



## ИЗВЕСТИЯ

списание на Икономически университет – Варна

<http://journal.ue-varna.bg>

## ЕТИЧНИТЕ ПРЕДИЗВИКАТЕЛСТВА ПРЕД ЕВРОПЕЙСКИЯ СЪЮЗ ПРИ ИЗПОЛЗВАНЕТО НА ИЗКУСТВЕНИЯ ИНТЕЛЕКТ

Емил РАДЕВ<sup>1</sup><sup>1</sup> Европейски парламент, Брюксел, Белгия. E-mail: [emil.radev@europarl.europa.eu](mailto:emil.radev@europarl.europa.eu)

JEL: K2

Резюме

**Ключови думи:**

изкуствен интелект,  
етични предизвикател-  
ства, етични принципи,  
дефиниция за ИИ

В настоящата статия се изследват някои аспекти при приложението на изкуствения интелект и свързаните с него етични предизвикателства пред Европейския съюз. Изследва се същността на изкуствения интелект – дефиниция и концепции за приложението му, както и европейската правна рамка в дефинирането на понятието. Във фокуса на разработката са поставени етичните предизвикателства в правната регулация на изкуствения интелект. В заключение и на база на извършения анализ са формулирани обобщения и са направени предложения.

© 2021 Икономически университет – Варна

**Цитиране:** РАДЕВ, Е. (2021) Етичните предизвикателства пред Европейския съюз при използването на изкуствения интелект. *Известия. Списание на Икономически университет – Варна*, 65 (3), с. 310 – 331.

### Въведение

През последните години и особено през 2020 г., която премина под знака на пандемията от COVID-19, прогресиращата цифрова трансформация се характеризира с все по-широко използване на технологиите с изкуствен интелект (ИИ). Автоматизацията на знанията и рутинните задачи, както и активното внедряване на системи за вземане на решения, имат потенциала да променят коренно живота на много хора и обществото като цяло.

От автоматизирани асистенти до автономни автомобили ИИ се развива бързо и носи редица ползи – стимулира иновациите, повишава икономическата производителност и дори жизнения стандарт. Опасенията относно рисковете и етичните проблеми, свързани с тези технологии, обаче са съвсем реални. В интервю

пред ВВС Стивън Хокинг заявява: „Развитието на ИИ може да означава края на човешката раса“.<sup>1</sup> А според Бил Гейтс ИИ е като ядрената енергия – „както обещаваща, така и опасна“.<sup>2</sup> Сред широката общественост тревогите за потенциала на ИИ да измества човека в някои видове дейности са очевидни.

Четвъртата индустриална революция (Индустрия 4.0), както е обяснено от ОИСР през 2017 г., се отнася до „използването в индустриалното производство на скорошни и често взаимосвързани цифрови технологии, които дават възможност за нови и по-ефективни процеси и които в някои случаи дават нови стоки и услуги. Свързаните технологии са много – от разработките в машинното обучение и науката за данни, които позволяват все по-автономни и интелигентни системи, до евтини сензори, които са в основата на интернет на нещата, до нови устройства за управление, които правят възможно второ поколение индустриална роботика“<sup>3</sup> (Andreeva, A., Yolova, G., p. 2019a) (Andreeva, A., Yolova, G., p. 2019b).

Обработвайки големи масиви от данни и свързвайки технологии, изкуственият интелект доведе до трансформации, които изискват да действаме значително по-адаптивно. ИИ повдига голям кръг от етични, правни и социални въпроси. (Andreeva, A., Yolova, G., Dimitrova, D., 2019) (Андреева, А., Йолова, Г., 2018) (Андреева, А. Йолова, Г., 2020) Ето защо се налага бързо да запълним белите петна в съществуващата регулаторна рамка. В центъра на правните норми, които е необходимо да изготвим, разбира се, трябва да са нуждите на хората. Голямото предизвикателство е как да осигурим лостове за контрол, без да спираме растежа, как да защитим европейските граждани, без да ограничаваме прекомерно бизнеса.

Понятието ИИ се употребява в различен смисъл. По всичко личи, че в представите си имаме различни разбираня за ИИ – от хуманоидни роботи като например в „Аз, роботът“ (Asimov, 1950), които могат да надминат човешкия интелект, до най-обикновени работи косачки. За да можем да регулираме тази материя, всъщност е необходимо да знаем какво точно регулираме. Това е още по-важно на фона на прогнозите, че в бъдеще животът на хората и машините ще бъде все по-тясно свързан. Ускорявайки трансформацията, ИИ осезателно се утвърждава като фактор за икономически растеж. Изкуственият интелект има преобразуващ ефект върху пазарите по света<sup>4</sup> и същевременно сам определя един бързо растящ пазар, за който се очакваше приходите да достигнат почти 50 млрд. долара през 2020 г.

