

**ADELE: Информационна платформа за анализ на съдебна практика чрез методите на изкуствения интелект**

# Предложение за хармонизирани правила относно изкуствения интелект в ЕС



Този проект е финансиран от Програма „Правосъдие“ (2014-2020) на Европейския съюз по Споразумение за отпускане на БФП № 101007420.

# Ролята на ЕС в регулирането на изкуствения интелект

- Съобщение на Комисията до Европейския парламент, Съвета, Европейския икономически и социален комитет и Комитета на регионите „Изкуствен интелект за Европа“ , Брюксел, 25.4.2018 г., COM(2018) 237 final
- Определение за ИИ: основни възможности и предизвикателства – Експертна група на високо равнище по въпросите на изкуствения интелект
- Насоки относно етичните аспекти за надеждния изкуствен интелект – високо равнище по въпросите на изкуствения интелект
- Списък за оценяване на надеждния изкуствен интелект
- Бяла книга за изкуствения интелект – атмосфера на доверие (2020 г.)
- Доклад относно последиците от изкуствения интелект и роботиката за безопасността и отговорността (2020 г.)

Европа, подготвена за цифровата ера: Комисията предлага нови правила и действия за високи постижения и доверие в изкуствения интелект, 21.04.2021 г.

# Фокус върху изграждането на доверие в ИИ

- Европейска правна рамка за ИИ за адресиране на основните права и рискове за безопасността, специфични за системите с ИИ
- Правила на ЕС за разрешаване на проблемите с отговорността, свързани с нови технологии, включително системи с изкуствен интелект
- Преразглеждане на секторното законодателство за безопасност (например, регламент относно машиностроителните изделия, относно общата безопасност на продуктите)

## В основата...

- ИИ не може да бъде отговорен или безотговорен
- ИИ е произведен резултат – артефакт, направен от човека
- Хората са изначално отговорни, пристрастни и несправедливи в своето поведение



- Проблемът е как да накараме хората да отчитат своите действия – както като физически лица, така и като компании

- Какво е ИИ? Как регулираме въпрос, по който не можем да дадем еднозначни дефиниции? [...]
- Спира ли или насърчава регулацията иновациите и напредъка? [...]
- Не са ли рисковете и съображенията прекалено различни за отделните секторни приложения, за да говорим за общ регулаторен подход? [...]
- Възможно ли е регулиране чрез стимули или инвестиционни програми, които подтикват организациите да следват определен тип приложения или технологични подходи? [...]

# Проект на Законодателен акт за изкуствения интелект

- Да гарантира, че системите с ИИ, които се пускат на пазара на ЕС и се използват, са безопасни и в съответствие с действащото законодателство в областта на основните права и ценностите на ЕС
- Да гарантира правна сигурност, която да улесни инвестициите и иновациите в ИИ
- Да подобри управлението и ефективното прилагане на съществуващата правна уредба в областта на основните права и изискванията за безопасност, приложими към системите с ИИ
- Да улесни развитието на единен пазар за законни, безопасни и надеждни приложения с ИИ и да предотврати разпокъсаност на пазара

