

**ПОЛИТИКА
ЗА ПРИЛОЖЕНИЕ НА ИЗКУСТВЕН ИНТЕЛЕКТ
В ТЕХНИЧЕСКИЯ УНИВЕРСИТЕТ – СОФИЯ**

Преамбюл

В съвременната академична и научна среда изкуственият интелект (ИИ) предлага значителни възможности за подобряване на образованието, научните изследвания, административните процеси и ежедневната дейност на висшите училища. В същото време, използването на ИИ поставя предизвикателства, свързани с академичната почтеност, етиката, защитата на личните данни и интелектуалната собственост.

Технически университет – София (ТУ - София), като водеща академична и научна институция, признава потенциала на ИИ за създаване на иновации и повишаване на качеството на обучение и изследвания. В този контекст Университетът счита за необходимо да установи ясни и отговорни насоки за използването на ИИ от страна на преподаватели, студенти, изследователи и служители.

Настоящата политика има за цел да гарантира, че технологиите с изкуствен интелект се използват по прозрачен, етичен и справедлив начин, в съответствие със законовите и регулаторни изисквания, както и с ценностите на ТУ-София – академична честност, отговорност и уважение към личната и научната неприкосновеност. Политиката цели да насърчи информираното и отговорно прилагане на ИИ в подкрепа на развитието на академичната общност, без да се компрометират основните принципи на висшето образование и научната етика.

Дефиниции

Изкуствен интелект	Изкуствен интелект (ИИ) е използването на цифрови технологии за създаване на системи, които могат да изпълняват задачи, които обикновено изискват човешка намеса. ИИ имитира човешкото мислене, но обработва информацията по-бързо и по-точно.
Системи с изкуствен интелект	„система с ИИ“ означава машинно базирана система, която е проектирана да работи с различни нива на автономност и която може да прояви адаптивност след внедряването си, и която, с явна или подразбираща се цел, въз основа на въведените в нея входящи данни, извежда начина на генериране на резултати като прогнози, съдържание, препоръки или решения, които могат да окажат влияние върху физическа или виртуална среда.
Академична почтеност	Академичната почтеност е качество, което демонстрира ценности като честност, доверие, справедливост, отчетност, грижа и уважение към другите и тяхната академична работа.
Конфиденциална информация	Конфиденциална информация означава всяка бизнес или техническа информация или резултат от изследвания, принадлежащи на ТУ-София, на член от работещите и обучаемите в ТУ-София, на сътрудници или на други трети страни, които не са публично известни или които са били предоставени или получени при задължение за запазване на информацията като поверителна. Това включва и защитена здравна информация.
Лична информация	„Лични данни“ означава всяка информация, свързана с идентифицирано или идентифицируемо физическо лице. Идентифицируемо лице е всяко лице, което може да бъде идентифицирано, пряко или косвено. Различните елементи от информацията, които събрани заедно, биха могли да доведат до идентифицирането на конкретно лице, също представляват лични данни. Лични данни са съгласно определението в член 4, точка 1 от Регламент (ЕС) 2016/679.
Входящи данни	„входящи данни“ означава данни, предоставяни на система с ИИ или пряко придобивани от нея, въз основа на които системата произвежда резултат.

I. Общи положения

1. ТУ-София желае да предостави насоки в тази политика, като вземе предвид позицията на професионалните, законовите и регулаторните органи, свързана с това как ИИ може да бъде от полза за изследователите, студентите и преподавателите, без да нарушава академичната почтеност.

2. Настоящата политика за приложение на ИИ урежда използването, създаването, модифицирането и внедряването на технологии и приложения с изкуствен интелект от служители, преподаватели, студенти и изследователи при изпълнение на техните функции за или от името на ТУ-София. Тъй като този документ може да бъде актуализиран на определени интервали от време, работещите и обучаемите в ТУ-София се насърчават редовно да преглеждат най-новата версия на тази Политика. Настоящата политика има за цел да установи правила и принципи за етичното и отговорно използване на системи с ИИ в рамките на ТУ-София, като гарантира прозрачност, справедливост и зачитане на основните права.

