

Злоупотребление ИИ и академическая честность

Юлия Кальчун-Набиева



StrikePlagiarism.com

ORIGINALITY IS A VALUE

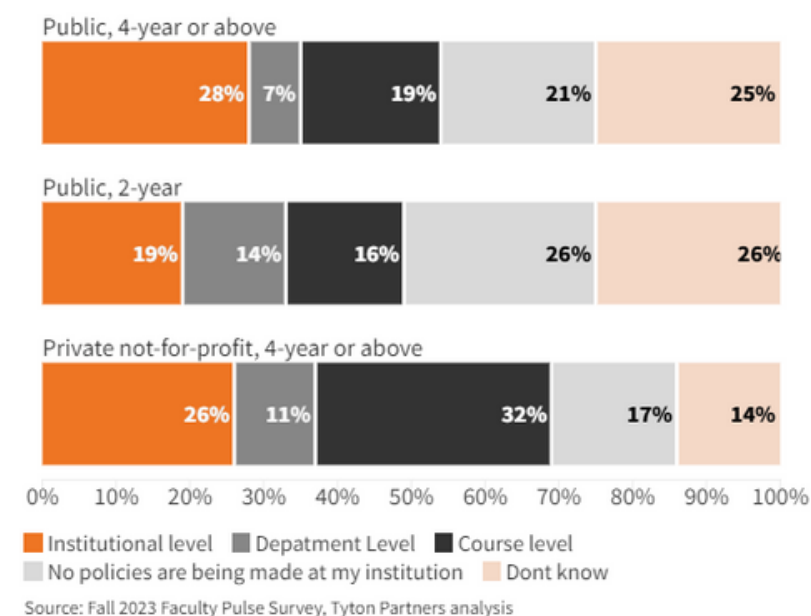


СТАРАЯ **ПРОБЛЕМА** – НОВЫЕ МАСШТАБЫ

- ▶ «Копировать – вставить» популярный метод написания работ
“Писатель-призрак” – Ghost writer и
- ▶ продажа готовых студенческих и научных работ через Интернет
- ▶ Расцвет **Thesis Mills** или **Essay Mills** – фирмы пишущие работы на заказ
- ▶ **“Коучинг”** за деньги
- ▶ **ИИ-плагиат** – новый способ обмана

AI policy-making

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



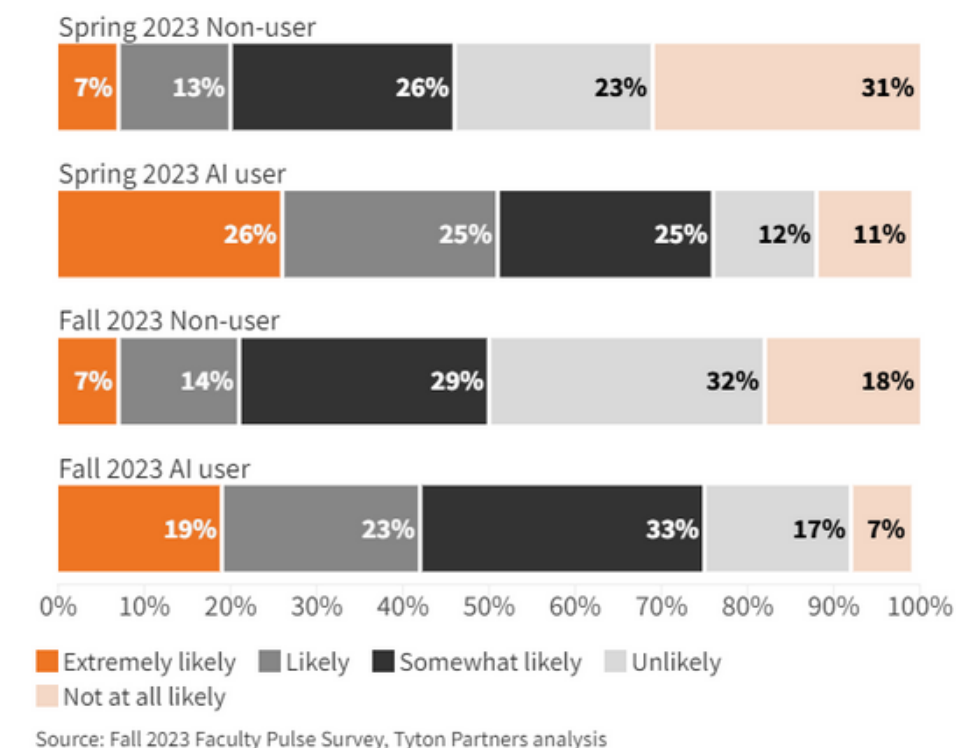
About a third (32 percent) of private, four-year institutions rely on instructors to set their own policy, although more than a quarter (26 percent) have institutionwide policies. Public four-year institutions are split on policies: 28 percent say policies at the institution have been set, 25 percent 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.

