https://doi.org/10.37708/ezs.swu.bg.v24i3.4

# ИЗКУСТВЕНИЯТ ИНТЕЛЕКТ В АКАДЕМИЧНА СРЕДА: ЕТИЧНИ, ТЕОРЕТИЧНИ И ПРАКТИЧЕСКИ РАКУРСИ

*Десислава Иванова*

Шуменски университет „Епископ Константин Преславски“, България
Email: des.ivanova@shu.bg

ORCID ID: 0009-0002-9107-4744

## ARTIFICIAL INTELLIGENCE IN ACADEMIC FIELDS: ETHICAL, THEORETICAL AND PRACTICAL PERSPECTIVES

*Desislava Ivanova,* Konstantin Preslavsky University of Shumen, Bulgaria

ABSTRACT: The object of this article is the role of artificial intelligence (AI) in the academic field. The objective of this article is to discuss different ways in which AI can be used in academic education and scientific research work. Specific policies in an international aspect, which try to regulate and balance this ratio, have been presented. The article outlines three specific aspects of AI in the academic educational and scientific research field. Artificial intelligence should not be ignored, but rather integrated through well-structured policies or approaches in academic practices. Only in this way is there a possibility of preserving integrity and apply critical measures against fraud. What makes the second specific of AI is that it can perceive simultaneously as a tool and as an object of influence. This article describes its function as a research and methodological tool, but what is more interesting is its presence as an active factor, which alters academic practices and predefines criteria of validity. The third aspect the article examines is the possibilities for an ethical and transparent AI and it elaborates on the fact that responsible use of AI is directly connected with creating policies that should regulate the extent of its use that guarantee transparency, methodological criticality, copyright management, declaration of content, etc. One of the conclusions is that the new reality enforces a change of perspective concerning the essence of tools which shape the competencies of modern individuals.

KEYWORDS: artificial intelligence, academic integrity, ethics, theoretical and practical perspectives, functions

**Въведение**

В съвременните образователни процеси и в научноизследователската дейност все по-често и по-осезаемо е присъствието на изкуствения интелект (ИИ) в различни форми и по различни начини. Естественият разум става все по-зависим от него, а въпросите, свързани с регулацията, етичните съображения и дилеми при използването му, стават все по-многобройни и провокират дискусии в много посоки.

В зоните на приложение на ИИ попада и образованието. От една страна, чисто технологичните възможности, които той предлага (автоматизирано създаване на текст, код, графика, мултимедийно съдържание и много други), улесняват обучителните и изпитните процедури. Появяват се нови предизвикателства към оценяването на

резултатите от обучението и въобще по нов начин започват да се позиционират ролите на обучаеми и обучители чрез посредничеството на ИИ.

Друга сфера, в която все по-активно „помага“ ИИ, е академичната научноизследователска дейност. Бързото и все по-усъвършенствано развитие на ИИ инструментите създава нови възможности, но и сериозни проблеми пред академичното писане и в много от аспектите на академичното образование. От етична гледна точка започва да се разколебава смисълът на категории като академична честност, отговорност, авторство. Границите между оригиналното и генерираното от ИИ съдържание се припокриват. Следователно е необходимо търсенето на някакъв баланс между иновативния потенциал на изкуствения разум, от една страна, и академичния ангажимент с почтеността при спазването на авторските права, от друга.

Тези (и други) размисли са повод за настоящото изложение, което, структурирано в три тези, се опитва да предложи една, разбира се, съвсем неизчерпваща проблема теоретична концептуална рамка на отношението ИИ : академично образование. Също така се коментират и конкретни международни регулаторни политики, чрез които университетските среди се опитват да управляват това отношение. Една от целите тук е да се насочи вниманието към различни стратегии, които се прилагат спрямо използването на изкуствения интелект в образованието и в научноизследователската работа. Не се предлага категорична позиция „за“ или „против“, а и не това е водещото разграничение. По-разумната принципна позиция е, че технологията трябва да подобрява, а не да подкопава академичното знание. Друга цел, която изложението си поставя, е да направи опит за набелязване на общи тенденции в международната регулация, а на тяхната основа да се оформят разумни перспективи за обогатяване и осмисляне на обучението както от гледна точка на технологиите, така и в етичен аспект. Смисълът е не само да бъдат посочени важните предизвикателства пред изследователските и образователни стандарти за академична автономия, автентичност и етика, но и да се дадат отговори (макар и частични и предварителни) на въпроса по какви начини и с какви роли ИИ присъства в академичната сфера, но и (което е по-съществено) какъв знак е това за профила на образованието в динамичния свят на дигиталните иновации, в който реално и виртуално присъстват почти равнопоставено.

