



РЕПУБЛИКА БЪЛГАРИЯ

Институт по публична
администрация



ИНСТИТУТ
ПО ПУБЛИЧНА
АДМИНИСТРАЦИЯ

ВЪТРЕШНИ ПРАВИЛА ЗА ИЗПОЛЗВАНЕ НА ИЗКУСТВЕН ИНТЕЛЕКТ В ИНСТИТУТА ПО ПУБЛИЧНА АДМИНИСТРАЦИЯ

2025

I. ОБЩИ ПОЛОЖЕНИЯ

Член 1. Настоящите правила уреждат използването на изкуствен интелект (ИИ) в дейността на Института по публична администрация (ИПА).

Примери за технологии на ИИ:

- 1) големи езикови модели (ГЕМ/Large Language Models/LLM), general pre-trained transformers (GPT) трансформери/трансформатори;
- 2) трансформатори/трансформери, които създават изображения въз основа на текстови или графични подсказки;
- 3) инструменти за превод и транскрипция, базирани на ГЕМ/LLM.

Член 2. Целта на тези правила е да се гарантира безопасното, етично и законосъобразно използване на ИИ системи в административните процеси на ИПА.

Член 3. Настоящите правила се прилагат за всички служители на ИПА, както и за външни организации, които използват ИИ технологии в рамките на съвместни проекти с Института.

II. ОСНОВНИ ПРИНЦИПИ НА ИЗПОЛЗВАНЕ НА ИИ

Член 4. Използването на ИИ в ИПА следва следните принципи:

1. Прозрачност. Потребителите трябва да знаят кога взаимодействат с ИИ, а не с човек. Ако ИИ генерира отговори на въпроси, това трябва да бъде ясно обозначено.
2. Отговорност. Човешкият надзор и експертиза са задължителни за всички решения, взети при използване на ИИ.
3. Сигурност и поверителност. Личните данни трябва да бъдат защитени в съответствие със Закона за защита на личните данни и на Общия регламент относно защитата на данните както и други приложими нормативни актове.
4. Недопускане на дискриминация. ИИ системите не трябва да водят до дискриминация въз основа на раса, пол, възраст и др.
5. Оценка на риска. Всички ИИ системи трябва да преминат предварителна проверка за рискове.
6. Киберсигурност: системите за ИИ се използват чрез профили, специално създадени за целта, които не са свързани с лични пощи. Където е необходимо, се координира създаване и ползване на профили за конкретната ИИ-система/ИИ-инструмент. Използват се само за целите на дейността на ИПА, не се предоставят за ползване и не се ползват на лични устройства.

III. ПРАВИЛА ЗА ВНЕДРЯВАНЕ НА ИИ В ИПА

Член 5. ИИ системите, използвани в ИПА, трябва да бъдат класифицирани по нивото на риск:

1. Минимален или никакъв риск:
 - 1.1. употреба на системи за филтриране на спам, социалните медии и интернет съдържанието, където специфични алгоритми разграничават и филтрират информация, за да я ориентират според конкретния потребител.;
 - 1.2. инструменти, задвижвани с ИИ за форматиране и редактиране на текст.
2. Ограничен риск
 - 2.1.1. употреба на чатботове;
 - 2.1.2. системи за превод на съдържание;
 - 2.1.3. инструменти за визуализация на данни.

3. Висок риск

- 3.1. ИИ системи, използвани за обработка и анализ на лични данни на служители, лектори и обучаеми;
- 3.2. ИИ системи, подпомагащи административни дейности, които засягат правата и свободите на служителите и гражданите.

4. Неприемлив риск

- 4.1. системи с ИИ за когнитивно-поведенческа обработка, прогнозиране в полицейската област, разпознаване на емоции на работното място и в образователните институции, както и социално оценяване;

Член 6. Всяка ИИ система трябва да премине предварителна проверка за съответствие с етичните и правните стандарти и да бъде регистрирана в база данни за мониторинг. Проверката се извършва спрямо Приложение 1.

IV. ОБУЧЕНИЕ И МОНИТОРИНГ

Член 7. Всички служители на ИПА преминават обучение за работа с ИИ технологии и принципите на етично използване.

Член 8. Създава се работна група за мониторинг и контрол на ИИ системите, която извършва периодични проверки и оценки на риска.

Член 9. ИПА поддържа регистър на внедрените ИИ системи в съответствие с изискванията на европейското и националното законодателство (Приложение 2).

