ДИГИТАЛНИ ОБРАЗОВАТЕЛНИ ТЕХНОЛОГИИ
Издание на Великотърновския университет „Св. св. Кирил и Методий“
2/2026

DOI: 10.54664/PDUJ4524

Венелина Цаневска[1]

УЯЗВИМОСТ НА ТРАДИЦИОННИТЕ ИЗПИТНИ ФОРМИ В ДИСТАНЦИОННОТО ОБУЧЕНИЕ В ЕРАТА НА ГОЛЕМИТЕ ЕЗИКОВИ МОДЕЛИ (LLMS)

Venelina Tsanevska

THE VULNERABILITY OF TRADITIONAL EXAMINATION FORMS IN DISTANCE LEARNING IN THE ERA OF LARGE LANGUAGE MODELS (LLMS)

Abstract: *This study examines the impact of generative artificial intelligence (GAI) — specifically text-generating Large Language Models (LLMs) — on assessment methods in a digital learning environment, illustrated by a specific case from the field of finance.*

The widespread adoption of GAI in the educational environment calls into question the validity of established forms of remote assessment. The paper analyzes the vulnerability of a multi-component written assignment in a finance course through a benchmark experiment: reference answers generated by three text-generating systems are compared with real student works in a distance learning format. The results reveal moderate to high levels of similarity with LLM-generated reference answers among a significant proportion of participants. This trend is characterized by distinct stylistic markers, structural similarities across assignments, and indications of possible AI reliance. The findings are discussed in terms of their implications for the reliability of remote assessment and teaching practice, and practical recommendations for adapting assessment to the changing technological environment are proposed.

Keywords: *large language models (LLMs), distance learning, examination forms, AI-resilient assessment, benchmark analysis, Bloom's taxonomy, AI-generated content, higher education.*

1. ВЪВЕДЕНИЕ

Актуалността на темата се обуславя от безпрецедентно бързото навлизане на генеративния изкуствен интелект в ежедневието, което за броени месеци го превърна в инструмент за масова, трудно контролируема употреба и от страна на студентите във висшите училища.

Появата на достъпни генеративни езикови модели след края на 2022 г. постави академичната общност пред предизвикателство, за което традиционните механизми на оценяване не са проектирани. Тази уязвимост е особено изразена в дистанционното обучение, тъй като при асинхронното онлайн оценяване контролът върху условията за изпитване е структурно невъзможен, а студентът разполага с неограничен достъп до технологични ресурси. Въпросът вече не е дали студентите използват технологии с генеративен изкуствен интелект (в частност големи езикови модели – LLMs), а в каква степен и с какви последствия за валидността на оценяването.

Настоящото изследване проверява уязвимостта на конкретно многокомпонентно писмено задание по финансова дисциплина в дистанционна форма на обучение чрез бенчмарк експеримент.

[1] *Венелина ЦАНЕВСКА* – главен асистент, доктор в катедра „Финанси и счетоводство“, Стопански факултет, ВТУ „Св. св. Кирил и Методий“, v.tsanevska@ts.uni-vt.bg

Публикацията представя резултатите от анализа, интерпретира идентифицираните маркери на технологично заимстване и предлага препоръки за адаптиране на оценяването в условията на генеративен изкуствен интелект.

2. ТЕОРЕТИЧНА РАМКА

2.1. Изкуствен интелект и LLM-базирани текстогенериращи системи

В съвременната научна литература изкуственият интелект (ИИ) се дефинира като софтуер, способен да открива причинно-следствени връзки в данни и да генерира прогнози, препоръки или решения на базата на непълна информация без предефинирани инструкции (МОН, 2024). Той често е описван като „дигитален мозък“, който имитира човешкото мислене и притежава способността непрекъснато да се усъвършенства чрез процеси на машинно обучение (МОН, 2024). Някои автори цитират Minski (Holmes, Bialik & Fadel, 2019: 194), който предлага класическо определение, според него изкуственият интелект е науката, която кара машините да правят неща, които биха изисквали интелигентност, ако бяха правени от хора[2].

Съвременните технологии с ИИ включват традиционни системи за анализ на данни и генеративни системи за създаване на ново съдържание, като към вторите се отнасят и текстогенериращите системи, базирани на големи езикови модели (LLMs).

Изкуственият интелект навлиза в образованието още в края на XX век под формата на системи за адаптивно обучение и автоматизирана проверка. На съвременния етап най-съществено значение за висшето образование придобиват генеративните технологии, базирани на големи езикови модели, които позволяват създаване на академично звучащ текст и поставят нови предизвикателства пред оценяването. Съществена промяна за висшето образование настъпва именно с тяхната поява. Това са системи, обучени върху огромни масиви от текстови данни, способни да генерират кохерентни и контекстуално релевантни текстове, да формулират аргументи и да възпроизвеждат характеристики на човешката езикова комуникация. Именно големите езикови модели имат пряко отношение към оценяването на писмени задания във висшето образование, тъй като техният основен продукт е структуриран текст – форматът, който традиционно се изисква от студентите при отворени въпроси и аналитични разработки. Поради това уязвимостта на писмените форми на оценяване е значително по-висока в сравнение с форми, изискващи устна защита, практическо изпълнение или лична аргументация.

В образователен контекст Holmes, Bialik и Fadel разграничават три измерения на ИИ – институционално, преподавателско и обучително, като настоящото изследване се фокусира върху обучителното измерение и по-конкретно върху използването на текстогенериращи системи от студентите в процеса на оценяване (Holmes, Bialik & Fadel, 2019).

Същевременно систематичният обзор на Zhai, Wibowo и Li предупреждава, че свръхзависимостта от ГИИ системи рискува да намали способностите на студентите за критично мислене, аналитично разсъждение и самостоятелно вземане на решения – именно умения, които висшето образование следва да изгражда (Zhai, Wibowo & Li, 2024: 28).

Теоретичната рамка, която обяснява защо студентите масово прибягват до ГИИ, е предложена от Davis (Davis, 1989: 319–340) чрез *Модела за приемане на технологии,* наричан в западната литература Technology Acceptance Model (TAM).

