



Ресурсен център Европа



С подкрепата на
програма "Еразъм+"
на Европейския съюз



07 март 2021

ПОЛИТИКА НА ЕВРОПЕЙСКИЯ СЪЮЗ ЗА ИЗКУСТВЕНИЯ ИНТЕЛЕКТ

EUROPEAN UNION ARTIFICIAL INTELLIGENCE POLICY

д-р КРАСИМИРА ВАСИЛЕВА,
председател на NABEC

Abstract

The paper examines the issue of EU approach to the artificial intelligence. Given the major impact that AI can have on our society and the need to build trust, it is vital that European AI is grounded in our values and fundamental rights such as human dignity and privacy protection.

голяма ефективност на системите за производство чрез прогнозна поддръжка, повече сигурност за европейците.

Страхове и притеснения

Изкуственият интелект се възприема като основен фактор за икономическото развитие и възможност за решаване на много предизвикателства, поради което и инвестициите в тези технологии растат постоянно. Голяма част от интереса се фокусира върху неговия потенциал в подкрепа на икономическия растеж. По-малко внимание се обръща на начина, по който той може да засегне хората, техните основни права. Това надига недоволство сред известни кръгове на обществото, които виждат назряване на етични, правни и социални проблеми с редица въпроси, много от които морални.

За добро е когато машините извършват рутинна, опасна или непривлекателна дейност вместо хората, които биха могли да се концентрират върху по-интересни задачи или да се наслаждават на повече свободно време, но това може да се окаже и лошо. Анализатори прогнозираят, че изкуственият интелект може да предизвика [масова глобална безработица](#), тъй като редица професии ще станат автоматизирани и няма да изискват човешка намеса.

Вярно е, че математическите и изчислителни конструкции в основата на изкуствен интелект не са заложили за негативно въздействие върху [правата на човека](#), но тяхното приложение и въвеждане в процеси може да доведе до такова. Изначално не е присъщо изкуственият интелект да дискриминира по пол, религия или друг признак и няма предразсъдъци, но на база получената информация той може да придобие съществуващи вече в обществото ни такива. Например система, на която е зададено да намери най-подходящите кандидати за позицията „програмист“, може да оптимизира информацията за тази професия и да приоритизира кандидатите мъже, т.к. в обществото ни тази професия се счита за „мъжка“.

Решенията на изкуствения интелект могат да повлияят на защитата на основните права на гражданите, включително чувствителни данни, дискриминация. Например случая на системите за наблюдение и лицево разпознаване, които все повече се поставят на обществени места, тема предизвикала сериозен дебат. При тях [дискриминацията](#) е една от многото "неочаквани последици", които може да предизвика технологията.

Непредвидената пристрастност в алгоритмите, особено машинното обучение, заплашва да се превърне в расови, сексуални и други пристрастия независимо, че учени се стремят изкуствения интелект да бъде обективен, рационален мислител, безпристрастен към конкретна раса, пол или сексуалност.¹

Много учени се опасяват, че изкуственият интелект може да се използва за **масово наблюдение**. Властите на държави използват комбинацията от технологии за лицево разпознаване и изкуствен интелект, за да ограничат престъпността. В което има смисъл, но следва да се има предвид и **неприкосновеността на личния живот** на гражданите. Китай е развил т.н. "социална кредитна система", която следи поведението на гражданите. Резултатите от нея определят достъпа до всичко - от самолетни полети до определени услуги за онлайн запознанства. Тоби Уолш, професор по изкуствен интелект в Университета в Нов Южен Уелс, заяви, че наблюдението ползващо изкуствен интелект се класира като с високо плашещи последици.

Според мнението на учени, биоинженерството, комбинирано с развитието на изкуствения интелект, може да доведе до разделение на човечеството на малка група от "**суперхора**" и внушителна група от по-нисши, безполезни хомо сапиенци.

Когато става дума за изследвания в рамките на биологията, физиката, химията и компютърните науки, учените са свикнали да разполагат с пълната информация или поне достатъчна такава, с която да може да се повторят експериментите и да се сверят резултатите. На практика, това се случва рядко и **липсата на прозрачност** затруднява признанието на автентичността на откритията в научните среди.