Исследование, проведенное Tyton Partners analysis, подтвердило **положительную динамику использования ИИ** студентами США. Более 70% студентов заявляют об активном использовании ИИ.

Using AI tools

Student likelihood to continue to use GenAI writing tools despite restrictions



Большая часть университетов **внедряет правила и практику использования ИИ**, как на уровне университета, так и на уровне факультетов.

ИИ в академическом процессе:

ПРОБЛЕМЫ:

Может выдать **фальшивую, неправдивую** информацию,

Не позволяет развивать **навыки поиска информации**, сравнения,

Формирует **привыкание**,

Может привести к **нарушению академической честности**,

Не имеет доступ к **самым новым данным**.

ВОЗМОЖНОСТИ:

Экономит время на **поиске** информации

Экономит время на **анализе**

Дает **развернутый ответ** из различных источников.

Что такое злоупотребление ИИ?

Примеры неправомерного использования ИИ включают, но не ограничиваются следующим:

Копирование или перефразирование разделов контента, созданного ИИ, так что работа больше **не является самостоятельной** работой студента,

Копирование или перефразирование целых ответов, сгенерированных ИИ **без ссылки на ИИ**,

Использование ИИ для выполнения частей работы, так что работа не отражает **самостоятельность** в анализе, оценке или расчетах,

Непризнание использования инструментов ИИ, когда они были использованы в качестве источника информации,

Фальсификация ссылок, когда текст ИИ выдают за авторский и подают **фейковую ссылку**,

Студенты должны продемонстрировать **СВОИ** собственные знания, навыки и понимание, **требуемые для данной квалификации**,

Отход от процедур и правил принятых университетом, о том **как правильно и неправильно** пользоваться инструментами ИИ.

Цитирование текста, сгенерированного ИИ

► Пример цитирования текста ИИ:

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

Студент должен сохранить копию вопроса(ов) и сгенерированного компьютером контента для справки и проверки подлинности в нестираемом формате (например, в виде скриншота) – приложить к работе.

► Другие действия, которые следует рассмотреть в связи с признанием использования ИИ, следующие:

Напоминание студентам о том, что, плохое цитирование, перефразирование копирование фрагментов текста может быть недобросовестной практикой, которая может повлечь за собой серьезные санкции, включая дисквалификацию работы.

В контексте использования ИИ студенты должны четко понимать, что является и что не является приемлемым в отношении признания содержания ИИ и использования источников ИИ. Например, недопустимо просто ссылаться на "ИИ" или "ChatGPT", так же как недопустимо указывать "Google", а не конкретный сайт и веб-страницы, к которым обращались;

Критерии оценки студенческих и научных работ на наличие контента, сгенерированного ИИ.

В соответствии с законодательством и международной практикой каждая работа, написанная студентом либо преподавателем **должна быть самостоятельной**, тем самым навыки и знания используемые автором при написании текста **не должны быть поставлены под сомнение**.

1. Тем самым каждая работа должна быть проверена на наличие контента сгенерированного ИИ.
2. **Фрагменты** работы, **написанные не автором, должны** быть правильно **оформлены** в виде **ссылок**.

Следуя этому правилу, если в работе имеются фрагменты, созданные инструментами ИИ, они должны быть оформлены в соответствии с международными стандартами.

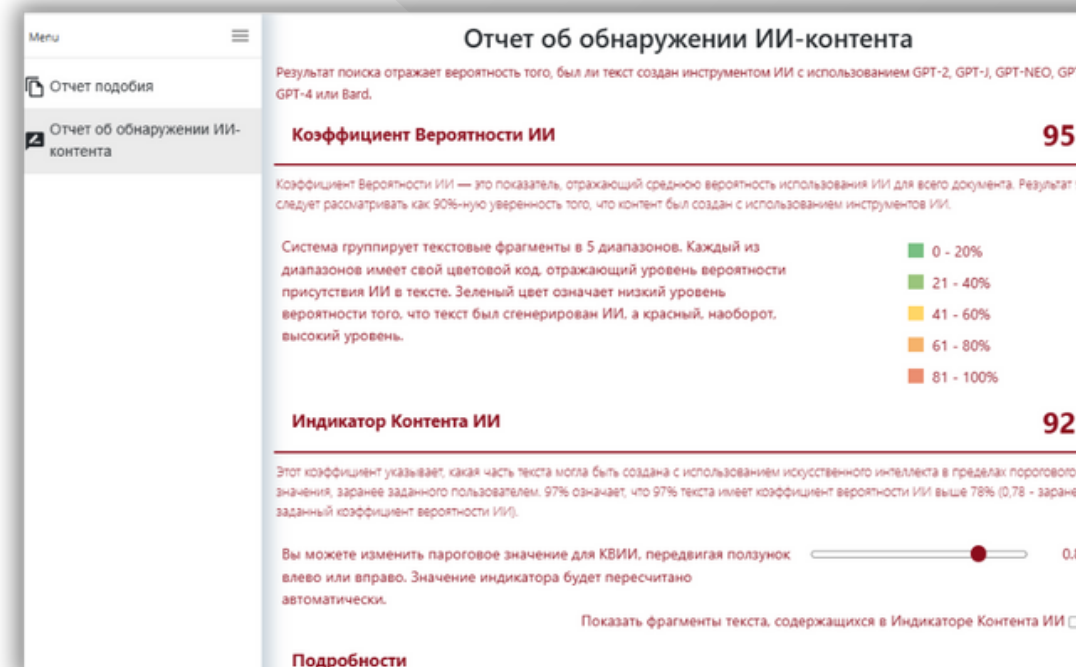
Пример:

Там, где инструменты искусственного интеллекта использовались в качестве источника информации, подтверждение использования инструментов ИИ должно содержать **имя** используемого источника и даты создания контента. Например: ChatGPT 3.5 (<https://openai.com/блог/chatgpt/>), 25.01.2023.

Потенциальные индикаторы использования ИИ

- » Использование по умолчанию **языка или лексики, не соответствующей уровню квалификации,**
- » **Отсутствие прямых цитат и/или использование ссылок** там, где они требуются или ожидаются,
- » **Включение ссылок, которые невозможно найти** или проверить (некоторые инструменты ИИ предоставляют ложные ссылки на книги или статьи реальных авторов)
- » **Отсутствие ссылок на события, произошедшие после определенной даты** (отражающей время составления источника данных инструмента ИИ), что может быть примечательно для некоторых предметов,
- » **Случаи неправильного/непоследовательного использования перспективы от первого и третьего лица,** когда сгенерированный текст остается неизменным,
- » **Разница в используемом стиле языка** по сравнению со стилем, используемым **студентом в классе** или в других ранее представленных работах,
- » **Отсутствие графиков/таблиц данных/визуальных средств** там, где они обычно ожидаются,
- » Содержание носит более **общий характер**, а не относится к самому студенту или к специализированному заданию или сценарию, если это требуется или ожидается
- » **Непреднамеренное включение студентами предупреждений или оговорок, сделанных ИИ,** чтобы подчеркнуть пределы его возможностей или гипотетический характер его результатов,
- » Включение сильно выраженных **нелогичных или уверенно неверных утверждений** в ином целостном контенте,
- » **Чрезмерно многословные или гиперболические высказывания,** которые могут не соответствовать обычному стилю кандидата.

Как работает ИИ-обнаружение?



Лингвистический анализ

Поскольку контент, созданный ИИ, генерируется на основе шаблонов, неудивительно, что в нем можно заметить повторяющиеся фразы, странный синтаксис или отсутствие нюансов, характерных для человеческого письма.

Статистический анализ

Многие платформы используют статистические модели для оценки сложности текста, структуры предложений и использования лексики, чтобы определить, был ли текст написан ИИ. Тексты, созданные ИИ, часто имеют однородную структуру предложений и длину, что отличает их от человеческого письма.