В контекста на гореизложеното настоящата статия **има за цел** се представят някои аспекти при приложението на ИИ и свързаните с него етични предизвикателства пред Европейския съюз.

За реализиране на поставената цел в статията са заложени следните изследователски задачи:

- да се изследва същността на ИИ – дефиниция и концепции за приложение

нието му;

- да се анализира европейската правна рамка в дефинирането на понятието;
- да се разгледат етичните предизвикателства в правната регулация на изкуствения интелект;
- на база на извършения анализ да се формулират изводи, обобщения и предложения с практическа насоченост.

**Методологията** на изследването включва комплексното прилагане на традиционни методи за правните изследвания: нормативен, сравнителноправен, индукция, дедукция и синтез и др.

## **1. Същност на изкуствения интелект – дефиниция и концепции за приложението му**

Въпреки интереса към ИИ от страна на научните среди, индустрията и публичните институции, все още няма стандартна дефиниция за това какво всъщност включва той. ИИ често се описва чрез подходи, основани на човешката интелигентност или интелигентността като цяло. Много определения се отнасят до машини, които се държат като хора или са способни на действия, изискващи интелигентност<sup>5</sup> (Russel, S., Norvig, P., 1955) (McCarthy, J., 2007) (Nilsson, N., 1998) (Fogel D.B., Fogel L.J., 1996) (Albus, J. S., 1991) (Luger, G.F., Stubblefield, W.A., 1993) (Winston, P. H., 1992) (Gardner, H., 1987) (Newell, A., Simon, H.A., 1976) (Bellman, R., 1978) (Minsky, M., Papert, S., 1969) (McCarthy, J., Minsky, M.L., Rochester, N., Claude, E., 1955). Според някои ИИ е област от компютърните науки, фокусирана върху създаването на интелигентни машини, които функционират и реагират като хора. Други термини, използвани за ИИ, могат да включват когнитивни изчисления или фрази, започващи с думи като „умни“, „интелигентни“, „предсказуеми“, „когнитивни“. ИИ понякога се използва в комбинация с големи данни и анализи в реално време за вземане на интелигентни решения относно поведението на потребителя или машината, но може да работи, като изучава как мисли човешкият мозък и как хората се учат, анализират и преодоляват проблеми.

Очевидно понятието „изкуствен интелект“ е свързано с компютърните науки, биологията, психологията, лингвистиката, математиката и инженерството. То търпи постоянно предефиниране с изключването на някои теми и с включването на нови, което донякъде е нормално предвид бързото развитие на технологиите, използващи ИИ.

Сериозните изследвания в тази област започват непосредствено след края на Втората световна война. През 1950 г. Алън Тюринг, известен с разбиването на нацисткия код ENIGMA, публикува „Изчислителни машини и разузнаване“. В ста-

тията си той предлага да отговори на въпроса „Могат ли машините да мислят?“ и въвежда Теста на Тюринг, за да определи дали компютърът може да демонстрира същата интелигентност (или резултатите от същата интелигентност) като човек.

За възникването на ИИ като нова самостоятелна област се приема лятото на 1956 г., когато Джон Маккарти и още няколко учени (Minsky, Shanon и Rochester) успяват да съберат на работна среща изследователи от САЩ, занимаващи се с теория на автоматите, невронните мрежи и интелекта. В обръщението си към участниците Маккарти за първи път официално употребява понятието „изкуствен интелект“. Той го определя като „науката и инженерството за създаване на интелигентни машини, особено интелигентни компютърни програми“. По думите му „това е свързано със сходната задача да се използват компютри за разбиране на човешкия интелект, но ИИ не трябва да се ограничава до методи, които са биологично наблюдаеми“ (McCarthy, 2007).

Човешката интелигентност също е трудна за дефиниране и измерване. Въпреки че има различни опити за количествено определяне (Gardner, 1983) (Neisser, U., Boodoo, G., 1996), обективната дефиниция на нещо толкова субективно и абстрактно като интелигентността (Kaplan, 2016) фалшиво създава впечатление за прецизност, която в действителност не може да бъде получена. Мнозина смятат, че ИИ е термин, отнасящ се до компютърни системи, които могат да усетят средата си, да мислят, евентуално да учат и да приемат действие в отговор на това, което усещат, или на своите цели.<sup>6</sup> Според С. Самоили (Samoili, 2020) повечето дефиниции, открити в научните изследвания, политическите или пазарните доклади са неясни и предлагат идеална цел, а не измерима изследователска концепция.