Моделът разграничава два основни фактора, които определят потребителското поведение при работа с дигитални системи (Davis, 1989: 319–320):

✓ възприета полезност – убеждението, че дадена система подобрява изпълнението на задачата;

✓ възприета лекота на употреба – убеждението, че взаимодействието с нея не изисква големи усилия.

Изследването потвърждава, че студентите са по-склонни да използват дадена технология, когато я намират за полезна, отколкото когато просто е лесна за употреба (Davis, 1989: 333). В

[2] Преводите на български език в статията са на авторката. – Бел. ред.

контекста на дистанционното обучение този механизъм се задейства максимално: LLM инструментите осигуряват бързи и полезни резултати при минимална инвестиция на усилия, неутрализирайки когнитивната бариера пред тяхното ползване. Следващи изследвания разширяват този извод, като установяват, че към полезността и лекотата на употреба се добавя и социалното влияние – студентите са по-склонни да използват ГИИ-инструменти и когато хората около тях правят същото (Raman, Achuthan & Nedungadi, 2026: 3). В контекста на българските висши училища тази динамика се наслагва върху продължаващия процес на дигитализация на образованието, при който технологичните инструменти навлизат по-бързо от институционалните рамки, регулиращи тяхната употреба. Слабата подготовка за използване на ИИ-решения не е само технологичен проблем – тя отразява и дефицити в знанието и обучението (Бянов, 2026: 711). Това важи в пълна сила и за академичната среда, където липсата на ясна регулация се очертава като катализатор за нерегламентираното ползване на ГИИ от студентите (Raman, Achuthan & Nedungadi, 2026: 4). Изследователи подчертават, че употребата на ГИИ в образованието поражда сериозни опасения относно академичната почтеност, надеждността на оценяването и развитието на критичното мислене у студентите (Noroozi, Soleimani, Farrokhnia & Banihashem, 2024: 374–375). Въпреки тези притеснения съвременните концепции за изкуствения интелект и неговите приложни форми вече са неразделна част от елементите, които се използват ежедневно, пораждайки коренно противоположни мнения сред критици и защитници. Съществено в новата парадигма е, че технологиите просто не могат да бъдат избегнати, поради което основната цел пред образователната система е те да бъдат управлявани и прилагани по начин, който максимално да увеличи техните позитивни ефекти върху учебния процес (Čolić, Sukić & Anđelković, 2025: 3714) .

Разграничението между персонализирани и генерични академични въпроси е ключово за разбирането на уязвимостта: персонализираните задачи са устойчиви срещу ГИИ, тъй като изискват контекстуална специфичност, непозната на системата – именно поради установената ѝ ограничена способност за контекстуално разбиране и персонализация (Baidoo-Anu & Ansah, 2023: 55).

2.2. ИИ-устойчиво оценяване и когнитивни равнища

В отговор на описаното предизвикателство се налага концепцията за т.нар. „*ИИ-устойчиво*“ *оценяване*. То е дефинирано като подход към оценяването, при който изпитните задачи запазват своята образователна стойност и не могат да бъдат изпълнени пълноценно без реалното участие на студента (Perkins, 2026: 4). Устойчивите оценителни задачи изискват процес или изпълнение, неразривно свързани с личното усилие на обучаемия, запазват автентичност чрез обвързване с реални приложения на знанието и включват механизми за верификация на реалното разбиране, а не само на „привидния“ верен отговор. Крайната цел е завършването на заданието да изисква активна когнитивна ангажираност – нещо, което не може да бъде делегирано на машина.

Концепцията за ИИ-устойчиво оценяване намира теоретична опора в **Таксономията на Блум** от 1956 г. (Anderson & Krathwohl, 2001), която разграничава шест равнища на когнитивна дейност: запомняне, разбиране, прилагане, анализ, оценяване и създаване. Логиката, която се следва, е, че колкото по-ниско е когнитивното ниво на една задача, толкова по-лесно тя се изпълнява от ГИИ вместо от студента. Следователно софтуерната уязвимост на изпитните инструменти се определя пряко от когнитивното ниво, което те тестват.



Фиг. 1. *Ревизирана таксономия на Блум*

Изследователи (Raman, Achuthan & Nedungadi, 2026: 21) установяват, че ГИИ-системите демонстрират висока ефективност именно на ниските равнища на таксономията (запомняне и разбиране), върху които масово са изградени университетските тестове, като същевременно генерират невярна или остаряла информация (МОН, 2024: 40). Следователно традиционните форми на оценяване са особено уязвими, докато извън обхвата на ГИИ засега остават задачи, изискващи лична позиция в актуален контекст, критична оценка на противоречиви източници или устна аргументация в реално време.

На национално и институционално ниво посочените рискове намират официално признание. Министерството на образованието и науката изрично предупреждава за рисковете от генериране на неточна, невярна или остаряла информация и задължава преподавателите да възпитават дигитална критичност у обучаемите (МОН, 2024). Рамката за компетентност на учителите на ЮНЕСКО дефинира три нива на компетентност (придобиване, задълбочаване и създаване) и изисква от преподавателите да развият специфични ИИ-оценъчни умения, за да гарантират дългосрочна академична почтеност чрез преформатиране на задачите (Miao & Cukurova, 2024: 25–27).

2.3. Дистанционното обучение като усилващ фактор

Дистанционното обучение по своята природа разполага студента в среда с неограничен достъп до технологични ресурси в хода на оценяването. При присъствено изпитване контролът е физически осъществим, докато при асинхронно онлайн оценяване той е структурно невъзможен.

Теоретична обосновка на феномена се намира в класическата *Теория за транзакционното разстояние* на Майкъл Мур (Moore, 1997: 22). Според нея по-ниската степен на институционална структура в програмите в дистанционно обучение корелира с по-високата автономия на студента и следователно с по-високия риск от злоупотреба с ГИИ. Процесите на дигитализация оптимизират технологичното управление, но и целят да формират изцяло нови поведенчески подходи относно начина, по който се използват или потребяват дадени услуги (Mancheva-Ali, 2021: 257). В съвременния образователен контекст това обяснява бързата адаптация на студентите, поставени в дигитална среда. Те изграждат нов поведенчески модел на ползване на образователната услуга,

при който нерегламентираното делегиране на задачи към ГИИ се превръща в най-краткия път към резултата.

