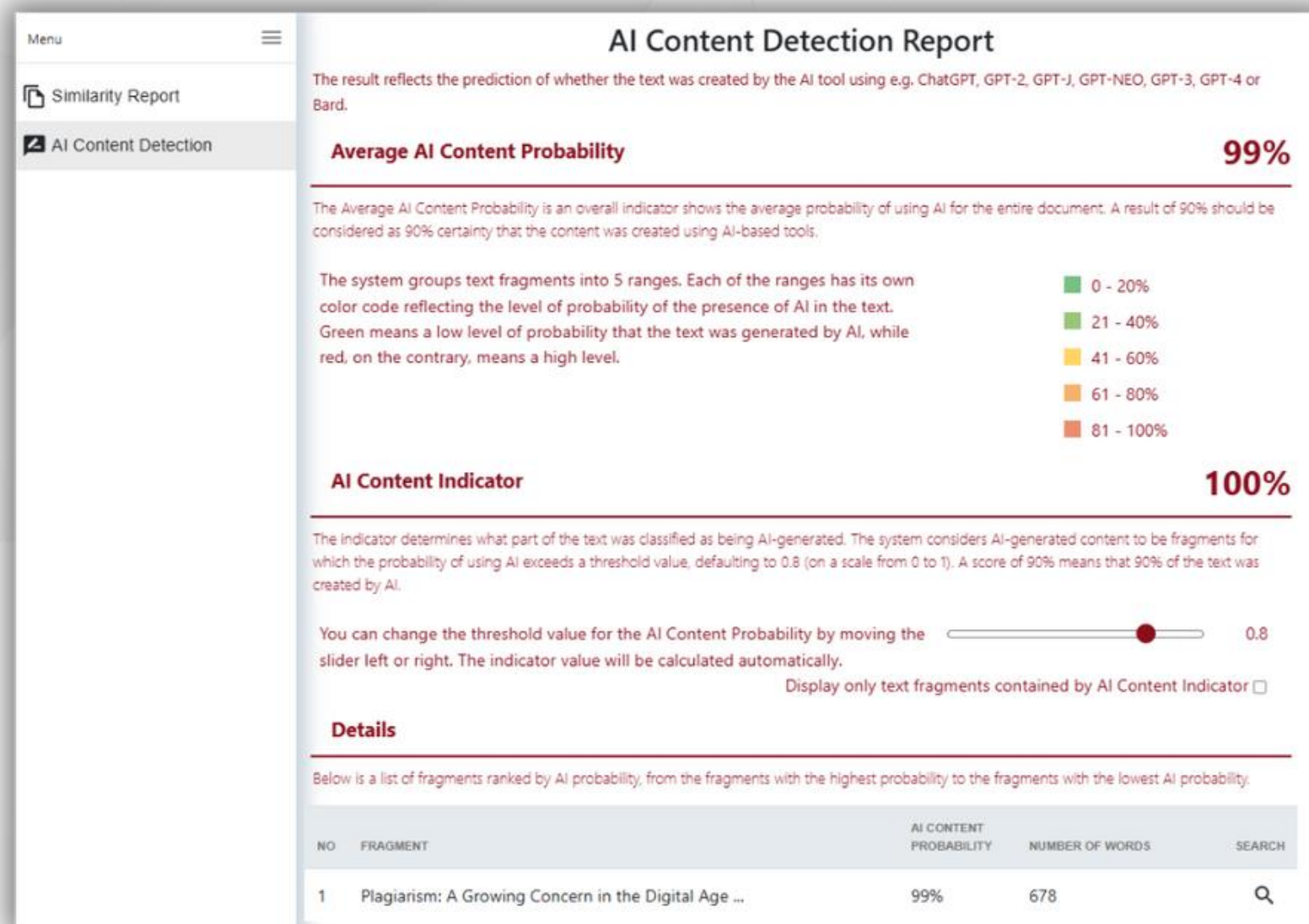
### ▶ **Други действия, които трябва да се обмислят във връзка с приемането на използването на ИИ, включват следното:**

Напомняне на учениците, че неправилното цитиране, перифразиране или копиране на текст е лоша практика, която може да доведе до сериозни санкции, включително дисквалифициране на работата.

В контекста на използването на изкуствен интелект учениците трябва ясно да разбират какво е приемливо и какво не, когато става въпрос за признаване на подобно съдържание. Например не е допустимо просто да се споменава „ИИ“ или „ChatGPT“, както не е допустимо да се посочва „Google“, вместо да се посочват конкретният сайт и уебстраници, до които е осъществен достъп.



# КАК РАБОТИ ОТКРИВАНЕТО НА СЪДЪРЖАНИЕ, ГЕНЕРИРАНО ОТ ИИ



**AI Content Detection Report**

The result reflects the prediction of whether the text was created by the AI tool using e.g. ChatGPT, GPT-2, GPT-J, GPT-NEO, GPT-3, GPT-4 or Bard.