Модели машинного обучения

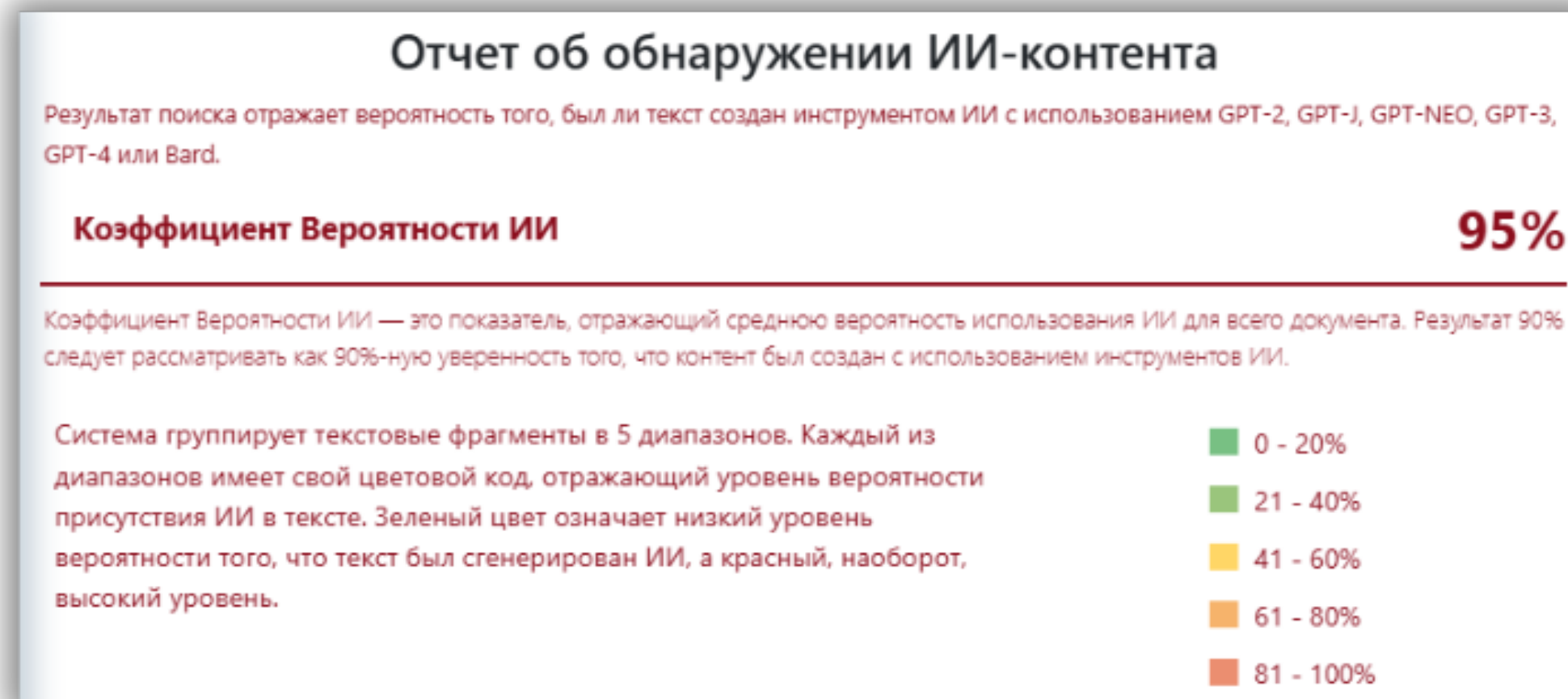
Модели машинного обучения (например, Originality.AI) обучаются различать тексты, созданные человеком и ИИ. Они используют различные признаки текста, включая вышеупомянутые методы, чтобы научиться выявлять характерные особенности, свидетельствующие о написании ИИ.

Ни одна система для обнаружения ИИ не может гарантировать 100% точность всегда.

Как интерпретировать отчет об обнаружении ИИ-контента?

»»» Наш отчет отражает как Коэффициент Вероятности ИИ (КВИИ), так и вероятность ИИ для каждого фрагмента текста, закрашивая фрагменты в разные цвета.

»»» Каждый цвет отражает определенную долю вероятности того, написан текст ИИ или человеком. Отчет показывает список фрагментов и Коэффициент Вероятности ИИ для каждого фрагмента.



Отчет об обнаружении ИИ-контента

Результат поиска отражает вероятность того, был ли текст создан инструментом ИИ с использованием GPT-2, GPT-J, GPT-NEO, GPT-3, GPT-4 или Bard.

Коэффициент Вероятности ИИ

95%

Коэффициент Вероятности ИИ — это показатель, отражающий среднюю вероятность использования ИИ для всего документа. Результат 90% следует рассматривать как 90%-ную уверенность того, что контент был создан с использованием инструментов ИИ.

Система группирует текстовые фрагменты в 5 диапазонов. Каждый из диапазонов имеет свой цветовой код, отражающий уровень вероятности присутствия ИИ в тексте. Зеленый цвет означает низкий уровень вероятности того, что текст был сгенерирован ИИ, а красный, наоборот, высокий уровень.

0 - 20%

21 - 40%

41 - 60%

61 - 80%

81 - 100%

Индикатор Контента ИИ

92%

Этот коэффициент указывает, какая часть текста могла быть создана с использованием искусственного интеллекта в пределах порогового значения, заранее заданного пользователем. 97% означает, что 97% текста имеет коэффициент вероятности ИИ выше 78% (0,78 - заранее заданный коэффициент вероятности ИИ).

Вы можете изменить пороговое значение для КВИИ, передвигая ползунок влево или вправо. Значение индикатора будет пересчитано автоматически.

0.8

Показать фрагменты текста, содержащихся в Индикаторе Контента ИИ

Подробности

Нижне представлен список фрагментов, ранжированных по вероятности ИИ, от фрагментов с самой высокой вероятностью до фрагментов с самой низкой вероятностью ИИ.

ПОРЯДКОВЫЙ НОМЕР	ФРАГМЕНТ	ВЕРОЯТНОСТЬ КОНТЕНТА ИИ	ЧИСЛО СЛОВ	ПОИСК
1	В Восточном Казахстане можно исследовать националь...	96%	268	🔍

Отдых в Казахстане — это уникальная возможность насладиться живописными пейзажами, познакомиться с древней культурой и ощутить гостеприимство местных жителей. Страна, известная своими широкими степями, величественными горными цепями и озёрами, предлагает туристам разнообразие активных и познавательных развлечений.