В българските висши училища за управление на дистанционното обучение широко се използва платформата за електронно обучение **Мудъл (Moodle)**. Тя предоставя разнообразни инструменти за оценяване: тестове, задания, форуми и др. При текстовите задания студентите качват текстов файл без реален контрол върху условията на работа, което прави неограниченото използване на ГИИ технически лесно и достъпно за всеки студент. Показателно е, че именно заданията с отворен писмен отговор (традиционно считани за по-трудно възпроизводими в сравнение с въпросите с избор) се оказват най-уязвими в дигитална среда, тъй като изходът на големите езикови модели е текст, структурно идентичен с очаквания студентски продукт.

3. МЕТОДОЛОГИЯ

В настоящото изследване емпиричният анализ е ограничен до текстогенериращи системи, базирани на големи езикови модели (Large Language Models, LLMs), като по-широкото понятие „генеративен изкуствен интелект“ (ГИИ) се използва в общ теоретичен и образователен контекст.

3.1. Обект, предмет, цел и задачи на изследването

Обект на изследване е многокомпонентно писмено задание като инструмент за оценяване по финансова дисциплина в условията на дистанционно обучение. **Предмет** на изследването е степента на уязвимост на многокомпонентното писмено задание към съдържание, генерирано от големи езикови модели (LLMs), и възможностите за неговото идентифициране.

Основната **цел** е да се провери емпирично доколко многокомпонентното писмено задание може да бъде възпроизведено чрез LLM-базирани текстогенериращи системи и да се идентифицират индикатори за вероятната им употреба в реални студентски разработки.

За изпълнение на целта са поставени следните задачи:

✓ изследване на теоретичните характеристики на LLM-базираните текстогенериращи системи и систематизиране на основните рискове и предизвикателства, които те поставят пред дистанционното обучение;

✓ симулиране на еталонни отговори чрез водещи големи езикови модели, съответстващи изцяло на условията на реално студентско задание;

✓ провеждане на сравнителен семантичен, структурен и стилистичен анализ между еталонните отговори и студентските текстове чрез комбиниран метод – ИИ-асистиран и експертен;

✓ очертаване на научно-практически препоръки за оптимизиране на изпитния дизайн в дигитална среда.

В изследването се използват следните методи: качествен текстов анализ, сравнителен анализ, бенчмарк симулация, обобщаване и интерпретация на резултатите.

Важно методологично уточнение е, че LLM-базираните текстогенериращи системи изпълняват двойна роля в изследването: те се използват както като обект на анализ (изследване на уязвимостта на разглежданото многокомпонентно задание), така и като помощен диагностичен инструмент при сравнителния анализ на текстове. Това е съзнателно решение на автора, при което тези системи се използват за предварително идентифициране на сходства и стилистични маркери. Експериментът е структуриран в две фази, като фаза № 1 установява теоретичната уязвимост на заданието, а фаза № 2 изследва дали и в каква степен тази уязвимост е реализирана в реалните студентски разработки. Двуфазовият дизайн позволява да се установи не само потенциалната технологична уязвимост на писменото задание, но и да се анализира степента на нейното реално проявление в студентските изпитни работи. Верификацията на резултатите е извършена от преподавателя чрез самостоятелен преглед на всички студентски разработки. Всяка студентска разработка е сравнена с еталонните отговори по три критерия – наличие на идентични формулировки, съвпадение на структурата на аргументацията и употреба на термини, липсващи от учебния материал. Крайната интерпретация на резултатите е извършена от автора въз основа на посочените критерии. Фокусът на изследването е върху идентифицираните индикации, а не върху окончателни и категорични заключения за всеки отделен случай.

3.2. Контекст и специфика на дисциплината

Обучението по финансови дисциплини в дистанционна форма изисква специфичен подход при избора на оценяващи инструменти, съобразен с характера на преподаваната материя. Изследването, направено в настоящата публикация, е с дисциплината „Централно банкерство“. Тя има за цел да формира стратегическо макроикономическо мислене у обучаемите чрез детайлно осмисляне на теоретичните основи и институционалната рамка на съвременните централни банки, същността на паричната политика, нейните инструменти и институционална независимост, механизмите за управление на валутните резерви, както и съвременните стандарти за макропруденциален надзор в контекста на интеграцията на БНБ в Еврозоната.

За разлика от приложните финансови курсове, където фокусът е върху количествените изчисления и финансовите коефициенти, тук водещо е концептуалното разбиране на причинно-следствените връзки в банковата сфера – формат, в който големите езикови модели са особено силни. Следователно от тази гледна точка дисциплината попада в умерено висока рискова зона. Компенсиращ фактор е динамичността на политическата среда: актуални решения на БНБ, ЕЦБ, отражение на геополитически рискове върху паричната политика и др. – теми, които изискват познаване на реални, текущи данни, с които текстогенериращите ГИИ-системи не разполагат системно. Именно в това се крие ключовият ресурс за проектиране на ИИ-устойчиво оценяване в тази дисциплина.

3.3. Участници и задание

А. Участници

През летния семестър на учебната 2025/2026 г. е направено изследване с всички студенти от специалност „Финанси“ (ДН), ОКС „магистър“, обучаващи се в дистанционна форма във Великотърновския университет „Св. св. Кирил и Методий“. В конкретния случай обучаемите са седем на брой.

Б. Задание

На студентите е дадено многокомпонентно задание за самостоятелна разработка в системата Мудъл за получаване на текуща оценка. То съдържа седем тематични раздела с различни формати на отговор. Отделните компонента са с нарастваща когнитивна сложност (Таблица № 1), проектирани да измерват различни равнища по таксономията на Блум и да представляват различна степен на устойчивост спрямо LLM-базирани текстогенериращи системи.

Компонентите са:

✓ Тестови въпроси от затворен тип (с избор между четири алтернативи), които са насочени към проверка на фактологически знания.

✓ Твърдения за определяне на логическа вярност (от типа „вярно/грешно“), които изискват от студента не просто механично разпознаване, но и кратка логическа обосновка на даден теоретичен казус или трансмисионен механизъм.