**Първа теза: ИИ като активен участник в академичната образователна среда**

За изграждането на стабилни и ясни етически рамки е необходимо на първо място да бъде признат и осмислен фактът, че ИИ е действащ и действителен (съ)участник в академичните практики както в образователен, така и в изследователски контекст. Неговото присъствие може метафорично

да се уподоби на англо-саксонския идиом “elephant in the living room”[1] („слон в хола“): той е видим и разпознаваем от общността, но публичното му валидиране е несигурно и поражда съмнения. Иначе казано, участието му е осезаемо в различни области (медии, технологии, изкуство, образование, медицина, администрация и пр.), но институционалното му инкорпориране като обект и субект създава проблеми и напрежение. В частност такава е ситуацията и в университетските среди, а именно: непризнаване на реалното му участие във формирането на различни аспекти от човешкото познание и креативните човешки умения и способности, които разширяват своите хоризонти именно благодарение, а не въпреки технологичните иновации. Активното (съ)участие на ИИ в академичните научноизследователски и образователни дейности отключва и още един морален и културен парадокс. Факт е, че използването му може да съкрати и улесни голяма част от изследователските процеси и образователните процедури. По този начин ще се „освободи“ човешки креативен и аналитичен капацитет. От другата страна обаче стоят нерешени въпросите за авторството, за академичната честност, за изследователската добросъвестност, за плагиатството и пр. В тази посока разсъждава изследователският екип на проучване, посветено на генеративния изкуствен интелект и адаптирането на университетските политики и ресурси за неговото използване (Wang et al., 2024). Генералният извод, до който те достигат, е, че са подкопавани академичната честност, оригиналността и авторството. Тези чисто академични елементи са поставени под реална заплаха и единственият изход е не игнорирането, а интегрирането на ИИ в образователния процес чрез структурирането и изпълнението на добре обмислени политики и/или подходи в академичните практики (Wang et al., 2024, p. 11). И нещо, което е по-важно – мерките срещу неетичното и безпринципно използване на ИИ трябва да се основат както на профила на съответната академична институция (технически, хуманитарен, медицински и пр.), така и да отразяват спецификите на всяка отделна дисциплина. Тогава е възможно преодоляването на скрития риск от обезценяването на качества като креативност, отговорност, етичност и пр. Казано по друг начин, в епохата на ИИ съхраняването на почтеността в академичните среди е свързано не само с усъвършенстването на технологиите за откриване и регистриране на измами, несъответствия и неетични действия, но и преди всичко с прецизиране и преосмисляне на целите и структурните компоненти на оценяването. Разбира се, проблемите с оценяването на знанията и уменията не са свързани с появата на изкуствения разум, те са старият, добре познат спътник на масовизацията и комерсиализацията на висшето образование. Неговото преформатиране от процес на интелектуално формиране в бизнес план за квалификации е понижило отдавна

[1] Литературният контекст на израза, както е известно, се свързва с баснята на Иван Крилов „Любопитният човек“ от 1814 г.

мотивационната ангажираност на студентите и е стимулирало у тях по-скоро инструментализирани нагласи към ученето. И в този смисъл, когато обучението се е превърнало в механизъм за придобиване на квалификации и степени, а оценките се възприемат като бюрократични пречки за тяхното постигане, студентите са по-склонни да предприемат етично неприемливи действия в изпитните процедури и ИИ активно участва в тях. Вероятно едно от възможните противодействия на този лавинообразен поток от измами е в изграждането или по-точно във връщането към етичните научни и педагогически модели, в *рехуманизирането* на оценяването чрез интеракция и персонализация, чрез насърчаване на критичното мислене, креативността и адаптирането на знанията към реални проблеми и житейски казуси. Отправна точка трябва да са не толкова задачите и ситуациите, свързани с механичен тип запаметяване и възпроизвеждане, а проектни дейности по конкретни, индивидуални случаи, устни презентации и креативни тренинги с отворени решения и др. Такива методически похвати трябва да са насочени към активирането на умения за индивидуална рефлексия у студентите и за стремеж към оригиналност. Така може да се намали рискът от злоупотреби при използването на ИИ, но и да се стимулира смислено, мотивирано ангажиране с учебния материал и с изпитните процедури у студентите.