Споделянето на алгоритмите също не е обичайна практика. Според доклада State of AI едва 15% от всички изследвания свързани с изкуствен интелект се споделя кода, с който са извършени. Основните и едни от най-добре познати инициативи за изкуствен интелект като OpenAI на Илон Мъск и DeepMind на Google държат в тайна софтуерните алгоритми. Така се публикуват хиляди научни трудове за изкуствен интелект, но без да има възможност за обективна проверка и потвърждение на резултатите.

¹ Например Тай, чатботът с изкуствен интелект на Microsoft, обучен да имитира потребителите, контактуващи с него онлайн, който за по-малко от ден бе вербуван в расизъм и език на омразата. Това принуди Microsoft да спре акаунта му, предизвиквайки сериозен дебат за възможността изкуственият интелект да бъде ръководен от предразсъдъци.

Развитието на изкуствения интелект често се свързва с иновативни военни машини и технологии. Някои се опасяват, че това може да завърши под формата на **война**. „Всякаква компютърна система, изкуствен интелект или не, която автоматично решава въпроси на живот и смърт – например дали да изстреля ракета – е действително плашеща идея“, казва Франк ван Хармелен, изследовател на изкуствения интелект към Vrije Universiteit Amsterdam. Има се предвид, че автоматизираните системи могат да направят грешка в преценката си, което да отприщи агресия между държавите.

Опасенията са, че роботи могат да идентифицират кого да убиват автоматично без намесата на човек в целия процес. Дронове убийци биха могли да бъдат програмирани със снимка и адрес, и след убиване на целта да се самовзривяват. Това са перфектни начини да убиеш някого, без да бъдеш разкрит. Много фирми и организации обещава да не работят по тази технология, но тя съществува.

Страхове и притеснения за опасност за човека се появяват и от факта, че изкуственият интелект може **да прогресира неконтролируемо** и какво ще се случи с това, кои данни може да ползва системата и как се събират и обработват? Дизайнерите може да знаят достатъчно добре как работи всеки отделен компонент на различни нива на сложните системи с изкуствен интелект, но дали винаги са наясно как биха функционирали всички компоненти заедно. Това означава, че системи няма да могат да бъдат разбрани напълно и на теория могат да функционират по непредсказуеми начини.

Съществува страх от **злоупотреба** дори и с най-съвършените машини, тъй като се разработват системи, които достигнат нивото на човешкия интелект. С всяко следващо поколение те успяват да се справят с грешките на предишното все по-добре. И това увеличава вероятността да се натъкнем на нежелани събития в симулациите, които да не могат да бъдат елиминирани, преди да влязат в реалния свят.

А до къде ще се стигне със създаването на „живи роботи“, изкуствени неживи организми. Вече са създадени „живи роботи“ от клетки на жаби. Как биха се развили в нашата среда и как биха взаимодействали със съществуващите организми?

С все по-широкото въвеждане на изкуствения интелект в различни процеси на вземане на решения, които директно засягат хората, възникват все повече

проблеми, свързани с опазването на правата на човека като: Кой носи отговорност, когато права на човека са нарушени в следствие решение, взето на база изкуствения интелект?; Възможно ли е системи, основани на изкуствения интелект, да са безпристрастни и недискриминационни?

Европейски подход към изкуствения интелект



В търсенето на баланс между свободата на иновациите, рисковете от тях и защитата на хората Европейския съюз разработи нови механизми за защита правата на човека и законова рамка, които да гарантират зачитането на съществуващите стандарти за спазване на правата на човека в контекста на новите технологии. Например [Общият регламент за защита на данните \(GDPR\)](#) от 2018 г. вече даде възможност за по-прозрачно използване на нашите данни от автоматизирани системи и техните разработчици. В България бе прието решението на Комисията за защита на личните данни, според което въвеждане на видеонаблюдение с възможност за лицево разпознаване, базирано на изкуствен интелект, в училище противоречи на правото на защита личните данни.