✓ Отворени теоретични въпроси, чиято цел е да оценят способността на обучаемите самостоятелно да структурират академичен текст, да анализират причинно-следствени връзки и да боравят свободно със специфичната финансова терминология.

Таблица 1. Структура на многокомпонентното задание

Тип въпроси	Характеристика	Ниво по Блум	Целеви умения	Уязвимост спрямо LLM-базирани системи
Затворени въпроси	Проверка на познаване на понятия, факти и механизми	Знание, разбиране	Разпознаване, репродукция	Висока – LLM-базираните системи генерират верни отговори с висока точност

Твърдения „Вярно/Грешно“ с обосновка	Верификация на твърдение и задължителна аргументирана обосновка	Разбиране, анализ	Критична преценка, разграничаване на верни и грешни твърдения, аргументация	Висока към умерена – LLM-базираните системи могат да генерират верен избор и обосновка; възможни са грешки при неправилно копиране.
Отворени въпроси	Свободен текст – обяснение, аргументация, приложение на теория	Анализ, синтез	Концептуално разбиране, аргументация	Висока – структурираният аналитичен текст е оптималният изход на LLM
Лично становище	Аргументирана позиция по актуален дискусионен въпрос с национално значение	Оценяване, създаване	Критично мислене, автентичност, персонализация	Умерена – LLM-базираните системи генерират шаблонни позиции; уязвимостта намалява при персонализиран контекст

3.4. Експериментално изследване: „ИИ срещу стандартния тест“

В контекста на дистанционната форма на обучение администрирането на изпитния процес по дисциплината традиционно разчита на структурирани задания под формата на стандартизирани изпитни шаблони, които студентите получават чрез системата Мудъл.

В онлайн среда този дизайн на оценяване притежава редица предимства, сред които висока степен на обективност, възможност за бърза обработка на резултатите при големи потоци от студенти и ясно проследяване на заложените срокове. Въпреки това именно неговата висока степен на структурираност го превръща в потенциален обект на уязвимост при навлизането на съвременните интелигентни технологии. Когато отговорите се базират на налична макроикономическа теория и предварително дефинирани правила, изпитният шаблон се оказва лесно предвидим за алгоритмите на LLM-базираните текстогенериращи системи, което налага детайлното изследване на този ефект в приложната част на публикацията.

Фази на изследването

Изследването е структурирано като пилотно научно-практическо проучване с подчертано качествен и непредставителен характер, ограничено до рамките на конкретно изпитно задание по финансова дисциплина. Емпиричното проучване се реализира в две последователни фази.

Фаза 1: Генериране на „еталонни отговори“ (бенчмарк)

В първата фаза на експеримента е приложен методът на *бенчмарк генериране* с цел създаване на контролни еталони за сравнение. Авторският изпитен шаблон по дисциплината е подаден едновременно и независимо към три различни текстогенериращи системи с безплатен достъп: OpenAI ChatGPT (версия GPT-5.5 Instant); Google Gemini (версия Gemini 2.5 Flash); Perplexity AI (версия Sonar). Симулацията е проведена през май 2026 г. Промптът е формулиран на български език, без допълнителни настройки или ограничения, и е подаден еднократно към трите системи за генериране на пълни отговори по условията на реалното студентско задание. Текстът на промпта е следният: „Влез в ролята на студент от специалност „Финанси“ в дистанционна форма на обучение и реши прикаченото изпитно задание по дисциплината „Централно банкерство“.

Всеки от моделите влиза в ролята на обучаем и генерира пълни отговори на тестовите въпроси, логическите твърдения, теоретичните казуси и практическите въпроси в рамките на няколко секунди. В резултат на симулацията са получени три еталонни отговора, генерирани от трите системи, които отразяват специфичния стил, структура, семантични модели и фразеология на всяка от тях. По-нататък в изложението те се обозначават като „еталонни отговори“ и служат като референтна база за последващия анализ.

Фаза 2: Сравнителен анализ между отговорите, генерирани от LLM-базираните текстогенериращи системи, и студентските разработки

Във втората, същинска фаза на изследването е реализиран *комплексен сравнителен анализ* в нова чат сесия. За тази фаза на изследването е избран отделен и независим от бенчмарк моделите инструмент – Anthropic Claude (версия Sonnet 4.6) – с цел анализиране на предоставените текстове и генериране на ориентировъчни показатели за степента на семантично и структурно сходство. Системата не е използвана за вземане на окончателни решения относно произхода на текстовете, а единствено за предварително идентифициране на сходства и потенциални маркери.

Към нея са подадени едновременно следните три масива от данни:

1. Оригиналното изпитно задание (шаблон);

2. Трите предварително генерирани „еталонни отговора“ от Фаза 1;

3. Реалните разработки, предадени от студентите в дистанционна форма на обучение (общо 7 текстови файла).

Специализираният промпт, който е подаден за тази част от изследването, е: „Анализирай сходства между студентски и ГИИ-отговори, характерни маркери, ниво на еднаквост между студентите“. Целта е да се изиска задълбочен, сравнителен софтуерен анализ, ориентиран в три ключови направления:

• Сходства между студентските разработки и еталонните отговори – системата сравнява съдържанието, логиката на аргументацията и използваната банкова терминология в студентските работи спрямо трите еталонни отговора. Това позволява да се идентифицират индикации за възможно заимстване на пасажи или концептуални рамки, сходни с тези в еталонните отговори.

• Характерни текстови маркери – проследява се наличието на специфични лингвистични маркери, характерни за големите езикови модели (прекалено гладък академичен изказ, специфични преходни думи, липса на печатни грешки, или обратното – изкуствено генерирани структури).

• Колаборация между студентите – анализира се нивото на сходство между студентски разработки. Системата идентифицира припокриващи се идентични текстове, сходни примери и др., което в дистанционна среда дава основание за допускане на нерегламентиран обмен на информация, споделяне на изпитни файлове или използване на едно и също запитване (промпт) към ИИ от различни лица.