Етичният образователен ефект е постижим например в следните аспекти: чрез изпълнението на персонализирани задачи, фокусирани в сферата на техните интереси, обучаемите ще валидират стойността на ученето и съответно на оценяването; ако студентите имат възможността да бъдат причислени към академична общност, която отхвърля измамата по интелектуални и етични причини, тогава те биха могли да развият своя стабилна самоидентичност и което е по-важно, критични нагласи към измамата като цяло и в частност – към нерегламентираното използване на ИИ. В подкрепа на тази теза за академичната интеграция на студентите може да послужи едно екипно проучване на университета в Ливърпул (Johnston et. al., 2024), което обобщава и прави анализ на студентските позиции „за“ и „против“ употребата на ИИ. На тяхна основа текстът на статията оформя ясни препоръки за промени в университетските документи по линия на академичната честност и почтеност в оценяването. Резултатите от анкетните карти и от сформираните фокус групи показват, че по-уверените в академичната си подготовка студенти са по-малко склонни да прибягват до инструменти на ИИ в обучението си. Същите тези студенти много по-малко подкрепят колегите си, които ги използват (Johnston et. al. 2024, pp. 8 – 11). Голяма част от студентите изразяват позицията, че са необходими университетски политики, които да регламентират целесъобразното използване на тези технологии (Johnston et. al., 2024, pp. 12 – 14). В обобщен вид анализите сочат необходимостта не от забрани, а от ясен регламент, който да осигури равен достъп до технологичните възможности на ИИ. Ако отново направим аналогия с идиоматичния образ “elephant in the living room”, ще обобщим, че „слонът“ вече съвсем не е само в хола, но *обживява*

и мисловните пространства на обучаваните и обучаващите. Само неговото включване, интегриране по силата на конкретни правила и регламенти, може да гарантира някакъв вид устойчивост в моралните дилеми, които той носи със себе си. И в тази връзка е много логична препоръката на М. Шарпълс (Sharples, 2022, pp. 1124 – 1125) преподавателите и студентите да използват отговорно и ангажирано ИИ инструментите, вместо да се ограничават с рестрикции и забрани. Ясно регламентираното участие на ИИ в образователния процес би улеснило и подготовката, и оценяването.

**Втора теза: ИИ едновременно като инструмент и като субект на влияние**

Тази втора теза е основана на позицията, че ИИ в академичната област и в частност в хуманитарното знание може да функционира паралелно по два начина: като технологичен инструмент, който спомага и разширява изследователските и преподавателските опции, и като реално влияещ субект, като активен фактор, който модифицира академичните практики, който преформулира критериите за валидност.

Функцията му на научноизследователски и методологичен инструмент може да се детайлизира и обясни в следните примерни аспекти: ИИ автоматизира рутинните изследователски задачи и етапите на емпирично равнище, т.е. може да извлича метаданни, да резюмира, да търси в корпуси и да селектира, да оформя библиографски данни и др.; изкуственият разум също така подпомага и улеснява значително изследователската и преподавателската работа по отношение на редакцията и превода, скъсява времето за достъп до материали и до големи бази от данни.

Интересната, но и по-сложната за разрешаване и съответно за регламентиране е втората функция на ИИ. Това е така, защото тук зоните на неговото влияние са свързани не само с краткотрайни, спорадични и спомагателни ефекти, а с механизми, които позволяват на ИИ да изгражда нови парадигми на познание, да конструира нови методологии и по този начин да налага нови изследователски норми и канони, които го превръщат във фактор в академичния процес, без да бъде осъзнаван като такъв. В подкрепа на тази теза могат да се посочат различни примери от академичната образователна практика. Един от тях е ситуация, в която студент отправя въпрос към изкуствения „помощник“ по определена тема. В този случай чрез ИИ той получава предпоставена тематична парадигма. Друг пример може да е свързан с вече наложилата се практика, в която изследовател търси резюмета на статии по дадена тема, област, казус, феномен и пр. Тогава ИИ създава модел, който обаче избира, пресява, пренарежда и синтезира концепции. И двата примера показват как може да функционира ИИ: като посредник на знанието, който активно участва в изграждането на нова епистемична среда и който регламентира, предзадава нейното популяризиране и използване. Казано по друг начин, той вече не е

в позицията само и единствено на технически помощник, а вече е динамичен участник при вземането на решения и във формирането на позиции. Това положение логично го превръща в субект, който има възможността да структурира и да определя границите на достъпа до интелектуалното пространство, да предпоставя решения, избори, внушения и пр. В бъдеще това може да стане част от мощна политика в ръцете на финансови и властови кръгове.

Влиянието на ИИ може да се посочи и в намесата му около формирането на когнитивните навици в процеса на обучение. Можем да вземем за очевиден пример начина, по който студентите четат. Целенасоченото, внимателното и задълбочено четене, което се стреми да прозре и да възприеме важни знакови текстови структури и по този начин да достигне до по-дълбоки смисли, вече е „излишен академичен лукс“. На неговото място се е настанило практичното резюмирано, но и доста по-повърхностно четене, при което четящите се фокусират върху крайния резултат (бързо изпълнение на поставена задача).

„Следите“ от влиянието на ИИ прозират и в друг процес: хомогенизирането на научноизследователския стил. Чрез ИИ се създават модели, които „услужливо“ предлагат стандартизиран академичен еталон – за структура на научна статия, за композиция на дипломен проект или дисертационен труд, за рецензия на научна статия, студия, монография и пр. Остава само „лесното“ решение потребителят изследовател да се възползва от предоставения му „калъп“.