V. САНКЦИИ И КОНТРОЛ

Член 10. В случай на неправомерно използване на ИИ в ИПА, ще се прилагат следните санкции:

1. Официално предупреждение за служители, които не спазват правилата.
2. Ограничаване или прекратяване на достъпа до ИИ системите.
3. Докладване на неправомерни практики на съответните надзорни органи.

VI. ИИ В ДЕЙНОСТТА НА ИПА

Член 11. Използването на ИИ в дейността на ИПА следва следните правила:

1. ИИ може да се използва за анализ на учебни материали, автоматизирани отговори на често задавани въпроси, обратна връзка от обучения от анкети и др., оценка на лекторите, създаване на персонализирани обучителни материали, както и за подпомагане на създаването на презентации и учебни ресурси.
2. При използване на ИИ за генериране на съдържание, за целите на дейността на ИПА, служителите в ИПА и лекторите на ИПА трябва да посочват дали текстовете, анализите и графиките са изцяло създадени от ИИ или са редактирани от човек. Всички ИИ генерирани материали трябва да бъдат проверени за точност и съответствие с официалните източници. При създаване на презентации с помощта на ИИ, трябва да се осигури балансиран и обективен подход, като се избягва изопачаване на фактите.
3. Служителите в ИПА и лекторите на ИПА носят пълна отговорност за проверката на достоверността и приложимостта на материалите, генерирани чрез ИИ. Те трябва да гарантират, че съдържанието не включва подвеждаща или неточна информация.
4. Не се допуска използване на ИИ за създаване на подвеждащо, манипулативно или неетично съдържание за публикуване, обучителни и информационни цели.
5. Служителите в ИПА и лекторите на ИПА трябва да информират всички заинтересовани страни (служител на ИПА, участници в обучения, лектори, мениджъри

на обучения, членовете на екипа по качеството на обучението), ако използват ИИ за генериране на съдържание.

6. При автоматизирани тестове и оценки, генерирани с ИИ, трябва да се осигури човешки контрол за валидиране на резултатите. Не се допуска използването на ИИ за автоматично генериране на оценки без участие на служител на ИПА.

7. При използването на ИИ за анализ на обратна връзка и резултати от обучаемите се гарантира, че личните данни са защитени и обработвани съгласно действащите нормативни актове и са напълно анонимизирани.

8. Екипът по качество на обучението преценява дали използването на ИИ от лекторите на ИПА за генериране на учебително съдържание съответства на настоящите правила.

8. Лекторите на ИПА задължително спазват „Етични правила и насоки относно използването на ИИ в учебителните дейности в ИПА“.

ДОПЪЛНИТЕЛНИ РАЗПОРЕДБИ

§1. Дефиницията на термините „лектор“, „обучаем“, „екип по качеството на обучението“, „мениджър на обучението“ е определена във Вътрешни правила за организация на учебната дейност и управление на качеството на учебния процес в Института по публична администрация

ЗАКЛЮЧИТЕЛНИ РАЗПОРЕДБИ

§1. Настоящите правила влизат в сила с тяхното утвърждаване от Изпълнителния директор

§2. Вътрешните правила се актуализират при необходимост, както и промяна в законодателството.

§3 В изпълнение на настоящите правила се утвърждават „Етични правила и насоки относно използването на ИИ в учебителните дейности в ИПА“.

ФОРМУЛЯР ЗА ОЦЕНКА НА РИСКА НА ИИ СИСТЕМИТЕ

Обща информация за ИИ системата

1. Име на системата:
2. Цел на използването:
3. Отговорно звено/служител:
4. Доставчик на ИИ технологията:
5. Дата на внедряване:

Оценка на риска

- Класификация по ниво на риск:
 - Минимален риск
 - Ограничен риск
 - Висок риск
 - Неприемлив риск

Анализ на риска

- Как системата влияе върху служителите/обучаемите/лекторите?
 - ИИ може ли да доведе до дискриминация?
 - Какви механизми за човешки надзор са предвидени?
1. Прозрачност
 2. Защита на личните данни
 3. План за мониторинг и одит
 4. Заключение и препоръки

Приложение 2

**РЕГИСТЪР НА СИСТЕМИ С ИЗКУСТВЕН ИНТЕЛЕКТ, ИЗПОЛЗВАНИ В ИПА,
ПО НИВО НА СИГУРНОСТ, ТИП И ПРИЛОЖЕНИЕ**

№	ИИ-система	Тип	Ниво на риск	Приложение в дейностите на ИПА
1				
2				
3				
4				
5				