**Average AI Content Probability** **99%**

The Average AI Content Probability is an overall indicator shows the average probability of using AI for the entire document. A result of 90% should be considered as 90% certainty that the content was created using AI-based tools.

The system groups text fragments into 5 ranges. Each of the ranges has its own color code reflecting the level of probability of the presence of AI in the text. Green means a low level of probability that the text was generated by AI, while red, on the contrary, means a high level.

- 0 - 20%
- 21 - 40%
- 41 - 60%
- 61 - 80%
- 81 - 100%

**AI Content Indicator** **100%**

The indicator determines what part of the text was classified as being AI-generated. The system considers AI-generated content to be fragments for which the probability of using AI exceeds a threshold value, defaulting to 0.8 (on a scale from 0 to 1). A score of 90% means that 90% of the text was created by AI.

You can change the threshold value for the AI Content Probability by moving the slider left or right. The indicator value will be calculated automatically.  Display only text fragments contained by AI Content Indicator

**Details**

Below is a list of fragments ranked by AI probability, from the fragments with the highest probability to the fragments with the lowest AI probability.

NO	FRAGMENT	AI CONTENT PROBABILITY	NUMBER OF WORDS	SEARCH
1	Plagiarism: A Growing Concern in the Digital Age ...	99%	678	<input type="text"/>

Никоя система за откриване на изкуствен интелект не може винаги да гарантира **100%** точност.

## ▶ Лингвистичен анализ

Тъй като съдържанието, създадено от ИИ, се генерира въз основа на шаблони, не е изненадващо, че в него можете да забележите повтарящи се фрази, странен синтаксис или липса на нюанси, характерни за човешкото писане.

## ▶ Статистически анализ

Много платформи използват статистически модели за оценка на сложността на текста, структурата на изреченията и употребата на лексика, за да определят дали текстът е написан от изкуствен интелект. Текстовете, създадени от ИИ, често имат еднаква структура и дължина на изреченията, което ги отличава от човешките текстове.

## ▶ Модели за машинно обучение

Моделите за машинно обучение (например Originality.AI) са обучени да правят разлика между текстове, написани от хора, и текстове, написани от изкуствен интелект. Те използват различни характеристики на текста, включително гореспоменатите методи, за да се научат да идентифицират характерни особености, които показват наличие на съдържание, генерирано от изкуствен интелект.



# КАК ДА ТЪЛКУВАМЕ ДОКЛАДА ЗА ОТКРИВАНЕ НА СЪДЪРЖАНИЕ, ГЕНЕРИРАНО ОТ ИЗКУСТВЕН ИНТЕЛЕКТ?



Нашият доклад показва както коефициента на вероятност за ИИ (AI Content Probability), така и вероятността за ИИ за всяка част от текста, като оцветява частите в различни цветове.



Всеки цвят представлява определен процент от вероятността за това дали текстът е написан от изкуствен интелект или от човек. Докладът показва списък с фрагменти и вероятността за съдържание, генерирано от ИИ за всеки фрагмент.

## AI Content Detection Report

The result reflects the prediction of whether the text was created by the AI tool using e.g. ChatGPT, GPT-2, GPT-J, GPT-NEO, GPT-3, GPT-4 or Bard.

**Average AI Content Probability**

**96%**

The Average AI Content Probability is an overall indicator shows the average probability of using AI for the entire document. A result of 90% should be considered as 90% certainty that the content was created using AI-based tools.

The system groups text fragments into 5 ranges. Each of the ranges has its own color code reflecting the level of probability of the presence of AI in the text. Green means a low level of probability that the text was generated by AI, while red, on the contrary, means a high level.

- 0 - 20%
- 21 - 40%
- 41 - 60%
- 61 - 80%
- 81 - 100%

Menu

Similarity Report

AI Content Detection

## AI Content Detection Report

The result reflects the prediction of whether the text was created by the AI tool using e.g. ChatGPT, GPT-2, GPT-J, GPT-NEO, GPT-3, GPT-4 or Bard.

**Average AI Content Probability** **99%**

The Average AI Content Probability is an overall indicator shows the average probability of using AI for the entire document. A result of 90% should be considered as 90% certainty that the content was created using AI-based tools.

The system groups text fragments into 5 ranges. Each of the ranges has its own color code reflecting the level of probability of the presence of AI in the text. Green means a low level of probability that the text was generated by AI, while red, on the contrary, means a high level.

- 0 - 20%
- 21 - 40%
- 41 - 60%
- 61 - 80%
- 81 - 100%

**AI Content Indicator** **100%**

The indicator determines what part of the text was classified as being AI-generated. The system considers AI-generated content to be fragments for which the probability of using AI exceeds a threshold value, defaulting to 0.8 (on a scale from 0 to 1). A score of 90% means that 90% of the text was created by AI.

You can change the threshold value for the AI Content Probability by moving the slider left or right. The indicator value will be calculated automatically.

Display only text fragments contained by AI Content Indicator

**Details**

Below is a list of fragments ranked by AI probability, from the fragments with the highest probability to the fragments with the lowest AI probability.