И двата примера водят към тезата, че ИИ вече е приел ролята на ментор, който задава и предлага генерирани ментални модели, а това е позиция, типична за лидерско влияние. Като допълнение идва и възможността на изкуствения разум да бъде равнопоставен участник в творчески и научноизследователски проекти, а именно: да генерира „прочити“, да структурира аргументи и контрааргументи, да поставя в центъра ключови теми и др. В никакъв случай тези негови действия не остават само на нивото на емпиричната помощ и „спестеното“ изследователско време. ИИ вече създава и репродуцира собствен интерпретативен изследователски ъгъл, който, разбира се, е формиран от наличните му бази от данни. Иначе казано, ИИ участва и направлява създаването на смисъл, а това е една от водещите функции в хуманитарното знание. Когато обаче неговите бази данни са ограничени, тогава всяка изследователска интерпретация е изправена пред риска от неадекватност. Оптимистичната гледна точка тук обаче е, че все пак крайното решение е в приоритетите на естествения човешки ум.

Позиционирането на ИИ в зоната на субективното влияние може да се види и в чисто технологичен аспект. Проблемът е, че той повтаря и съсредоточава вниманието върху автори, школи, течения и др., които са достатъчно добре представени в дигиталните архиви. И обратно – логично е да маргинализира такива, които се появяват рядко в базите данни. Именно

заради посочените зависимости от алгоритмите ИИ действа като дигитален механизъм за културна и/или образователна селекция, но и за сегрегация и *деформация*.

Участието на ИИ в образованието е свързано и с идеята за целенасоченото и ползотворно системно влияние, което увеличава шанса за подобряването на ученето чрез адаптивни и индивидуализирани модели (Zawacki-Richter et al., 2019). Тази констатация е обвързана с вече познатите интелигентни системи за обучение, които са най-надеждните предимства на ИИ (Peng, Ma & Spector, 2019). Очакванията в тази насока са свързани с появата на нова парадигма на адаптивно, персонализирано обучение, чиито основни елементи са: индивидуални характеристики, индивидуални постижения, личностно развитие и адаптивност (Peng, Ma & Spector, 2019, p. 6). Авторите твърдят, че „тази стратегия е неразделна от технологиите за овластяване“ и споменатите четири специфики могат да я развият „като ефективна педагогика, подкрепена от технологии, която може адаптивно да приспособява преподаването“ (Peng, Ma & Spector, 2019, p. 6). От гледна точна на тезата за ИИ като субект на влияние това изследване е пример как интелигентните системи за обучение могат да присъстват не само като инструменти, които подпомагат процеса на преподаване, но и могат и да го моделират според индивидуалните качества на обучаемите и по този начин се превръщат в активен фактор, който преформатира педагогическите практики.

Изводът, който може да се направи дотук, е, че в академичната сфера и особено в хуманитаристиката ИИ е субект на влияние не по силата на някакво имагинерно лидерско присъствие, а по-скоро защото притежава възможностите да структурира достъпни познавателни алгоритми, да задава нормативни принципи за валидация и важност, да формира когнитивни системни умения и методологични компетентности, да насочва целево вниманието към едни корпуси от текстове, анализи, източници и пр. и в същото време да изтласква в периферията други, да съучаства в продуцирането на интерпретации, наративи, проучвания, експерименти и пр. С голяма доза увереност може да бъде окачествен като културен и технологичен източник на идеи, модели, структури и др., който променя концепцията на хуманитарното знание и като цяло влияе върху академичната екосистема. За да обобщим, ще си послужим отново с идиоматичната метафора “elephant in the living room“. В контекста на казаното дотук може да се прибави и още една характеристика на „слона“ в нашия хол: освен, че е очевиден и познат наш гост, той вече е и институционално влиятелен. Проблемът е в нашата нагласа към него – неосъзнатото му (а може би и целенасочено) игнориране може да е причина за създаването на ситуации, в които се пренебрегват важни и фундаментални изисквания за изследователска и образователна автентичност, оригиналност и задълбоченост.

**Трета теза: Етичен и прозрачен ИИ**

Последната трета теза коментира възможностите и реално създадените условия за отговорно и етично обусловено използване на ИИ. Тук се включват вече създадени и прилагани (или отчасти прилагани) политики, които се опитват да регламентират рамките и вариантите на неговото приложение, да дадат някакви гаранции за яснотата и видимостта при работата с него, за критичност в методологията. Тези стратегии предлагат и дефиниции за границите на авторството и на авторските права, като и на създаденото оригинално (или не) творческо съдържание.