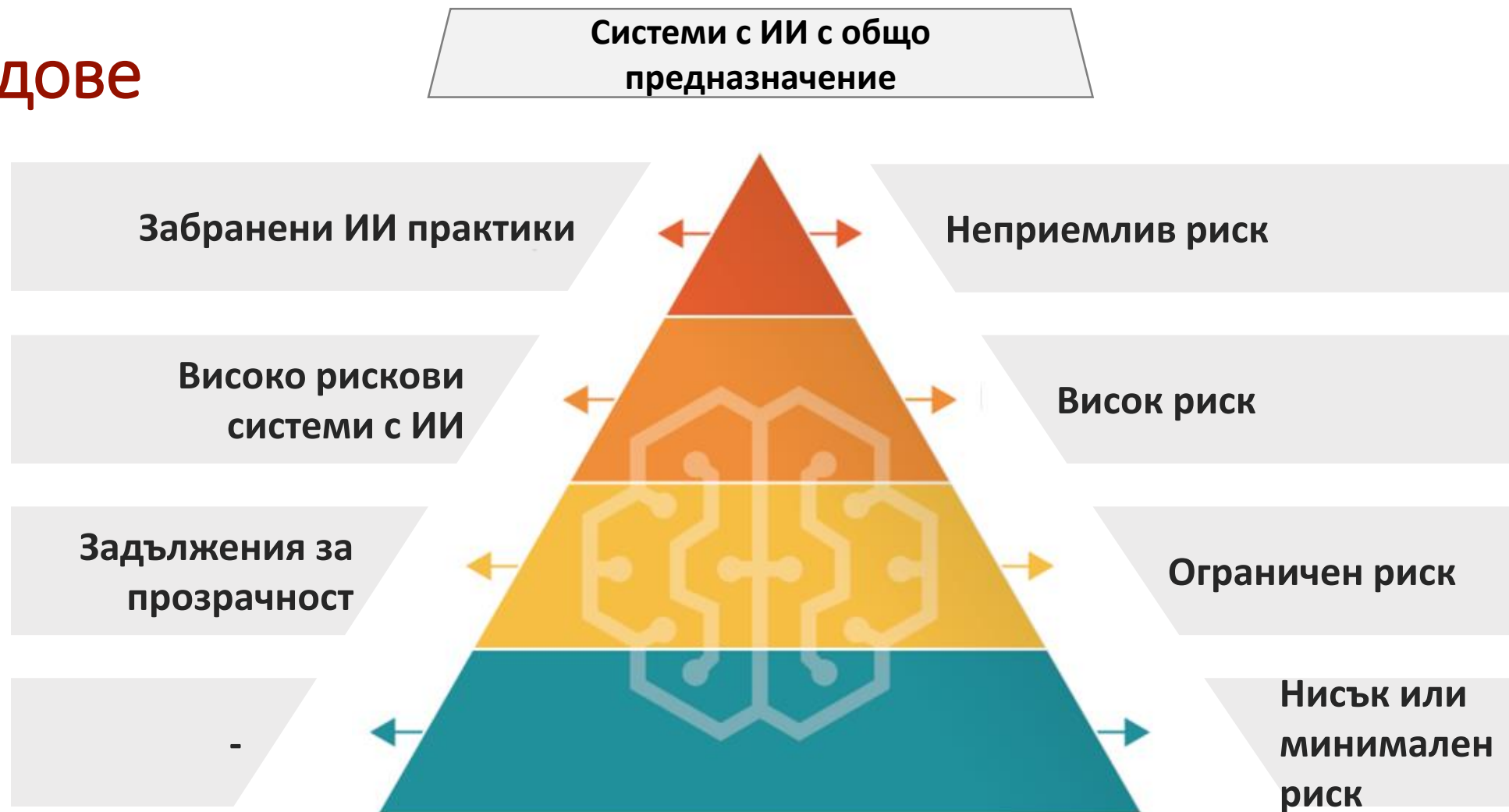
# Прилага се по отношение на...

- Доставчиците, които пускат на пазара или въвеждат в експлоатация системи с ИИ в ЕС, независимо дали тези доставчици са установени в ЕС или в трета държава
- Ползвателите на системи с ИИ, намиращи се на територията на ЕС
- Доставчиците и ползвателите на системи с ИИ, намиращи се в трета държава, когато резултатът, получен от системата, се използва в ЕС
- Вносителите и дистрибуторите на системи с ИИ
- Производители на продукти, които пускат на пазара или пускат в експлоатация система с ИИ, заедно с техния продукт и под собствено име или търговска марка
- Упълномощени представители на доставчици, които са установени на територията на ЕС

## **Изключения:**

- Продукти или системи, попадащи в обхвата на други актове
- Системи с ИИ, разработени или използвани изключително за военни или научни цели

# Видове



# Забранени практики (1)

- Система с ИИ, която си **служи със сублиминални техники отвъд рамките на съзнаването от човека**, с цел съществено да измени поведението на дадено лице по начин, който причинява или може да причини на същото или на друго лице физически или психологически вреди
- Система с ИИ, която **използва което и да е от уязвимите места на конкретна група лица**, дължащи се на тяхната възраст, увреждане, с цел съществено да измени поведението на дадено лице, принадлежащо към тази група, по начин, който причинява или може да причини на същото или на друго лице физически или психологически вреди



## Забранени практики (2)

- Система с ИИ за **оценка или класифициране на физически лица в течение на определен срок на база на тяхното социално поведение или известни или прогнозирани лични или личностни характеристики**, като социалният ранкинг води до едното или и двете от следните:
  - увреждащо или неблагоприятно третиране на определени физически лица или групи от такива лица в социални контексти, които нямат връзка с контекстите, в които първоначално са били генерирани или събрани данните;
  - увреждащо или неблагоприятно третиране на определени физически лица или групи от такива лица, което е неоправдано или непропорционално на тяхното социално поведение или на тежестта, която му се отдава.

## Забранени практики (3)

- Системи за **дистанционна биометрична идентификация в реално време в обществено достъпни места за целите на правоприлагането или от името на правоприлагащите органи**, освен ако и доколкото такова използване е строго необходимо за една от следните цели:
  - целево издирване на конкретни потенциални жертви на престъпления;
  - предотвратяване на конкретна и значителна заплаха за живота или физическата безопасност на физически лица или за терористично нападение;
  - локализиране или идентифициране на извършител или заподозрян в извършването на престъпление, за целите на провеждане на наказателно разследване, наказателно преследване или изпълнение на наказателна санкция за престъпления, наказуеми в съответната държава членка с лишаване от свобода или мярка за задържане с максимален срок не по-малко от 3 (или 5) години\*.