1. Природные чудеса
Казахстан — это место, где встречаются горы, степи, озёра и пустыни. На юге страны раскинулись величественные Тянь-Шань и Алтай, привлекающие альпинистов и туристов, желающих полюбоваться альпийскими лугами и живописными озёрами, такими как озеро Большое Алматинское или озеро Сайрам-су. В Восточном Казахстане можно исследовать национальные парки, такие как Катон-Карагай, с его первозданной природой.

2. Культурные и исторические достопримечательности
Казахстан богат историческими памятниками. Алматы, бывшая столица, известна своими культурными объектами, включая музеи, театры и архитектурные памятники. Не менее интересен и Нур-Султан (Астана) с его современными зданиями и музеями, такими как Национальный музей Казахстана и Байтерек — символ столицы.

Любители истории могут посетить исторические комплексы, такие как мавзолей Ходжи Ахмеда Ясави в Туркестане, или ознакомиться с памятниками древности в районе Сарыарка, которые внесены в список Всемирного наследия ЮНЕСКО.

3. Активный отдых
Казахстан — отличное место для активного отдыха на свежем воздухе. Туристы могут заниматься пешими походами, альпинизмом, горными и водными видами спорта, а также катанием на лошадях, что является важной частью местной культуры. Для любителей экстремальных видов спорта открыты возможности для парашютного спорта и каякинга в степях и горах.

4. Гостеприимство и кухня
Казахская кухня представляет собой удивительное сочетание вкусов и традиций. Одним из самых популярных блюд является казы — казахская колбаса, а также бешбармак — мясное блюдо с лапшой. Местные рестораны и кафе предлагают не только национальную кухню, но и современные интерпретации блюд, которые порадуют даже самых требовательных гурманов.

Кроме того, казахи славятся своим гостеприимством, и туристы часто ощущают себя как дома, независимо от того, в каком уголке страны они находятся.

Если текст зеленого цвета, то вероятность того, что он написан машиной минимальна, если красного цвета – максимальна.

Эти цвета невозможно изменить вручную, принять или отклонить их, изменяя КВИИ.

Вероятность того, что текст написан машиной проверяется модулями и алгоритмами, которые являются лучшими на данный момент.

Преимущества нашего модуля поиска контента ИИ

- **Большое количество языков** (узнайте весь список в службе поддержки),
- **Скорость** проверки,
- Проверка документов **большого размера**, т.е. до 1000 стр.,
- Отражение **процента вероятности** не только для всего текста, но **для каждого фрагмента**, отражая самый подверженный ИИ текст.
- **Пороговое значение** для индикатора ИИ, что позволяет отразить фрагменты с **наивысшим процентом** вероятности ИИ,
- **Пять диапазонов вероятности**, со своим **цветовым кодом**, чем темнее текст, тем выше вероятность ИИ,
- **Отдельный отчет для ИИ**, отчет контента ИИ можно также скачать,
- Аккуратность и **точность** обнаружения превышает **95%**.

Измените **ПОДХОД** к студенческим письменным работам:

- ▶ Избегайте простых вопросов, основанных на фактах,
- ▶ Используйте случаи и примеры, специфические для конкретной группы студентов,
- ▶ Разбейте большие задачи на меньшие этапы,
- ▶ Попросите о мета- или рефлексивных высказываниях,
- ▶ Попросите поделиться личными примерами,
- ▶ Используйте источники, к которым ИИ не имеет доступа,
- ▶ Используйте изображения в подсказках,
- ▶ Используйте альтернативные задания.



Критерии оценки студенческих и научных работ на наличие контента, сгенерированного ИИ.

3. При анализе работы необходимо обратить внимание на **фрагменты**, у которых **Коэффициент Вероятности ИИ превышают 80%**. Необходимо уточнить у студента **источник этих фрагментов**.
4. В случае, если **студент забыл оформить ссылку**, оформил ссылку с нарушениями правил университета, необходимо **вернуть работу на исправление**.
5. Если окажется, что **результат** поиска **контента ИИ** является неточным и **не соответствует истине**, достаточно добавить **обоснование принятия работы преподавателем** по причине **отсутствия нарушения** принципов академической честности.

ЧТО МЫ ХОТИМ?