Прекаленото опростяване на концепцията за интелигентност, която е необходима, за да се дефинира или дори да се развие ИИ, е илюстрирано от Russel, Norvig (2010) и подчертано от Експертната група на високо ниво по изкуствен интелект, създадена от Европейската комисия (HLEG, 2019). Поставя се фокус върху рационален ИИ и, следователно, обмисляне на критерий спрямо идеалното представяне: „Системата е рационална, ако прави „правилното нещо“, като се има предвид това, което знае“ (Russel, Norvig, 2010).

Оформилото се през 1956 г. понятие за ИИ например понастоящем е официално дефинирано във Франция като „теоретичната и практическа междудисциплинарна област, която има за предмет разбирането на механизмите на познавателната и разсъдъчната функция, както и тяхната имитация посредством хардуер и софтуер, за целите на подпомагането или заместването на човешките дейности“ (доклад на Министерството на въоръжените сили от 2019 г., „ИИ в услуга на отбраната“, с. 3). Същата идея за „симулиране на човешките познавателни функ-

ции“ намираме и в руското определение (член 5 от президентския указ от 10 октомври 2019 г.). Тези две определения обаче засягат основно използването на ИИ във военната област.

В Испанската стратегия за изкуствения интелект от 2019 г. „ИИ може да се определи като наука и инженерство, което позволява проектирането и програмирането на машини, способни да изпълняват задачи, изискващи интелигентност. Вместо да постига обща интелигентност, настоящият ИИ се фокусира върху това, което е известно като специфичен ИИ, което води до много важни резултати в много области на приложение като обработка на естествен език или изкуствено зрение; обаче от научна и основна и приложна изследователска гледна точка общият ИИ остава основната цел, която трябва да бъде постигната, т.е. създаването на екосистема с интелигентни многозадачни системи.”

В Националната стратегия за ИИ на Германия от 2018 г. се посочва, че не съществува общовалидна или последователно използвана от всички заинтересовани страни дефиниция. Дефиницията за ИИ, използвана за стратегията за ИИ на федералното правителство, се основава на следното разбиране: „В много абстрактни термини изследователите на ИИ могат да бъдат разпределени в две групи: „силен (широк)“ и „слаб (тесен)“ ИИ. „Силен“ ИИ означава, че системите с ИИ имат същите интелектуални способности като хората или дори ги надвишават. „Слабият“ ИИ е фокусиран върху решаването на специфични проблеми с помощта на методи от математиката и компютърните науки, при което разработените системи могат да се самооптимизират. За тази цел аспектите на човешкия интелект са картографирани и официално описани, а системите са проектирани да симулират и подпомагат човешкото мислене”.<sup>7</sup>

В масовата употреба изкуственият интелект се отнася до способността на компютър или машина да имитира възможностите на човешкия ум – учене от примери и опит, разпознаване на обекти, разбиране и реагиране на езика, вземане на решения, решаване на проблеми и комбиниране на тези и други способности за изпълнение на функции, които човекът може да извършва, като поздрав на гости на хотела или шофиране на кола.

## **2. Европейската правна рамка в дефинирането на изкуствения интелект**

В своя резолюция от 16 февруари 2017 г., съдържаща препоръки към Комисията относно гражданскоправни норми за роботиката, Европейският парламент ясно заявява, че следва да бъде установено общоевропейско определение за „интелигентни“ автономни роботи, включително по целесъобразност определения

на неговите подкатегории, като се вземат предвид конкретни характеристики: способност да придобиват автономност посредством датчици и/или чрез обмен на данни с околната среда (взаимосвързаност) и анализ на тези данни.<sup>8</sup>

Според дефиницията, използвана в съобщението на Европейската комисия „Изкуствен интелект за Европа“,<sup>9</sup> понятието „изкуствен интелект“ включва системи, които показват интелигентно поведение чрез анализ на средата и предприемане на действия – с известна степен на самостоятелност – за постигане на конкретни цели. Системите с ИИ могат да бъдат изцяло основани на софтуер, действащи във виртуалния свят (гласови асистенти, анализ на изображения, софтуер, търсачки, системи за разпознаване на реч и лице и др.), или да бъдат вградени в хардуерни устройства (модерни роботи, автономни коли, дронове, интернет на нещата и др.).

Експертната група на високо равнище на ЕС по въпросите на ИИ дава в своя документ „Определение за ИИ: основни възможности и дисциплини“ следното определение: „Системите с изкуствен интелект (ИИ) са софтуерни (а вероятно и хардуерни) системи, създадени от хора,<sup>10</sup> които с оглед на дадена сложна цел действат в рамките на физическото или цифровото измерение, като възприемат заобикалящата ги среда чрез събиране на данни, тълкуват събраните структурирани или неструктурирани данни, разсъждават въз основа на познанието или обработват информацията, получена от тези данни, и вземат решение за предприемане на най-доброто (добрите) действие (действия) за постигане на дадената цел. Системите с ИИ могат или да използват символно представени правила, или да усвояват цифров модел и могат да адаптират поведението си, като анализират начина, по който средата е засегната от предишни техни действия.