Изброените принципни позиции спрямо използването и прилагането на ИИ в образователния, научноизследователския, социалния, публичния и пр. контексти са заложени в *Препоръката* на ЮНЕСКО (UNESCO, 2022). Този документ е един от първите опити за създаване на регламент в световен план и фокусът в него е върху етичното използване на ИИ. Тъй като *Препоръката* е с глобално предназначение, насоките и предписанията в нея са общи и валидни за държавите и институциите. С текста си ЮНЕСКО цели да създаде и предложи една полезна регулаторна база, която да даде гаранции, че ИИ „ще работи за доброто на човечеството, индивидите, обществата, околната среда и екосистемите, както и за предотвратяване на вредата. Целта е също така стимулиране на мирното използване на AI системи“ (UNESCO, 2022, p. 14). *Препоръката* поставя акценти върху безопасността и сигурността (UNESCO, 2022, p. 14), прозрачността (UNESCO, 2022, p. 22), отговорността и отчетността (UNESCO, 2022, p. 22), осведомеността и повишаването на нивото на грамотността (UNESCO, 2022, p. 23). Не е пропуснато и важното уточнение, че ИИ е и реална възможност в развитието на научните знания, на академичната практика, но тази възможност води след себе си както ползи, така и ограничения и рискове и научните общности трябва да са наясно с тях (UNESCO, 2022, p. 35).

С регулацията на ИИ от различни перспективи се ангажира и *Регламентът на Европейския парламент и на Европейския съвет* от 13 юни 2024 г. (Регламент (ЕС)/ Reglament 2024/1689, 2024). В края на дългото си заглавие този обемен (повече от 140 страници) документ е назован *Закон/акт за изкуствения интелект* и всъщност това е първият законодателен текст с подробни насоки за регулация на ИИ на наднационално равнище. Прилагането на Регламента ще влезе в сила през лятото на 2026 г. и оптимистичният знак в този смисъл е, че страните членки на ЕС ще имат стимул и посока за изработването на собствени национални регламенти. Настоящото изложение няма за цел да се спира подробно на този документ, а само да посочи ключови акценти от него. Целта му е ориентирана в няколко посоки: „... да насърчи приемането на ориентиран към човека и надежден изкуствен интелект (ИИ)“ (Регламент/Reglament (ЕС) 2024/1689, 2024, с. 1); да „… гарантира високо равнище на защита на здравето, безопасността, фундаменталните права … срещу вредните ефекти от ИИ системите в Съюза и подпомогне иновациите“ (Регламент (ЕС)/

Reglament 2024/1689, 2024, с. 1). Въпреки че не е ориентиран конкретно към хуманитаристиката, социалните науки и културно-историческото наследство, в Регламента могат да се открояят и потенциални ползи за изследователските академични среди. Актът поставя стандарти с етичен и правен характер, изисквания за прозрачност, отчетност, надзор, респект към фундаменталните човешки права – тяхното прилагане гарантирано ще улесни разработването на изследователски проекти и проучвания.

През септември 2024 г. е публикувана *Рамковата конвенция за изкуствения интелект и правата на човека, демокрацията и върховенството на закона* (Council of Europe Framework Convention, 2024). Това е още една регулаторна инициатива на Съвета на Европа, в чиито приоритетни зони са включени: защитата на човешкото достойнство, на човешките равни права, защитата на личните данни, на неприкосновеността и ценността на човешкия живот, нетолерирането на дискриминацията и пр. *Рамковата конвенция* е с функциите на международен регулаторен акт с глобално значение и в този смисъл неговата цел е да предложи сигурни гаранции, че използването на ИИ няма да бъде в конфликт с правата на човека и тяхната неприкосновена защита. В положителна посока може да се коментира и фактът, че споразумението не се обвързва с конкретни технологии (предвид лавинообразната скорост на тяхната поява това е наистина практичен и умен ход), а поставя акцент върху константни морални принципи.

Цитираните и коментирани глобални документи комбинират правен контрол с етични стандарти, а вероятно само по този начин може да се осигури необходимият баланс между технологичните иновации, неприкосновената свобода и защитата както на човешките права, така и на моралните общественозначими ценности. На сходни принципи са стъпили и други международни регламенти, например *Глобалният дигитален пакт (Global Digital Compact)*, приет като анекс към *Споразумението за бъдещето (Pact for the Future)* на ООН от 2024 г. (United Nations, 2024, pp. 37 – 51). Този документ настоява да се определят общи принципи за „отворена, справедлива, устойчива и сигурна дигитална среда“, в която ИИ е управляван „в интерес на човечеството“ (United Nations, 2024, pp. 37 – 38). Визионерската му стойност е ясно видима в целта, която е поставена: да не се разчита на фрагментарни национални регулации, а на международен консенсус, който е необходим особено при глобални и трансгранични предизвикателства в съвместната работа на човешкия и изкуствения интелект. Предложен е и международен ангажимент, свързан с обучения, достъп до ресурси, инфраструктура, локални решения, особено за страни с ограничени ресурси (United Nations, 2024, p. 45). И тук, както и в останалите документи, се откроява идеята, че приложенията на ИИ са много, разнообразни и различни, а ефектът им върху човешките достойнство и права, както и върху демократичните стандарти като цяло може да варира. Разбира се, в полза на обективността е нужно да се спомене, че всички тези