Навыки анализа и проверки **фактов**,

Развитие **креативности**,

Умение отображать **свои идеи** и **мысли**,

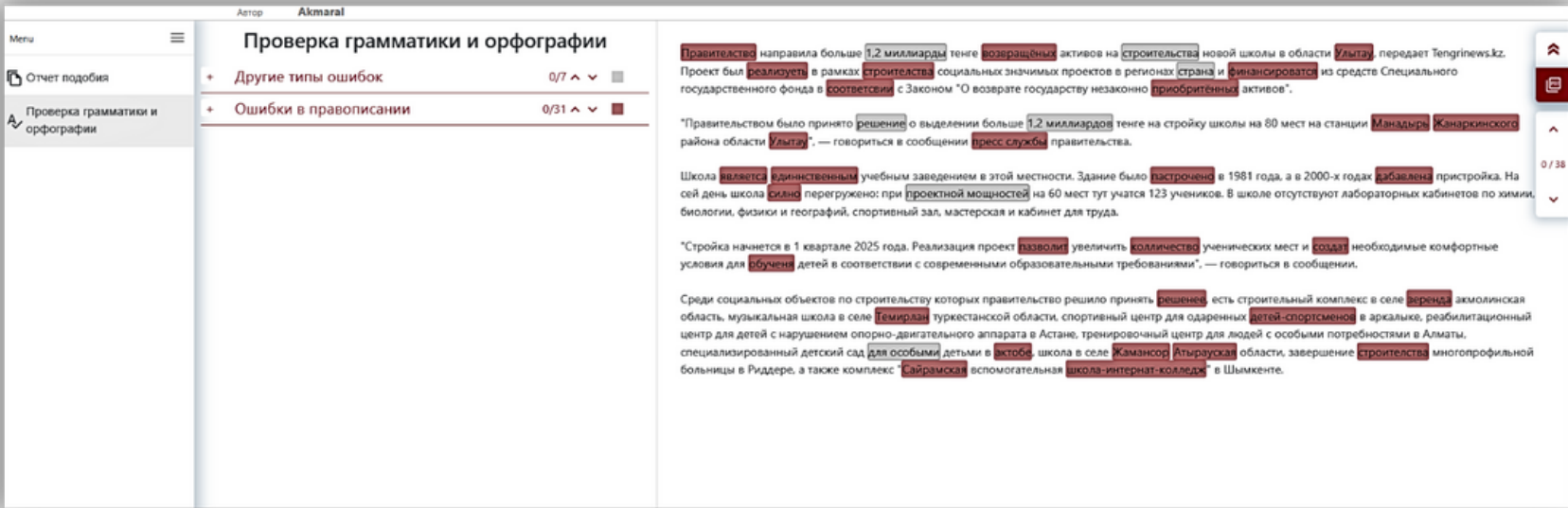
Отображать чьи-то идеи = обогащать
свои,

Перефразировать = обогатить
словарный запас,

Поиск данных = обогащение своих
знаний,

Быть более **эффективным на рынке
труда** = обогатить свое сообщество!

Дополнительный функционал StrikePlagiarism.com



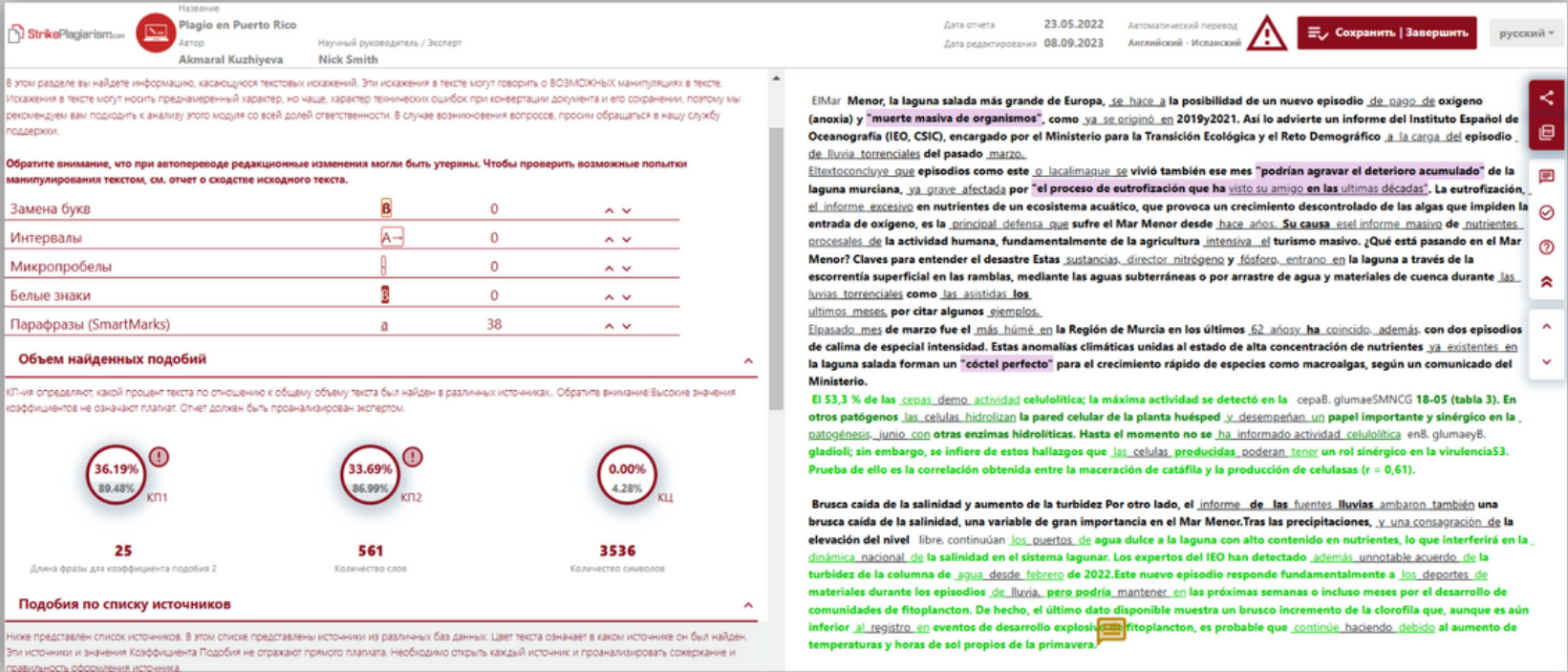
Проверка грамматики:

- Поддерживает более 30 языков,
- Отображает подсказки для исправлений,
- Простая навигация по тексту,
- Отображает ошибки по категориям, где каждая категория имеет свой цветовой код.

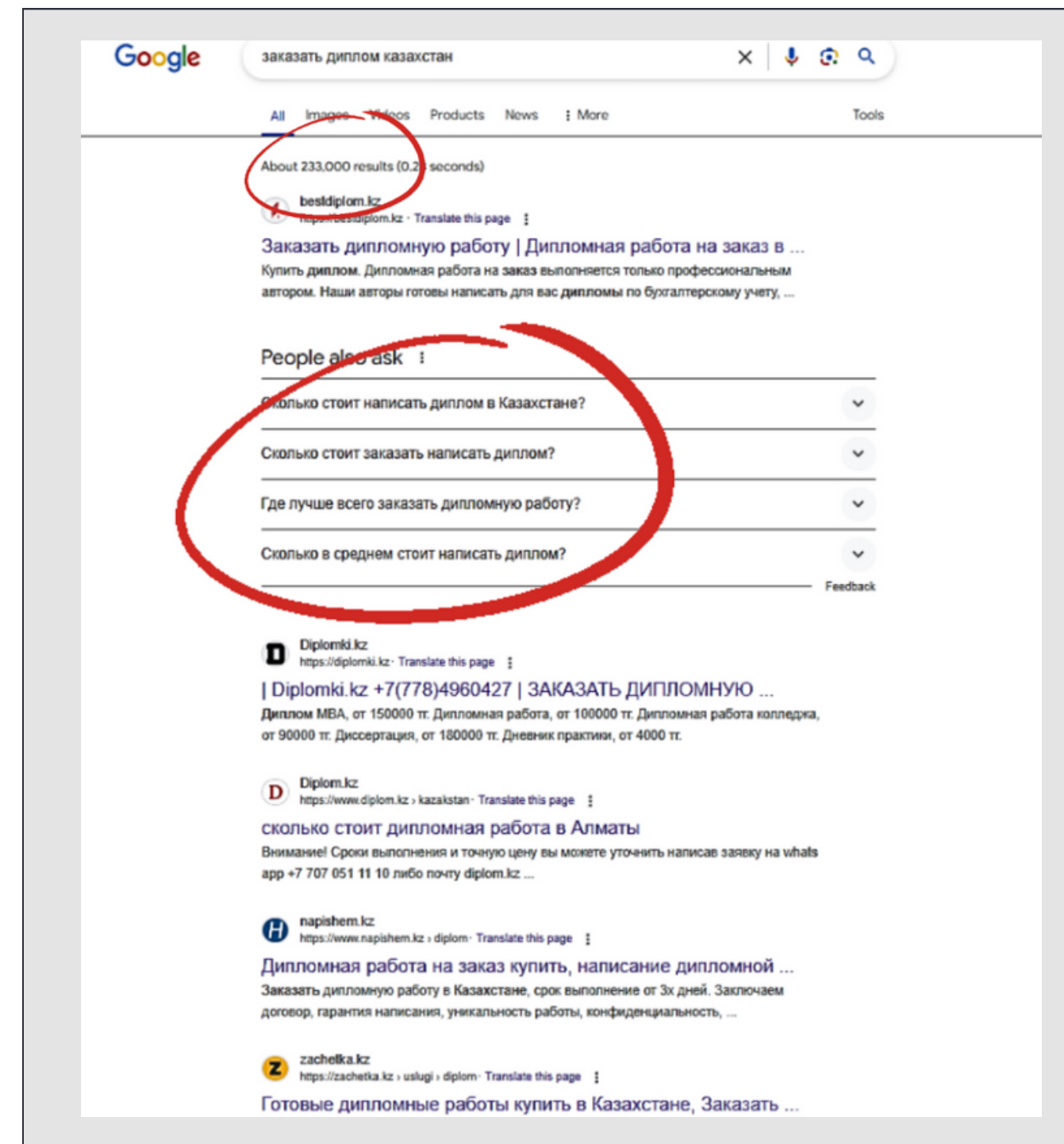
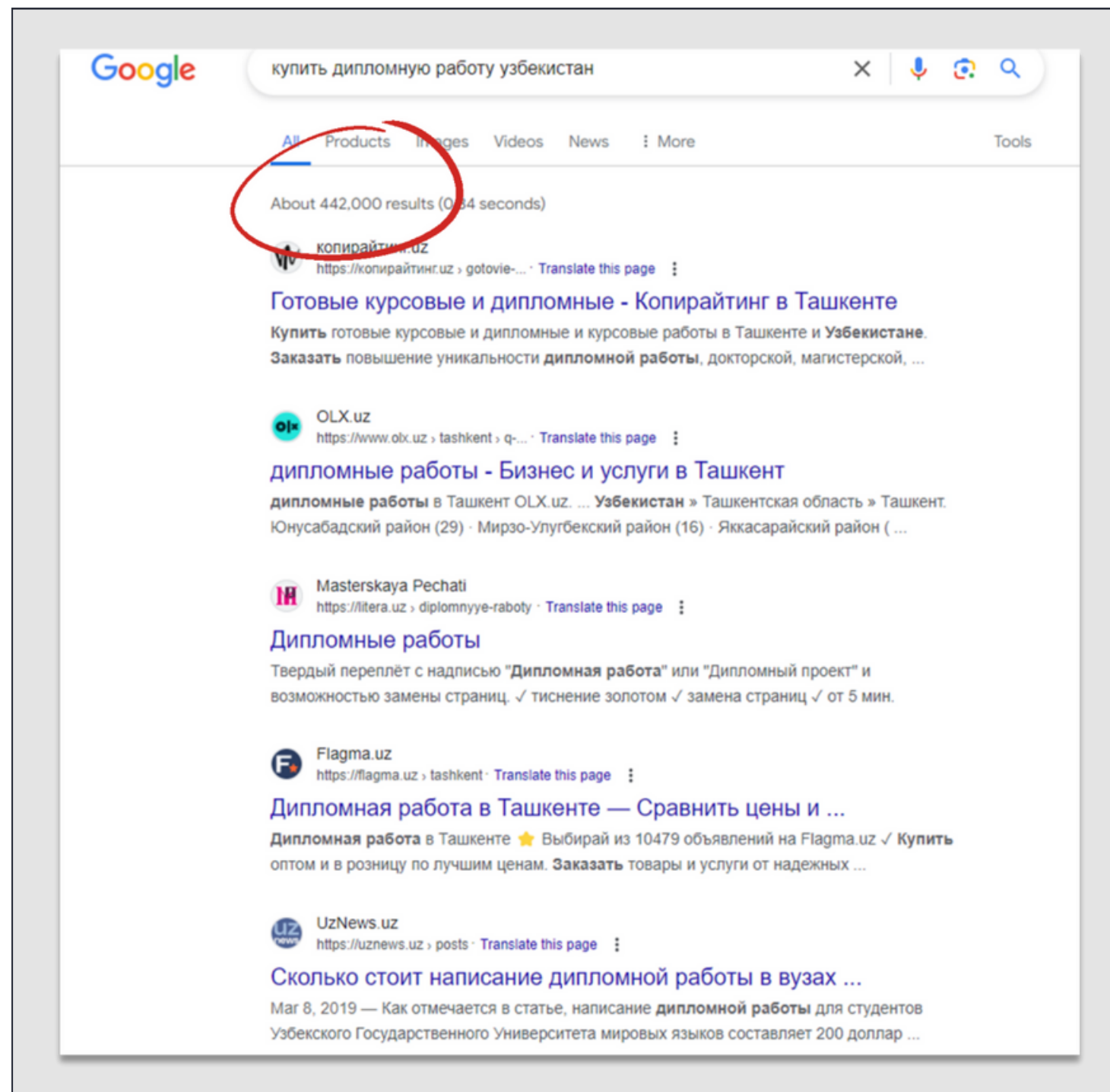
Поиск переводного плагиата:

- Поддерживает более 100 языков,
- Поддерживает большие файлы,
- Высокая скорость.

Грамматическая проверка, между тем, позволяет обнаружить **обход системы антиплагиата**. Например, автор добавляет ошибку, чтобы нарушить порядок фраз, что не позволяет системе найти сходства.



Мельницы диссертаций все еще серьезная проблема для высшего образования



The background features a large red triangle in the top right corner and a grey triangle in the bottom right corner, meeting at a diagonal line.

Спасибо за внимание

contact@strikeplagiarism.com

Kazakhstan@strikeplagiarism.com