3. Изразът "изкуствен интелект (ИИ)" се използва в този документ, за да се подчертае основният термин, при който системи с изкуствен интелект могат да се прилагат и/или разработват в основните дейности на Техническия университет – София (ТУ-София).

4. Принципи на приложението на изкуствения интелект:

а) Университетът ще подобрява дигиталните възможности на студентите и персонала за да им даде възможност да разберат предимствата и недостатъците на използването на генеративен ИИ.

б) Университетът ще насочва студентите и персонала към приемливото и неприемливо използване на ИИ в студентския опит за да гарантира, че академичната строгост и почтеност се поддържат.

в) Университетът ще гарантира, че ИИ се възприема етично за да подобри възможностите за преподаване и учене и да подпомогне оценяването. Университетът ще продължава да проучва най-добрите практики за изкуствен интелект и да ги споделя във всички програми.

5. Всеки потребител на ИИ в университета трябва да бъде запознат с потенциалните рискове и ограничения на технологията.

6. Използването на ИИ не може да замества критичното мислене и аналитичните способности на студентите и преподавателите.

II. Основни принципи за работа със системи с изкуствен интелект

7. Да се избягва въвеждането на конфиденциална или лична информация в ИИ системи, освен ако има ясни технически и правни гаранции за сигурността на данните или е получено информирано съгласие. Насърчава се използването на анонимизирани или синтетични данни при възможност.

8. Да не се въвежда лична информация: Работещите и обучаемите в ТУ - София не трябва да въвеждат каквато и да е лична информация за служители, студенти, преподаватели или други заинтересовани страни в инструменти, използващи изкуствен интелект, освен когато това е разрешено от съответните отговорни лица.

Да не се въвежда информация, която нарушава интелектуалната собственост: Работещите и обучаемите в ТУ - София трябва да са наясно с условията, при които използват инструментите за ИИ. Всички работещи и обучаеми трябва да зачитат правата на интелектуална собственост с цел защита на тези права. Задължение на отделните потребители е да гарантират, че входните и изходните данни на техните системи с ИИ са надлежно защитени по причини като закони за авторското право и патентите, разпоредби за защита на данните и престъпления, свързани с кражба на самоличност. Такива инструменти за ИИ могат да включват големи езикови модели (например ChatGPT), платформи за машинно обучение (например AWS SageMaker), софтуер за разпознаване на изображения (например Google Cloud Vision) или други.

Да се потвърждава точността на резултатите, предоставени от генеративен изкуствен интелект и други подобни инструменти: Работещите и обучаемите в ТУ - София трябва да проверяват точността на информацията от системи с изкуствен интелект, преди да разчитат на тази информация. Възможно е генерираното от ИИ съдържание да бъде неточно, пристрастно или изцяло измислено (понякога наричано "халюцинации"). Да се има предвид, че такова генерирано от ИИ съдържание може да съдържа материали, защитени с

авторски права. Ползвателят на генерираните данни носи отговорност за всяко съдържание, което публикува и което включва материал, генериран от ИИ.

Да се проверяват резултатите от системи с изкуствен интелект и други подобни инструменти за пристрастност: Работещите и обучаемите в ТУ - София трябва да преценят дали данните, въведени в системи с изкуствен интелект и други подобни инструменти, както и изходът от тях, произвеждат решения, които могат да доведат до неравностойно въздействие върху лица съгласно приложимото законодателство, като раса, етническа принадлежност, национален произход, възраст, сексуална ориентация, статус на увреждане и др.

Да се оповестява използването на инструменти за системи с изкуствен интелект: Работещите и обучаемите в ТУ - София, които използват системи с изкуствен интелект за изготвяне на писмени материали или други работни продукти, трябва да оповестят, че тези материали и този работен продукт се основават на или произтичат от използването на ИИ. Винаги да бъде прозрачно, ако се разчита на резултати от подобни инструменти.