Въз основа на тези критерии моделът извежда ориентировъчни категории на текстуално и концептуално съвпадение (представени в Таблица № 2). Получените категории са резултат от оценка на семантично, структурно и лексикално сходство между студентските разработки и еталонните отговори, извършена от използвания ИИ-инструмент.

Изследването следва смесен методологичен подход, който съчетава качествен текстов анализ с обобщаващи количествени характеристики. Основната стратегия е сравнително-аналитична: реалните студентски отговори се съпоставят систематично с бенчмарк отговорите на трите системи по предварително дефинирани критерии. Изследването има пилотен характер и не претендира за статистическа представителност, а целѝ постигане на дълбочинна диагностична стойност.

4. РЕЗУЛТАТИ

4.1. Резултати от бенчмарк симулацията

В рамките на изследвания конкретен формат на дистанционно оценяване резултатите от бенчмарк симулацията потвърждават високата степен на уязвимост при конкретния формат на задание в дистанционна форма на изпитване.

При въпросите от затворен тип трите използвани LLM-базирани системи генерират верни отговори на всички въпроси. При твърденията от типа „Вярно/Грешно“ с последваща аргументация моделите идентифицират логическата вярност на съответните твърдения и предоставят академично издържана обосновка, която в 90% от случаите съвпада с предварително заложените критерии за оценка. Установено е, че ГИИ се справя по-трудно единствено с въпроси, изискващи интерпретация на специфични актуални данни (например от последния тримесечен бюлетин на

БНБ) – наблюдение, което подчертава важността на периодичната актуализация на изпитните материали.

Съпоставката на бенчмарк отговорите с реалните студентски разработки разкрива различни степени на сходство и индикации за вероятно технологично заимстване (Таблица № 2). При пет от седемте участници е идентифицирано умерено до високо сходство с поне един от трите еталонни отговора. При част от тях се наблюдават високи степени на текстово и концептуално съвпадение на посочените аргументи с еталонните отговори, включително идентично структуриране на дефиниции. Двама участници (Студент № 6 и Студент № 7) демонстрират разработки с характеристики на автентично студентско писане: по-кратък и по-малко „гладък“ изказ, както и типични студентски грешки. При един от тях е забелязана оригинална хронологична аргументация с конкретни исторически данни, които не се срещат в нито един от еталонните отговори.

Таблица 2. Профил на сходство и идентифицирани маркери по участници

Участник	Риск	Сходство	Основен идентифициран маркер
Студент № 1	Висок	Много високо	• дословни фрази, идентични с изхода на OpenAI ChatGPT; • индикации за директно заимстване в множество компоненти
Студент № 2	Висок	Високо	• високо сходство с еталонния отговор на OpenAI ChatGPT, с индикации за последваща собствена редакция; • оригинални елементи в отделни подтеми
Студент № 3	Умерен	Умерено-високо	• структура на OpenAI ChatGPT и преформулиране; • вероятен машинен превод (системно главни букви при институционални имена)
Студент № 4	Умерен	Умерено	• индикации за вероятна употреба на текстогенерираща ГИИ с диагностични грешки – допусната инверсия при верни/грешни отговори
Студент № 5	Умерен	Умерено	• смесен подход: признаци за вероятно използване на LLM-базирана текстогенерираща система при отворените въпроси и лекционен материал при затворените; • самостоятелна, макар и грешна преценка
Студент № 6	Нисък	Ниско	• използван лекционен материал; • по-кратки и по-малко „гладки“ отговори; • типични студентски грешки
Студент № 7	Нисък	Много ниско	• оригинална хронологична структура, уникален подход; • нулево сходство с другите участници

Важно е да се уточни, че посочените категории на сходство представляват ориентировъчни индикатори, генерирани въз основа на извършения софтуерен сравнителен анализ. Те служат като допълващ инструмент към качествения експертен анализ, а не като самостоятелно доказателство за произхода на текста. Всички констатации са верифицирани от преподавателя чрез самостоятелен преглед на студентските разработки и сравняване с идентифицираните от системата маркери.

4.2. Идентифицирани маркери на LLM-генериран текст

Направеният анализ позволи изолиране на характерни стилистични модели, присъщи за големите езикови модели. Например – прекалено гладък академичен изказ, специфични преходни фрази и съюзи, които рядко се срещат в естествената студентска реч, липса на типичните в он-

лайн формат правописни, пунктуационни или граматически грешки. Наличието на тези маркери позволява да се направи обосновано научно допускане за възможно LLM-генерирано съдържание в съответната разработка.

Най-високо сходство с еталонния отговор на OpenAI ChatGPT е установено при четирима от участниците. При част от участниците се наблюдават индикации за възможно пряко използване на генериран текст, докато при други еталонният отговор изглежда по-скоро като отправна точка за последваща редакция. Тази система най-често се асоциира с характерни фрази при описание на институционални промени и с термини, които не присъстват в учебния материал. Сходство с еталонния отговор на Google Gemini се наблюдава при двама участници, а с Perplexity – вероятно само в комбинация с останалите при един случай. Два случая са особено интересни. Студент № 3 е изписал имената на банки и институции с главни букви навсякъде в текста, което е характерен белег за автоматичен превод от английски, който не е редактиран след това. Студент № 4 е допуснал инверсия при верните и грешните отговори, което дава основание за допускане, че таблица, генерирана от LLM-базирана текстогенерираща система, е пренесена без критична преработка на съдържанието.

4.3. Сходство между участниците

Един от най-важните административни резултати от изследването е засичането на висока степен на сходство между самите студентски работи (Таблица № 3). Анализът идентифицира индикации за сходни пасажи, повтарящи се примери и сходни структури в текстовете на различните участници. В условията на дистанционна форма това насочва към индикации за нерегламентиран обмен на изпитни материали или използване на едно и също запитване към LLM-базирана текстогенерираща система от различни студенти – и в двата случая това компрометира индивидуалния характер на оценяването. Важно е да се отбележи обаче, че съответствието „участник–еталонен отговор“ е значително по-голямо от съответствието „участник–участник“, което по-скоро насочва към независимо използване на сходен модел на ИИ, а не директно споделяне на работи между обучаемите.