документи имат и някои ограничения към настоящия момент. На първо място излиза фактът, че дори да са подписани на международно ниво (каквато е *Рамковата конвенция* на Съвета на Европа), те все още са в един предварителен етап и предстои тяхната ратификация, приемането на национално законодателство и пр. Друго ограничение е свързано с това, че някои документи (като например *Глобалният дигитален пакт*) не са юридически, а имат по-скоро политическа и/или морална тежест и в този смисъл тяхното прилагане може да се случи само с волевата инициатива на съответните държави. Тези и още някои съображения налагат извода, че глобалната сцена на действието, развитието и ефективното прилагане на ИИ все още е фрагментирана.

Подобна е ситуацията и в академичните среди. На проблемните зони в прилагането на ИИ в университетската практика и в науките са посветени много международни изследвания. Включително са структурирани и обзорни публикации с библиометрични показатели, които се стремят да оформят картографски изследователските търсения. Някои от тях са подредени по ключови думи, по конфликтни области, по зони и степени на приложение и влияние на ИИ, по класифициране на неговите разнородни роли във висшето образование (вж. например Bozkurt et al., 2021; Ouyang, Zheng & Jiao, 2022López-Chila et al., 2023). Вече са направени и важни изводи за резултатите и последващите реакции от приложението на ИИ в педагогическата дейност, за стратегическите механизми в подкрепа на обучението. Също така са поставени и важни проблеми около някои нововъзникнали (а може би позабравени) етични казуси и морални избори. Търсят се и адекватни решения в полза на отговорната работа с ИИ в университетите и в този аспект подходите са разнообразни и вариативни. На международно ниво те варират от създаването на централизирани университетски политики към структурирането на наръчници и правила на факултетно ниво и се стигне до документи, които са предназначени за конкретни специалности, дисциплини и пр. Разбира се, изброените в настоящото изложение подходи, методи, стратегии, форми и пр. варианти на академичен контрол и управление на използването и прилагането на ИИ нямат претенциите за изчерпателност. Изборът им не е целенасочен, на първо място е обусловен от тяхната видимост, т.е. от тяхната публична достъпност като дигитални ресурси, а на второ място той е свързан с факта, че университетите са международно разпознаваеми и авторитетни. Целта е преди всички да се маркират и анализират приложими практики, добре работещи модели, но и общи проблемни зони.

Както вече беше уточнено, в международните университетски политики около интегрирането на ИИ в образованието и научните изследвания се наблюдават различни подходи. Един от тях е децентрализираното и автономното използване на ИИ. Например университетът в Йейл регламентира достъпа и работата с ИИ, като се съобразява със спецификата на конкретната дисциплина. Добавено е обаче,

че „неоторизираното използване на ИИ се смята за академично нарушение“ (Yale University, 2023), което показва, че все пак съществуват някакви общовалидни критерии. По подобен начин университетът в Станфорд подчертава, че водещи в работата с ИИ са самостоятелността и институционалната независимост на преподавателите при вземаните от тях решения. Университетският наръчник за политиката на ИИ посочва, че те „са свободни да определят свои собствени правила за използването на генеративни ИИ инструменти“ (Stanford University, n.d.). И двете университетски ръководства се обединяват около идеята за академичната свобода, от една страна, и за персонализирания контрол, от друга. Полезният ефект на действие се търси в гъвкавостта спрямо различни и специфични характеристики за конкретни дисциплини. Вероятно може да се окаже проблем, че като че ли е свалена институционалната отговорност и е оставена в прерогативите на индивидуални избори и еднолични решения, които невинаги могат да бъдат обективни и в полза на човешкия интелект.

Някои университети предпочитат създаването на общоинституционални документи по защитата на личните данни, безопасността и прозрачността. Целта е да са посочени ясно критериите и те да са еднакви и общовалидни за цялата академична общност. По този начин по-сигурно се намалява рискът от злоупотреби, а едновременно с това работят в синхрон както етични, така и правни норми. Вероятно такава е била идеята около създаването на „Наръчник за безопасно и отговорно използване на инструментите на ИИ“ в Оксфордския университет (University of Oxford, 2025). Документът се ангажира с цялостна институционална подкрепа в полза на продуктивността и безопасната среда при използването на ИИ. Текстът е подробен, информативен, с ясно установени правила и права. Структуриран е в осем раздела с инструкции за етичното използване, за оценяването, за употребата на ИИ в изследователската дейност и в образователния процес. Присъстват и конкретни инструменти на ИИ, които университетът поддържа и до които студентите имат достъп чрез университетските си профили. Изрично е уточнено, че работата с тях е възможна след консултация с преподавателите. Оксфордският наръчник е добър пример за балансираното прилагане на обективни рестрикции, но и на добре обмислени и премерени разрешени употреби на ИИ в академичната образователна среда.