За да се изградят стабилни модели в основата на системи базирани на изкуствен интелект, висококачествените данни са ключов фактор за подобряване на производителността. [Регламентът относно управлението на данните](#), приет от Европейската комисия на 25 ноември 2020 г., насърчава обмена на данни между секторите и държавите членки.

Тези документи са важна стъпка в стремежа на Комисията за изграждане на доверие и са крачка напред за осигуряване на правна яснота в приложенията, основани на изкуствен интелект.

[Европейската стратегия за изкуствен интелект](#)² и [Европейският координиран план за изкуствения интелект](#)³ поставят доверието като предпоставка за осигуряване на ориентиран към човека подход към изкуствения интелект.

През април 2019 г. Комисията публикува Съобщение на тема "[Изграждане на доверие в ориентирания към човека изкуствен интелект](#)", в което се

² 2018, <https://eur-lex.europa.eu/legal-content/BG/TXT/HTML/?uri=CELEX:52018DC0795&from=EN>

³ 2019, <https://data.consilium.europa.eu/doc/document/ST-6177-2019-INIT/bg/pdf>

подчертават основните изисквания и концепцията за надежден изкуствен интелект, въведена от Експертната група на високо равнище по изкуствен интелект (AI HLEG) в техните [Насоки относно етичните аспекти за надежден изкуствен интелект](#).⁴ Тези две концепции определят визията на Комисията за изкуствен интелект – *„Доверието в ориентиран към човека ИИ ще се изгради в резултат на утвърждаването на ценностите и принципите, добре установена нормативна уредба и етични насоки, които съдържат основни изисквания“*.

„Неприкосновеността на личния живот и защитата на данните ще определят равнището на доверие на гражданите и потребителите в ИИ. Собствеността върху данните, контролът и използването им от страна на предприятията и организациите са въпроси, които до голяма степен предстои да бъдат регулирани“.

На 9 февруари 2020 г. Европейската комисия оповести своята стратегия за бъдещата законодателна рамка, която да осигури и ускори развитието на високите технологии в рамките на Европейския съюз – [„Бяла книга за изкуствения интелект: Европа в търсене на високи постижения и атмосфера на доверие“](#)⁵. С нея Комисията поставя началото на широк европейски дебат по въпроса и се фокусира върху рисковите и „високорискови“ случаи при приложение на технологията. Разработките, които попадат в тази категория, ще са обект на по-строги правила и регулации. Това е първата стъпка от един дълъг процес на създаване на правила, които да защитят гражданите срещу злоупотреби с изкуствен интелект.

В тези документи ЕС залага [единен европейски подход](#) към развитието на технологиите с изкуствен интелект, изразяващ се в:

- [подкрепа за иновациите](#) и насърчаване на въвеждането на новите технологии и внедряването на изкуствен интелект;
- [подготовка](#) за произтичащите социално-икономически промени;
- осигуряване на [етическа и правна рамка](#), с бъдеща регулаторна рамка на ЕС, която да определя видовете правни изисквания, с особен акцент върху приложенията с висок риск.

⁴ <https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top>

⁵ https://ec.europa.eu/info/files/white-paper-artificial-intelligence-european-approach-excellence-and-trust_bg/

Според Насоките, **надеждният изкуствен интелект** трябва да има следните характеристики:



- да е **законосъобразен**, като се придържа към всички приложими закони и подзаконови актове;
- да е **етичен**, като гарантира зачитането на етичните принципи и ценности;
- да е **стабилен** от техническа и от социална гледна точка, тъй като и при най-добри намерения е възможно системи с изкуствен интелект да причинят неумишлено вреди.

Посочени са и 4-те **етични принципи**, произтичащи от основните права, които следва професионалистите, работещи в областта на изкуствения интелект, да спазват и за да се гарантира, че системите с изкуствен интелект са разработени, внедрени и използвани по надежден начин:

- **Зачитане на свободата и автономността на хората**

Основните права, на които се основава ЕС, са насочени към гарантиране на зачитането на свободата и автономността на хората. Хората, които си взаимодействат със системите с изкуствен интелект, трябва да са в състояние да продължат цялостно и ефективно да се самоопределят и да могат да участват в демократичния процес в нашите общества. Системите с изкуствен интелект не следва необосновано да подчиняват, принуждават, заблуждават, манипулират, уговарят или подбират хората.