NO	FRAGMENT	AI CONTENT PROBABILITY	NUMBER OF WORDS	SEARCH
1	Plagiarism: A Growing Concern in the Digital Age ...	99%	678	<input type="text"/>

**Plagiarism: A Growing Concern in the Digital Age**

Plagiarism is a serious issue that has gained prominence in recent years, particularly with the rise of digital technology and the increasing accessibility of information. It involves the act of using someone else's work, ideas, or expressions without proper acknowledgment, presenting them as one's own. In the context of academic, professional, and creative fields, plagiarism undermines the value of original work, distorts intellectual property rights, and can lead to severe legal consequences.

**The Different Forms of Plagiarism**

Plagiarism can take various forms, ranging from direct copying of text to more subtle types of intellectual theft:

**Direct Plagiarism:** This is the most blatant form, where an individual copies someone else's work word-for-word without citation or permission.

**Paraphrasing Plagiarism:** In this case, the individual rephrases someone else's ideas or work but still fails to give proper credit to the original source.

**Self-Plagiarism:** This occurs when an individual reuses their own previously published work or research without acknowledging that it has been used before.

**Mosaic Plagiarism:** This involves blending someone else's work with one's own without appropriate citation, often creating a patchwork of plagiarized content.

**Accidental Plagiarism:** Sometimes, individuals may commit plagiarism unintentionally, usually due to ignorance or poor knowledge of citation rules.

**Causes of Plagiarism**

The rise of plagiarism can be attributed to several factors:

**Easy Access to Information:** The internet has made it easier for individuals to copy and paste content from websites, articles, or research papers. With a simple search, vast amounts of information are readily available.

**Pressure to Succeed:** In academic and professional environments, there is often immense pressure to produce high-quality work quickly. Some individuals resort to plagiarism as a shortcut to meet deadlines or to achieve success.

Ако текстът е зелен, вероятността да е написан от т.нар. езиков модел е минимална, а ако е червен – максимална.

Тези цветове не могат да се променят ръчно, да се приемат или отхвърлят чрез промяна на стойностите. Вероятността текстът да е написан от ИИ се проверява чрез модули и алгоритми, които са най-добрите в момента.





## Предимствата на нашия модул за търсене на съдържание, генерирано от изкуствен интелект

- »»» Голям списък на езиците (пълният списък на езиците можете да получите, като се свържете със отдела за поддръжка).
- »»» Скорост на проверката.
- »»» Точността на откриване е над 94%.
- »»» Проверка на големи документи.
- »»» Процент на вероятност не само за целия текст, но и за всеки фрагмент, за да се определи най-засегнатата част.
- »»» Прагова стойност за индикатора ИИ, показваща най-засегнатите части.
- »»» Пет диапазона на вероятност с цветни кодове, като колкото по-тъмен е текстът, толкова по-голяма е вероятността за ИИ.
- »»» Отделен доклад, показващ резултатите от проверката за ИИ.



# КРИТЕРИИ ЗА ОЦЕНКА НА СЪДЪРЖАНИЕ, ГЕНЕРИРАНО ОТ ИИ

- Фрагменти от произведения, които не са написани от автора, трябва да бъдат правилно цитирани и да се добави позоваване на ИИ.
- Пример: ChatGPT 3.5 (<https://openai.com/blog/chatgpt/>), 25.01.2023 г.
- Обърнете внимание на фрагментите, които надхвърлят 80% вероятност, изяснете с автора източника на тези фрагменти.
- Дайте възможност на ученика, да коригира работата, ако няма сериозно нарушение на образователната етика.

# КАКВО ИСКАМЕ?

Критично мислене и умения за **проверка на факти.**

Развитие на **креативността.**

Умение да отразяват **собствените идеи и мисли.**

Да **перифразираш** -> да обогатиш лексиката си.

Да **намираш информация** -> да обогатиш знанията си.

Да **обогатиш знанията си** -> да развиеш умения за критично и аналитично мислене

**Да отразяваш чужди идеи = да създаваш свои.**



## **ПРОМЕНЕТЕ ПОДХОДА СИ КЪМ ПИСМЕНИТЕ РАБОТИ НА УЧЕНИЦИТЕ:**

- ▶ Избягвайте опростени въпроси, основани на факти.
- ▶ Използвайте конкретни казуси и примери, специфични за определена група ученици.
- ▶ Разделете големите задачи на по-малки стъпки.
- ▶ Изисквайте мета или рефлексивни изказвания.
- ▶ Използвайте източници, до които изкуственият интелект няма достъп.
- ▶ Предложете на учениците да споделят личен опит.
- ▶ Използвайте изображения в насоките, свързани със задачата.
- ▶ Приложете алтернативни задачи.



**БЛАГОДАРЯ ЗА ВНИМАНИЕТО!**

**Доника Гавова**

[dgavova@strikeplagiarism.com](mailto:dgavova@strikeplagiarism.com)