# Високо рисковни системи с ИИ

- Система с ИИ се счита за високорискова, когато са изпълнени следните две условия:
  - системата с ИИ е предназначена да се използва като защитен елемент от продукт или самата тя е продукт, попадащ в обхвата на законодателството на Съюза за хармонизация;
  - продуктът, чийто защитен елемент е системата с ИИ, или самата система с ИИ като продукт трябва да премине оценяване на съответствието от трета страна с оглед на пускането на пазара или въвеждането в експлоатация на посочения продукт в съответствие със законодателството на Съюза за хармонизация.
- ИЛИ
  - Системи с ИИ, посочени в **приложение III**, включително относно „правораздаване и демократични процеси“ - за системи с ИИ, предназначени за използване от съдебен орган и/или от тяхно име при тълкуване на фактите или правото и при прилагане на закона към конкретен набор от факти.

# Изисквания към високо рискови системи с ИИ

- Система за управление на риска
- Използвани данни и управление на данните
- Техническа документация
- Поддържане на регистри (записи за събития)
- Прозрачност и предоставяне на информация на ползвателите
- Човешки надзор
- Точност, надеждност и киберсигурност
- Мониторинг от доставчиците след пускането на пазара и план за мониторинг след пускането на пазара за високорискови системи с ИИ
- Докладване на сериозни инциденти и неизправности

# Задължения за прозрачност за някои системи с ИИ

- Системите с ИИ, предназначени да взаимодействат с физически лица, са проектирани и разработени по такъв начин, че физическите лица да бъдат информирани, че взаимодействат със система с ИИ, освен ако това не е очевидно от обстоятелствата и контекста на употреба.
- Ползвателите на система за разпознаване на емоции или на система за биометрична категоризация информират физическите лица, които са обект на нейното действие.
- Ползвателите на система с ИИ за генериране или обработка на изображения, аудио или видео съдържание, които до голяма степен наподобяват съществуващи лица, обекти, места или други субекти или събития и навярно някой би приел за автентични или достоверни („дълбинен фалшификат“), са задължени да разкрият, че съдържанието е било изкуствено създадено или обработено.
- Изключения

# Управление и прилагане

- **Европейски съвет по изкуствен интелект**
  - Състои се от по един представител от всяка държава членка
  - Европейският надзорен орган по защита на данните участва като наблюдател
- Всяка държава членка създава и/или определя поне един нотифициращ орган и поне един надзорен орган (**национални компетентни органи**) с организация на работа, която гарантира обективността и безпристрастността на техните дейности и задачи.

# Други нетехнически методи

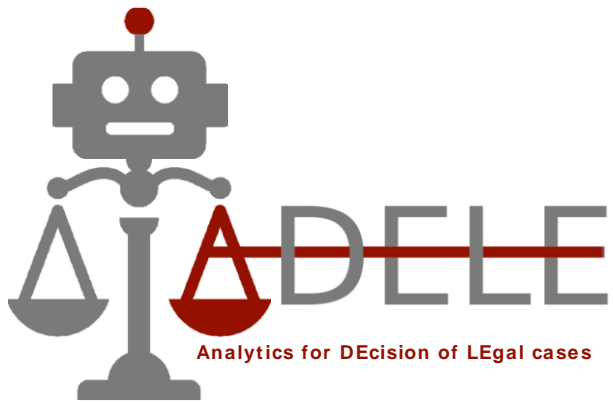
- Кодекси за поведение
- Стандартизация
- Сертифициране
- Отчетност чрез рамки за управление
- Образование и осведоменост за насърчаване на етичните нагласи
- Участие на заинтересованите страни и социален диалог
- Многообразие и екипи за приобщаващо проектиране
  - Значение на пол, културен произход, етническа принадлежност, религия, език
  - Когнитивно многообразие за по-добро вземане на решения
  - Мултидисциплинарност – етично, социално, икономическо въздействие на ИИ
  - Образование – ИИ отвъд инженерната дисциплина и програмирането

# Инициативи на Съвета на Европа

- European Ethical Charter on the use of AI in judicial systems (2018)
- Ad hoc Committee on Artificial Intelligence – CAHAI (2019-2021)
  - [Towards regulation of AI systems](#) ('12.2020)
  - [Possible elements of a legal framework on artificial intelligence, based on the Council of Europe's standards on human rights, democracy and the rule of law](#) ('12/2021)
- Committee on Artificial Intelligence – CAI (2022+)
  - **Convention on artificial intelligence, human rights, democracy and the rule of law (в процес на разработка)**
- CEPEJ-CYBERJUST Advisory Board on AI (2022+)







Александра Цветкова, Фондация „ЛИБРе“  
[alexandra.tsvetkova@libreresearchgroup.org](mailto:alexandra.tsvetkova@libreresearchgroup.org)



Този проект е финансиран от Програма „Правосъдие“  
(2014-2020) на Европейския съюз по Споразумение за  
отпускане на БФП № 101007420.