Върху институционалната политика за защита на личните данни стъпва и „Ръководство за използване на инструментите на генеративния ИИ“ на Масачузетския институт по технологии (Massachusetts Institute of Technology, n.d.). Включено е задължителното изискване за откритост и ясна декларация при използването на ИИ.

Важна инициатива за безпроблемното и отговорно използване на ИИ е включването на дигиталната грамотност в образователния процес. На тази основа, комбинирана с ясни критерии за академична добросъвестност и почтеност, са изработени някои университетски регламенти за използването

на ИИ. Например в сайта на University College в Лондон (UCL) присъстват достатъчно подробни и ясни, конкретни насоки за използването на ИИ в обучението и в последващото оценяване (University College London, n. d.). На първо място по значимост е разделът „Генеративният ИИ и академична почтеност“. Важно е обаче да се отбележи, че в сайта е отделено място и на специален обучителен онлайн модул по академична добросъвестност. Предлага се и актуална информация за други видове измами и форми на академично неправомерно поведение и пр. Акцентът като цяло е върху етичната гледна точка и върху развиването на политики за академична честност и добросъвестност. Подобна е позицията на най-стария университет във Франция – Сорбоната, чийто регламентиращ използването на ИИ текст е със заглавие: „Изкуственият интелект и аз в университета“ (Sorbonne Université, 2024). По много неформален, стилово близък до обучаемите начин, който проличава още от заглавието на документа, са посочени възможните и разрешени употреби на изкуствения разум. Под формата на въпроси и отговори в 1 л. ед. ч. ненатрапчиво и елегантно са поставени акценти върху информираността, критичното мислене и проверката на резултатите. Не забраните и санкциите са водещи перспективи, а отговорното използване като следствие от дигиталната грамотност и научната добронамереност.

Разликите в цитираните документи се отнасят повече към зоните на тяхното приложение, към делегираните отговорности в използването на ИИ ресурсите. При всички случаи обаче генералната тенденция ги обвързва в общия стремеж към ясни граници, към етичност и към академична почтеност. Предписани са и някои ограничителни действия срещу плагиатството, дезинформацията и срещу обезценяването на авторството и креативността във всичките ѝ форми. Във всички прегледани документи обаче ясно проличават подходи, които целят интеграцията на ИИ, а не абсолютното му отричане.

**Заключение**

Пределно ясно е, че за изкуствения интелект в академичната общност има още какво да се научи от самата академична общност и това вече се случва, както настоящият текст се опита да покаже. Ясно е също така, че правилният път за университетските структури е в отделянето на достатъчно адекватни ресурси от всякакво естество (интелектуално, финансово, морално и пр.) за проучването на ИИ в неговите разнообразни форми. За да изпълни изследователската и образователната си мисия, е необходимо висшето образование да бъде в проактивна позиция, а не в пасивното, дистанцирано наблюдение на „слона в хола“ и в очакването за спуснати отнякъде инструкции за „правилна“ употреба на ИИ в научните изследвания или в образователните дейности. Защото всъщност „слонът е в хола“ (дори и в случаите на видима неосъзнатост) и все повече ще се трансформира във видима част от реалността. Не е необходимо създаването и

мултиплицирането на драматични, апокалиптични сюжети, по-важно е да се осъзнае, че новата реалност налага и промяна на възгледите за същността на инструментите, които формират компетенциите на съвременния човек. Все по-отчетливо се налага необходимостта „не само да можем да използваме ИИ (дори когато не го осъзнаваме), но и повече да разбираме неговите широки приложения, полезност, ограничения и пристрастия“ (Southworth et al. 2023, p. 3).

**БИБЛИОГРАФИЯ:**

**Регламент (ЕС) 2024/1689 (2024).** Регламент (ЕС) 2024/1689 на Европейския парламент и на Съвета от 13 юни 2024 година за установяване на хармонизирани правила относно изкуствения интелект и за изменение на регламенти (ЕО) № 300/2008, (ЕС) № 167/2013, (ЕС) № 168/2013, (ЕС) 2018/858, (ЕС) 2018/1139 и (ЕС) 2019/2144 и директиви 2014/90/ЕС, (ЕС) 2016/797 и (ЕС) 2020/1828 (Акт за изкуствения интелект). *Официален вестник на Европейския съюз,* L 1689, <https://eur-lex.europa.eu/legal-content/BG/TXT/PDF/?uri=OJ:L_202401689> (Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance). (2024). *Official Journal*, L 1689. <http://data.europa.eu/eli/reg/2024/1689/oj>).