Да се спазват правата на интелектуална собственост на трети страни: Работещите и обучаемите в ТУ - София не трябва да представят резултатите, генерирани чрез инструменти на ИИ, като свои собствени. Ако се цитират, перифразират или заимстват идеи от резултати, да се потвърди, че те са точни и че не плагиатстват съществуваща работа на друга страна или не нарушават по друг начин правата на интелектуална собственост на трета страна.

Да не се използват системи с изкуствен интелект за създаване на злонамерено съдържание: На работещите и обучаемите в ТУ - София е забранено да използват системи с изкуствен интелект и други подобни технологии за генериране на злонамерено съдържание, като например зловреден софтуер, вируси, софтуерни червеи и троянски коне, които могат да заобиколят мерките за контрол на достъп, въведени от ТУ-София или от трета страна, за да се предотврати неоторизиран достъп до съответните мрежи.

Управление на личните данни за обучение на модели: Когато е възможно, се препоръчва използването на настройките, с които се ограничава използването на въведените данни за по-нататъшно обучение на модела. Университетът ще предостави насоки за това как да се конфигурират такива системи.

III. Изкуствен интелект и академична почтеност

В ТУ-София, обща отговорност на всички преподаватели, служители, студенти, докторанти и специализанти е да се насърчава интелектуалната честност и научната почтеност, които могат да бъдат подкопани с използването на съдържание, генерирано от изкуствен интелект, което се представя за собствена разработка.

Препоръки за използването на ИИ от студенти:

а) При липса на ясно изявление от преподавателя, студентите следва да се консултират относно допустимостта на използване на ИИ. Университетът насърчава преподавателите да посочват конкретни рамки, в които ИИ може да бъде използван за обучение и развитие на умения.

б) Студентите се насърчават да разговарят с преподавателите си относно техните очаквания за използването на инструменти с ИИ.

в) В учебни програми, където това е уместно, използването на ИИ ще бъде не само разрешено, но и насърчавано като част от развиването на дигитални компетентности и критично мислене. Студентите ще бъдат обучавани как отговорно да използват ИИ, включително как да анализират, валидират и цитират генерирано съдържание.

г) Студентите могат да използват ИИ за академични цели, при условие, че ясно декларират кога и как са го използвали.

Препоръки за използването на ИИ от преподаватели:

а) Като минимум се препоръчва преподавателите да споделят ясни очаквания в началото на всеки семестър чрез учебната програма, разпространението на правилата и обсъждането в занятията относно подходящото използване на инструменти за ИИ. Преподавателите следва да насърчават студентите да се обръщат към тях, когато имат нужда от подкрепа, вместо да рискуват потенциално нарушение на академичната почтеност, както и да признават и цитират всяко използване на приложения на ИИ.

б) Преподавателите се насърчават да експериментират с използването на ИИ в преподаването, включително за разработване на учебно съдържание,

интерактивни помощници и автоматизирана обратна връзка, при спазване на етичните принципи.

в) Преподавателите се насърчават да разработят политика на всеки курс относно използването на инструменти с ИИ и какво преподавателите смятат за подходящо и неподходящо в техните занятия.

г) Да се консултират с Дирекция образователни технологии при необходимост относно насоки и съображения за инструментите на ИИ, позволени за използване.

д) Да се използват инструменти за откриване на съдържание, генерирано от ИИ, при проверка на курсови работи, курсови проекти и дипломни работи.

е) Използването на ИИ за автоматизиране на административни задачи е допустимо, ако не противоречи на действащите университетски политики.

Препоръчва се провеждането на обучения и семинари за правилно използване на ИИ в академичната дейност.

IV. Изкуствен интелект и научноизследователска и развойна дейност

Лицата, които използват системи с ИИ в научните изследвания, трябва да бъдат прозрачни по отношение на използването му в описанието на методите, благодарностите или другаде, както е уместно.