Таблица 3. Сходство между участниците

Двойка участници	Сходство	Интерпретация
Студент № 1 ↔ Студент №2	Високо	• идентична таблична структура и сходни формулировки; • вероятно независима употреба на една и съща LLM-базирана текстогенерираща система, а не директно споделяне
Студент № 3 ↔ Студент № 6	Умерено	• почти идентична таблична организация; • вероятен общ голям езиков модел или споделен шаблон за попълване
Студент № 4 ↔ Студент № 5	Ниско-умерено	• сходен речник при отворените въпроси; • вероятно общ лекционен източник или сходна текстогенерираща ГИИ система
Студент № 7	Минимално	• нулево значимо сходство с нито един друг участник; • индивидуален подход

Студент № 7 е единственият обучаем, който се отличава, тъй като няма значимо структурно или езиково сходство с нито един друг обучаем, нито с някой от еталонните отговори. Нивото на детайлност и оригиналността на подхода надвишават характеристиките на всички еталонни отговори в съответните компоненти, което може да се разглежда като индикация за по-висока степен на самостоятелност на разработката.

5. ДИСКУСИЯ

Резултатите от изследването показват, че интегрирането на големите езикови модели в дистанционната образователна среда може да окаже съществено влияние върху валидността, надеждността и автентичността на оценяването. Следва да се подчертае, че използването на големите езикови модели за оценка на сходство между текстове има спомагателен характер и не може да замести независима експертна или техническа верификация. В рамките на настоящото изследване се наблюдават индикации, че традиционното задание може частично да загуби способността си да измерва реалния процес на разбиране.

5.1. Уязвимост на стандартните изпитни шаблони в дигитална среда

Валидността на оценяването предполага, че изпитните задания измерват именно знанията и уменията, които преподавателят целѝ да оцени.

В случая резултатите показват, че:

✓ големите езикови модели могат да решават лесно тестови въпроси, включително въпроси от типа „Вярно/Грешно“, както и въпроси с вариативни отговори;

✓ кратките аналитични въпроси, които традиционно се считат за по-трудни, също се оказват в рамките на възможностите на текстогенериращите ГИИ-системи, които генерират структурирани, логични и академично звучащи отговори;

✓ студентските разработки, които съвпадат с еталонните бенчмарк отговори, поставят въпроса дали оценяването измерва реално разбиране, или умение за използване на ИИ.

Следователно тези системи са особено ефективни именно при предвидими, добре структурирани задачи с широка формулировка – точно тези, които доминират в стандартните изпитни шаблони на по-теоретичните дисциплини. Колкото по-общ е въпросът, толкова по-компетентен е отговорът на изкуствения интелект. Това противоречи на съвременните парадигми във висшето образование, според които интегрирането на активни стратегии за преподаване, студентски ориентираното обучение и иновативните образователни среди са жизненоважни за изграждането на специалисти, способни да се адаптират към сложната и динамична пазарна реалност (Kostadinova, 2025: 284). Вместо това, традиционното задание може да не измерва в достатъчна степен това, за което е проектирано – то рискува да отразява в по-голяма степен способността на обучаемия да работи с такива инструменти, отколкото реалното му разбиране.

5.2. Намалена надеждност на оценяването

Смисълът на оценяването е да разграничи студентите според действителните им знания. Резултатите от анализа обаче очертават притеснителна реалност: висока степен на хоризонтална еднаквост между разработките, силно застъпени машинни структури и почти пълна липса на типичните стилови и логически пропуски и индивидуални грешки, характерни за самостоятелните студентски разработки. Едновременното отпадане на тези индикатори практически отнема от преподавателя основните му ориентири за автентичност и води до изравняване на резултатите.

Следователно, когато при значителна част от разработките се наблюдават индикации за частично или значително използване на генерирано съдържание, оценката рискува да не изпълнява в достатъчна степен основната си функция (а именно, да диференцира реалните нива на подготовка), тя вече измерва качеството на използвания модел. Получава се ситуация, в която студент, вложил сериозен труд в самостоятелна разработка с леки несъвършенства, може да получи по-ниска оценка от студент, предал текст, генериран с помощта на ИИ. Това е не само методологичен, но и етичен проблем, тъй като системата създава предпоставки за неблагоприятно третиране на автентичния студентски труд, а възнаграждава заимстването.

5.3. Проблемът с „унифицирания стил“

Една от най-важните интерпретации е свързана с т.нар. „унифициран стил“ на моделите за изкуствен интелект: последователна структура, академични фрази, балансирани аргументи и

липса на индивидуален изказ. Когато група от студенти използва сходни безплатни LLM-модели и сходни промптове, системите връщат близки семантични ядра. Резултатът е масово генериране на „текстове близнаци“ и високо ниво на колаборация между отделните участници. Този унифициран стил заличава индивидуалния почерк, критичното мислене и личната авторска позиция на студента, превръщайки писмените работи в еднообразен софтуерен продукт. Оказва се парадоксално, че именно несъвършенствата в самостоятелните студентски разработки могат да се считат за индикатор за самостоятелен труд и реална когнитивна ангажираност.

Това поставя специфични предизвикателства пред дисциплина като „Централно банкерство“, където се очаква студентът да демонстрира разбиране на причинно-следствени връзки, а не да възпроизвежда дефиниции. В рамките на настоящото изследване това е видимо в сравнението между разработките с висок и нисък риск: разработките на Студент № 6 и Студент № 7 са разпознаваеми именно по конкретността, по личния отпечатък, дори и по грешките си, докато унифицираният стил на останалите подсказва липсата на собствен почерк.

5.4. Влияние върху преподавателската практика

Навлизането на големите езикови модели в образователната система променя ролята на преподавателя по начин, който не е предвиден в традиционните академични функции, като го натоварва с неприсъщи до момента административни отговорности. Вместо да се фокусира върху менторството и оценяването на реалното студентско развитие, той е принуден да отделя значителна част от времето си за проверка на автентичността на предаденото съдържание. Анализът на текстове „на ръка“ за маркери за ИИ съдържание и сходства между разработките е времеемък и увеличава административната натовареност на преподавателя извън традиционните нормативни рамки, превръщайки се в сериозен източник на професионално прегаряне (т.нар. burnout).