**Bozkurt, A., Karadeniz, A., Baneres, D., Guerrero-Roldán, A. E. & Rodríguez, M. E. (2021).** Artificial intelligence and reflections from educational landscape: A review of AI studies in half a century. *Sustainability*, 13 (2) (2021), 800, <https://doi.org/10.3390/su13020800>.

**Council of Europe Framework Convention (2024).** Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (CETS No. 225), <https://rm.coe.int/1680afae3c>.

**Grassini, S. (2023).** Shaping the Future of Education: Exploring the Potential and Consequences of AI and ChatGPT in Educational Settings. *Education Sciences*, *13*(7), 692. https://doi.org/10.3390/educsci13070692.

**Johnston, H., Wells, R. F., Shanks, E. M., Boey, T. & Parsons, B. N. (2024).** Student perspectives on the use of generative artificial intelligence technologies in higher education. *International Journal for Educational Integrity*, 20, 2, 1-21. https://doi.org/10.1007/s40979-024-00149-4.

**López-Chila, R., Llerena-Izquierdo, J., Sumba-Nacipucha & N. J. Cueva-Estrada (2023).** Artificial intelligence in higher education: An analysis of existing bibliometrics. *Education in Science*, 14 (1), 47. https:// doi.org/10.3390/educsci14010047.

**Massachusetts Institute of Technology (n.d.).** Guidance for use of Generative AI tools. Avalable at: <https://ist.mit.edu/ai-guidance>.

**Ouyang, F., Zheng, L. & Jiao, P. (2022).** Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020. *Education and Information Technologies*, 27 (6), pp. 7893-7925. https://doi.org/10.1007/s10639-022-10925-9.

**Peng, H., Ma, S., & Spector, J. M. (2019)**. Personalized adaptive learning: An emerging pedagogical approach enabled by a smart learning environment.// *Smart Learning Environments*, 6 (1), 1 – 14.

**Sharples, M. (2022).** Automated essay writing: An AIED opinion. // *International Journal of Artificial Intelligence in Education*, 32 (4), 1119 – 1126. https://doi.org/10.1007/s40593-022-00300-7.

**Sorbonne Université (2024).** L'IA et moi à l'université. Avalable at: <https://www.pantheonsorbonne.fr/formation/lia-et-moi-luniversite>.

**Southworth, J., Migliacco, K., Glover, J., Reed, D., McCarty, C., Brendemuhl, J. & Thomas, A. (2023).** Developing a model for AI Across the curriculum: Transforming the higher education landscape via innovation in AI literacy.// *Computers and Education: Artificial Intelligence*, 4, 100127; https://doi.org/10.1016/j.caeai.2023.100127.

**Stanford University (n.d.).** Generative AI Policy Guidance, <https://communitystandards.stanford.edu/generative-ai-policy-guidance>.

**UNESCO (2022).** Recommendation on the Ethics of Artificial Intelligence, <https://unesdoc. unesco.org/ark:/48223/pf0000381137/PDF/381137eng.pdf.multi>.

**United Nations (2024).** Global Digital Compact. – In: Pact for the Future, Global Digital Compact and Declaration on Future Generations. Adopted at the Summit of the Future, pp. 37–51., <https://www.un. org/sites/un2.un.org/files/sotf-pact_for_the_future_adopted.pdf>.

**University College London (n.d.).** Engaging with Generative AI in your education and assessment <https://www.ucl.ac.uk/students/exams-and-assessments/assessment-success-guide/engaging-generative-ai-your-education-and-assessmentKU Leuven>.

**University of Oxford. (2025).** Guidance on safe and responsible use of GenAI tools. <https://www.ox.ac.uk/students/life/it/guidance-safe-and-responsible-use-gen-ai-tools>.

**Wang, H., Dang, A., Wu, Z., Mac, S. (2024).** Wang, H., Dang, A., Wu, Z., & Mac, S. (2024). Generative AI in higher education: Seeing ChatGPT through universities' policies, resources, and guidelines. *Computers and Education: Artificial Intelligence*, *7*, 1–13. Avalable at: <https://arxiv.org/pdf/2312.05235v3>.

**Yale University (2023).** Guidelines for the Use of Generative AI Tools. <https://provost.yale.edu/news/guidelines-use-generative-ai-tools>.

**Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019).** Systematic review of research on artificial intelligence applications in higher education–where are the educators?. // *International Journal of Educational Technology in Higher Education*, 16 (1). https://doi.org/10.1186/s41239-019-0171-0.

Copyright © 2026 Ivanova. This is an open-access publication distributed under the terms of the Creative Commons Attribution 4.0 International Licence CC BY 4.0