- **Предотвратяване на вреди**

Системите с изкуствен интелект и средите, в които те функционират, трябва да бъдат безопасни и сигурни. Това включва защитата на човешкото достойнство, както и психическата и физическата неприкосновеност. Те трябва да бъдат технически стабилни и да гарантират, че не са отворени за злонамерена употреба.

- **Справедливост**

Системите с изкуствен интелект трябва да гарантират равномерното и справедливо разпределение както на ползите, така и на разходите, а също и липса на несправедливи предубеждения, дискриминация и заклеймяване по отношение на лицата или групи лица. За да е възможно това, необходимо е да бъде установен

субектът, отговорен за вземането на решението, а процесите на вземане на решения да бъдат обясними.

- **Обяснимост**

Процесите на изграждане на системи, базирани на изкуствения интелект трябва да бъдат прозрачни, възможностите и целта ѝм да бъдат открито съобщавани, а решенията да бъдат разяснявани до степента, в която е възможно на пряко или косвено засегнатите от тях.

Европейската комисия определи и **седем основни изисквания за постигане на надежден изкуствен интелект** /фигура 1./. Всичките те са еднакво значими, подкрепят се взаимно и следва да се прилагат и оценяват от различните участници в през целия жизнен цикъл на система с изкуствен интелект. Комисията призовава за въвеждането на **европейски стандарти за безопасност** и за развиването на стабилна **европейска система за сертифициране**, основана на процедури за изпитване, които биха позволили на предприятията да потвърдят надеждността и безопасността на своите системи с изкуствен интелект.

Фигура 1: Взаимодействие на 7-те изисквания за създаване на надежден изкуствен интелект



Комисията призовава достъпа до данни „да бъде управляван и контролиран по подходящ начин“, включително зачитане на правото на неприкосновеност на личния живот. Тя насочва своите усилия към създаване на регулации по въпросите за собствеността върху данните, контрола над тях и използването им от страна на предприятията и организациите.

Особен акцент се поставя върху прилагането на изкуствен интелект в системи, който ще спомагат за намаляване на емисиите на парникови газове, по-специално в промишления, транспортния, енергийния, строителния и селскостопанския сектор.

За да може промените, които изкуственият интелект налага в нашия живот и на които сме свидетели, да се контролират, Комисията работи за изграждане на стратегии за тяхното прогнозиране, преди да се извършат евентуални реструктурирания и да се предлага продължаващо обучение на всички работници.

Европейската комисия препоръча на държавите членки да засилят разпоредбите на законодателството си, за да защитят основните права спрямо изкуствения интелект, който може да бъде източник на грешки и дискриминация, които в момента трудно могат да се оспорят.

Европейската комисия обяви плановете си да сформира група с основна област изкуственият интелект, чиято задача ще бъде да се анализират етичните проблеми, свързани с тази технология. До края на 2021 год. групата трябва да създаде [проектно ръководство за етична разработка и употреба на изкуствен интелект](#) на базата на фундаменталните права на европейските граждани.

Комисията желае да пренесе своя [подход към изкуствения интелект на международната сцена](#) и ще продължи да играе активна роля, включително в рамките на Г-7⁶ и Г-20.⁷

⁶ Г-7 - Групата на седемте - е група от 7-те икономически най-развити държави – Великобритания, Германия (Западна Германия до 1991 г.), Италия, Канада, САЩ, Франция, Япония, които заедно представляват около 65% от БВП в света. От 1997 до 2014 г. форумът е наричан Г8, Г7 плюс Русия.

⁷ Г-20 - Групата на двайсетте - е група от 19-те страни с най-големи икономики в света плюс Европейския съюз, представян от държавата председател на Съвета на ЕС (освен в случаите, когато страната председател е член на голямата осморка и по такъв начин вече е представена в G20). Това са Австралия, Аржентина, Бразилия, Великобритания, Германия, Индия, Индонезия, Италия, Канада, Китай, Мексико, Русия, Саудитска Арабия, САЩ, Турция, Франция, Южна Корея, ЮАР, Япония и Европейският съюз.