Настоящото изследване илюстрира и друго измерение на проблема, което рядко се споменава в литературата – **асиметрия между усилието на преподавателя и студента**. В реалната университетска практика разработката на качествено, научно издържано, многокомпонентно задание, обхващащо различни формати и когнитивни нива, изисква от преподавателя няколко работни дни (както е и в конкретния случай), докато отговорите на студентите, ползвали LLM-базирани текстогенериращи инструменти, могат да бъдат генерирани в рамките на минути, а може би и секунди. Преподавателят се оказва в позиция, в която времето за създаване и последващ технологичен контрол на задачите надвишава в пъти времето, необходимо за тяхното „решаване“ от страна на обучаемите. Този феномен създава скрита, некомпенсирана административна тежест в дигитална среда.

Предизвикателството пред академичната общност не е как да забрани изкуствения интелект, а как да трансформира изпитните шаблони така, че те да стимулират автентичното критично мислене. Конкретните препоръки за адаптиране на оценяването са систематизирани в следващата част.

5.5. Технологични и педагогически предизвикателства

Наред с установените промени в надеждността на оценяването и преподавателската практика използването на големи езикови модели в образователния процес поставя и някои предизвикателства, които следва да бъдат отчетени при адаптирането на моделите за оценяване.

Основно място заемат технологичните предизвикателства. Наличните програми за автоматично засичане на съдържание, създадено с ИИ, не са надеждни за текстове на български език. Те грешат в двете посоки – от една страна, невинаги засичат реално заимстване, а от друга страна – погрешно маркират автентична студентска работа като генерирана от ИИ. Това прави автоматизираната детекция рискован инструмент за вземане на решения и прехвърля отговорността за диагностика на почтеността върху преподавателя, който трябва да се опре на собствен качествен анализ. Именно заради това е удачно системите за изкуствен интелект да са помощен инструмент, а не основен. Допълнително предизвикателство представлява тяхната динамика, тъй като те се

обновяват непрекъснато, стават все по-добри в имитирането на човешки стил и все по-трудни за разпознаване. Маркерите, валидни днес, може да не са валидни след няколко месеца.

Другият важен аспект са педагогическите предизвикателства**.** Когато студентът делегира аргументацията на ГИИ, той пропуска именно процеса, в който се изгражда аналитичното мислене. Финансовите дисциплини изискват способност за интерпретация на данни, оценка на риска и вземане на решения в условия на неопределеност – умения, които не се развиват чрез четене на отговорите, генерирани от LLM-базираните текстогенериращи системи, колкото и качествени да са те. Рискът е дългосрочен, тъй като съществува вероятност студентът да не изгради реална компетентност. Съществува и обратен педагогически риск – свръхреакцията. Тотална забрана на използване на ИИ е едновременно неприложима на практика и контрапродуктивна. Тя лишава студентите от инструмент, който ще бъде неизбежна част от тяхната бъдеща среда. Задачата на висшето образование не е да изолира студентите от технологията, а да ги научи да я използват критично и отговорно.

6. ПРЕПОРЪКИ ЗА ОПТИМИЗИРАНЕ НА ОЦЕНЯВАНЕТО

В контекста на настоящата публикация могат да се формулират следните препоръки, които успешно да се имплементират в дистанционната форма на обучение:

➢ Модификация на писмените казуси

Основната препоръка е преминаване към персонализирани задания, при които контекстът е специфичен и непознат за текстогенериращите ГИИ-системи. За икономическите дисциплини това означава обвързване с актуални данни и решения, изискване за анализ на конкретна национална ситуация, както и формулировки, изискващи лична позиция, която не може да бъде генерирана автентично от модел, обучен върху общи данни. Това намалява полезността на безплатните големи езикови модели, тъй като те са склонни да генерират по-абстрактни и остарели отговори.

➢ Устно изпитване като стандартна практика

Устната защита остава най-надеждният инструмент за проверка на самостоятелно усвоените знания. Кратък синхронен разговор върху предадената разработка е практически невъзможно да се симулира от студент, който я е написал с ГИИ без реално осмисляне. В контекста на дистанционното обучение това може да се реализира чрез видеовръзка, без да изисква физическо присъствие. Важно е да се отбележи, че устната верификация не трябва да се разбира като наказателна мярка, а като стандартен компонент от оценяването, приложен към всички обучаеми.

➢ Регламентирана употреба вместо забрана

В дългосрочна перспектива по-продуктивен подход от забраната е регламентираната употреба на ГИИ с изискване за прозрачност. Възможно решение е студентът да декларира употребата на ГИИ и да описва как е използвал съответния ИИ. Така ГИИ се превръща от заплаха за оценяването в обект на критично мислене, което е и по-реалистична позиция спрямо технологична реалност, която не може и не трябва да бъде игнорирана. Този подход изисква ясна институционална рамка на ниво университет (факултет или катедра), която няма забранителен характер, а регулативен, т.е. каква е допустимата употреба на ГИИ, как се декларира, какви са последствията при нарушение. Без такава рамка преподавателят остава сам с проблем, който е системен по своята природа. В допълнение всеки преподавател може да формулира конкретни правила за своята дисциплина – кога ГИИ може да се използва, кога е недопустим и как трябва да се декларира употребата му.

Студентите, от своя страна, следва да бъдат обучени как да използват текстогенериращи ГИИ-системи етично, как да проверяват достоверността на генерираната информация и как да развиват собствен стил и аналитични умения.

➢ Големите езикови модели като партньор на преподавателя в проверката

Както в конкретния случай, те могат да бъдат в помощ на преподавателя, тъй като той може да ги използва за сравнение на стилове (генериране на еталонни отговори и сравнение с маркерите на конкретни системи). Второто им приложение е като диагностични инструменти, които могат

да откриват сходства. Преподавателят може да ги използва за откриване на повторения между отделни разработки, анализ на структура, идентифициране на шаблонни фрази.