Установено е, че системите с ИИ, наред с други неща, генерира цитати на несъществуващи статии от несъществуващи автори, а също така се използва за генериране на изображения за експерименти, които никога не са били провеждани. Изследователите носят отговорност за точността на всяко съдържание, създадено от ИИ, което е включено в резултатите от изследването.

От изследователите се очаква да следват политиките на списанията, финансиращите органи и професионалните общества, чрез които докладват за своите изследвания.

Изследователите следва да оценят риска от въвеждане на непубликувани данни в ИИ системи, като при наличие на договорни или правни рискове използват вътрешни, защитени версии на ИИ системи. Насърчава се използването на локални модели или модели с възможност за деактивиране на обучението.

Изследователите трябва да избягват да въвеждат или да използват като входни данни поверителна информация на други страни. Системите с ИИ може да не осигуряват защита на Конфиденциална информация и използването им може да създаде потенциал за нарушаване на поверителни договорни ангажименти. Това включва напр. непубликувани ръкописи или предложения за финансиране, които изследователите може да бъдат помолени да рецензират. Това включва и личната информация на участниците в изследванията.

Изследователите трябва да имат предвид, че резултатите от системи с изкуствен интелект могат да нарушат правата на трети страни, тъй като генерираните отговори са събрани от вече създадени произведения.

Да се спазват изискванията за Високорискови системи с ИИ по член 6, параграф 2, от РЕГЛАМЕНТ (ЕС) 2024/1689 НА ЕВРОПЕЙСКИЯ ПАРЛАМЕНТ И НА СЪВЕТА от 13 юни 2024 година.

Разработването на системи с ИИ, различни от високорискови системи с ИИ, да бъде в съответствие с изискванията на РЕГЛАМЕНТ (ЕС) 2024/1689 НА ЕВРОПЕЙСКИЯ ПАРЛАМЕНТ И НА СЪВЕТА от 13 юни 2024 година.

V. Санкции при нарушение

За спазването на политиката и разрешаване на възникнали с приложението й казуси се сезира Комисията по етика, избрана от АС на ТУ-София и назначена от ректора.

Всички констатирани нарушения на настоящата политика се разглеждат от Комисията по етика на ТУ-София.

При неспазване на споменатите в политиката норми и принципи на използване на ИИ се носи дисциплинарна отговорност съгласно Етичния кодекс на ТУ – София и Кодекса на труда.

В случай на академични измами, извършени с помощта на ИИ, студентите могат да получат наказания като анулиране на задания, понижаване на оценки или отстраняване от ТУ – София.

Санкциите могат да включват и временно ограничаване на достъпа до ИИ инструменти в рамките на университета.

При постъпване на работа в университета прекият ръководител е длъжен да запознае служителя с разпоредбите на тази политика.

При всяка промяна в тази политика, утвърдена от Академичния съвет на ТУ–София, всички ръководители са длъжни да запознаят пряко подчинените си служители с разпоредбите на актуализираната политика в седемдневен срок.

Окончателно решение по въпросите от настоящата политика взема ректорът на ТУ–София.

VI. Заключение разпоредби

Настоящата политика е разработена в съответствие със:

- а) Закона за висшето образование;
- б) РЕГЛАМЕНТ (ЕС) 2024/1689 НА ЕВРОПЕЙСКИЯ ПАРЛАМЕНТ И НА СЪВЕТА от 13 юни 2024 година;
- в) Правилника за устройството и дейността на ТУ – София;
- г) Етичен кодекс за поведение на академичния и преподавателския състав на Технически университет – София;
- д) Други действащи нормативни актове в Република България;
- е) Настоящата политика може да бъде допълвана с вътрешни наредби и инструкции, които да регулират конкретни аспекти на използването на ИИ.

Настоящата политика влиза в сила от датата на одобрението му от АС на ТУ - София.