➢ Повишаване на ИИ-компетентността на преподавателя

Разпознаването на съдържание, генерирано от LLM-базирани текстогенериращи системи, е умение, което може да се развие и систематизира. Познаването на характерните маркери на различните системи, изграждането на сравнителна база от автентични студентски работи и споделянето на опит между колеги са практически стъпки, достъпни без допълнително финансиране или институционална подкрепа. Това означава, че диагностичната компетентност на преподавателя не е еднократно придобита – тя изисква постоянно надграждане.

7. ЗАКЛЮЧЕНИЕ

В рамките на настоящото пилотно изследване се установи, че уязвимостта на дистанционното оценяване към използването на големи езикови модели не произтича единствено от недостатъчна сложност на заданията, а и от самата им природа: задачите, изградени върху възпроизвеждане на знание, се оказват особено уязвими към LLM-базирано генериране. Двуфазовият дизайн на изследването потвърди теоретичната уязвимост и очерта сериозни индикации за нейното реално проявление.

Последствията надхвърлят конкретния случай. Оценяване, което не може да разграничи самостоятелния труд от технологично генерирания, рискува да загуби своята основна функция – да удостовери реално придобита компетентност. Асиметрията на усилието между преподавател и студент допълнително задълбочава проблема, като натоварва системата с невидима административна тежест.

Препоръките, изведени от изследването, обхващат адаптиране на изпитния дизайн, регламентиране на употребата на текстогенериращи системи и развитие на диагностичната компетентност на преподавателя.

Настоящото изследване има диагностична стойност и целѝ да илюстрира конкретен проблем, а не да предлага окончателни обобщения за всички дистанционни изпитни практики. То има своите ограничения – малка извадка от студенти, обучаващи се в дистанционна форма, една дисциплина, един изпитен формат, които очертават и посоките за бъдещи проучвания. Бъдещите изследвания биха могли да разширят обхвата към различни дисциплини и форми на обучение, да установят дали регламентираната употреба на ГИИ реално намалява заимстването и да проследят как нарастващата грамотност на студентите по отношение на ГИИ се отразява върху тяхното академично поведение във времето.

Интегрирането на генеративния изкуствен интелект в образованието е необратим процес. Предизвикателството не е как да го спрем, а как да трансформираме изпитните шаблони в онлайн среда така, че да продължат да измерват онова, за което са създадени – реалното критично мислене и автентичната компетентност на студента.

ЛИТЕРАТУРА / REFERENCES

Бянов 2026: Бянов, И. Приложение на изкуствения интелект в обучението на земеделски стопани. – В: *Дигитална трансформация на образованието – проблеми и решения*. Четвърта научно-практическа конференция с международно участие, Русе, 711. ISSN 3033-0610. // **Byanov 2026:** Byanov, I. Prilozhenie na izkustvenia intellekt v obuchenieto na zemedelski stopani. – V: *Digitalna transformatsia na obrazovanieto – problemi i reshenia*. Chetvartha nauchno-prakticheska konferentsia s mezhdunarodno uchastie, Ruse, 711.

МОН 2024: Министерство на образованието и науката. Насоки за използване на изкуствен интелект (проект, януари 2024 г.). София // **MON 2024:** Ministerstvo na obrazovanieto i naukata. Nasoki za izpolzvane na izkustven intellekt (proekt, yanuari 2024 g.). Sofia.

Anderson, Krathwohl 2001: Anderson, L. W., D. R. Krathwohl. A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. New York: Longman.

Baidoo-Anu, Owusu Ansah 2023: Baidoo-Anu, D., L. Owusu Ansah. Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning. – In: *Journal of AI*, 7(1), 52–62.

Čolić, Sukić & Anđelković 2025: Čolić, V., E. Sukić, M. Anđelković. AI-Supported Learning Methodologies: Theoretical Review and Empirical Research on the Educational Process Impact. // *TEM Journal,* Vol. 14, No. 4, 2025, pp. 3714–3727. DOI: 10.18421/TEM144-75.

Davis 1989: Davis, F.D. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. – In: *MIS Quarterly,* 13(3), 319–340.

Holmes, Bialik, Fadel 2019: Holmes, W., M. Bialik, Ch. Fadel. Artificial Intelligence in Education: Promise and Implications for Teaching and Learning. Boston: Center for Curriculum Redesign.

Kostadinova 2025: Kostadinova, N. From Education to Practice in Tourism: Insights from Bulgaria. Educational Foundations and Teaching Innovations in Tourism Studies. – In: *Socio-Economic Analyses (Социално-икономически анализи)*, 2/2025 (28), 283–302. – https://doi.org/10.54664/VCYC9887

Mancheva-Ali 2021: Mancheva-Ali, O. The digital economy and its impact on consumption in the regions of Bulgaria. – In: *Advances in Economics, Business and Management Research*, 181, 256–261. – https://doi.org/10.2991/aebmr.k.210710.043

Miao, Cukurova 2024: Miao, F., M. Cukurova. AI Competency Framework for Teachers. Paris: UNESCO.– https://doi.org/10.54675/ZJTE2084

Moore 1997: Moore, M.G. Theory of transactional distance. – In: D. Keegan (Ed.). *Theoretical Principles of Distance Education*. London: Routledge, 22–38.

Noroozi, Soleimani, Farrokhnia, Banihashem 2024: Noroozi, O., S. Soleimani, M. Farrokhnia, S.K. Banihashem. Generative AI in Education: Pedagogical, Theoretical and Methodological Perspectives. – In: *International Journal of Technology in Education (IJTE)*, 7(3), 373–385.

Perkins 2026: Perkins, G. Resilient assessment in the age of AI: authentic design and the case for verbal examinations in business education. – In: *Assessment & Evaluation in Higher Education.* – https://doi.org/10.1080/02602938.2026.2644516

Raman, Achuthan, Nedungadi 2026: Raman, R., K. Achuthan, P. Nedungadi. Generative AI Integration in Education: Theoretical Review and Future Directions Informed by the ADO Framework. – In: *Information,* 17(3), 241. – https://doi.org/10.3390/info17030241

Zhai, Wibowo, Li 2024: Zhai, C., S. Wibowo, L. D. Li. The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. – In: *Smart Learning Environments*, 11, 28. – https://doi.org/10.1186/s40561-024-00316-7