Като научна дисциплина ИИ включва няколко подхода и техники като например машинно самообучение (специфични примери, за което са дълбокото самообучение и обучението с утвърждение), машинно разсъждение (което включва планиране, изготвяне на график, моделиране на познанието и разсъждение, търсене и оптимизация) и роботика (която включва контрол, възприятие, сензори и задвижващи механизми, както и интегрирането на всички други техники в киберфизически системи)“.

Същото определение е посочено и в „Насоки относно етичните аспекти за надежден ИИ“ от 8 април 2019 г., пар. 143: „Системите с изкуствен интелект (ИИ) са софтуерни (а вероятно и хардуерни) системи, създадени от хора, които с оглед на дадена сложна цел действат в рамките на физическото или цифровото измерение, като възприемат заобикалящата ги среда чрез събиране на данни, тълкуване на събраните структурирани или неструктурирани данни, обосноваване въз основа на знание или обработване на информацията, получена от тези данни, и взе-

мане на решение за предприемане на най-доброто (добрите) действие (действия) за постигане на дадената цел. Системите с ИИ могат или да използват символно представени правила, или да усвояват цифров модел и могат да адаптират поведението си, като анализират начина, по който средата е засегната от предишни техни действия“.<sup>11</sup>

Въпреки множеството аспекти на ИИ и съответно липсата на обща дефиниция, съществуват редица общи черти, които наблюдаваме в анализираните дефиниции. Този израз на общи аспекти предполага те да се разглеждат като основни характеристики на ИИ: възприемане на околната среда, включително разглеждане на сложността в реалния свят; обработка на информация, в това число събиране и интерпретиране на входни данни (под формата на данни); вземане на решения (включително разсъждения и учене) – предприемане на действия, изпълнение на задачи (адаптация, реакция на промени в околната среда и др.) с определено ниво на автономност; постигане на конкретни цели, което се счита за основната причина за системите за изкуствен интелект.<sup>12</sup>

Дефиницията, предложена от експертната група, може да се счита за изключително техническа за различни аудитории и цели, но това е много изчерпателно определение. То включва аспектите на възприятие, разбиране, интерпретация, взаимодействие, вземане на решения, адаптиране към поведението и постигане на целите, докато други определения не ги разглеждат в толкова детайли.

В свои доклади от 20 октомври 2020 г.<sup>13</sup> Европейският парламент също се опитва да даде единна дефиниция на това какво е система с ИИ: „Изкуствен интелект“ означава система, която е или базирана на софтуер, или вградена в хардуерни устройства, и проявява интелигентно поведение, като наред с другото събира и обработва данни, анализира и тълкува заобикалящата я среда и като предприема действия с известна степен на автономност за постигането на конкретни цели“.<sup>14</sup>

С други думи, има различни аспекти, които трябва да се разглеждат в различните дефиниции на ИИ като предизвикателство към регулирането.

По време на пандемията ползите от технологиите с изкуствен интелект излязоха на преден план. COVID кризата на практика ускори значително процеса на дигитализация в световен мащаб. С напредъка на технологиите обаче изникват и много въпроси за социално-икономическия ефект, въпроси от правно и етично естество. Също както индустриалната, така и дигиталната революция изисква обществото и правото да се адаптират към промените. Налага се бързо да запълним празнотите в настоящото законодателство. В центъра на правната рамка, която е необходимо да изготвим, разбира се, трябва да са нуждите на хората.

### **3. Етичните предизвикателства в правната регулация на изкуствения интелект**

Използването на ИИ за наблюдение или дори прогнозирането на човешкото поведение рискува заклеяване, засилване на съществуващите стереотипи, социална и културна сегрегация и изключване, подкопавайки индивидуалния избор и равните възможности. Потенциалът на ИИ за овластяване, рисковете в резултат на социална оценка, базирана на използването на ИИ, както и достъпността на услугите на ИИ дават основания за загриженост относно човешката безопасност, здраве и сигурност, свобода, неприкосновеност на личния живот, достойнство, самоопределение и недискриминация.

През последните години се провеждат все повече дискусии относно изкуствения интелект и етиката. Изследване на етичните аспекти на ИИ, публикувано през 2019 г., например идентифицира 84 документа, съдържащи етични принципи или насоки за ИИ (Jobin, A., Lenca, M., Vayena, E., 2019). Проучването стига до заключението, че има относително единомислие в световен мащаб за поне пет подхода с етичен характер, основани на принципи, и това са: 1) прозрачност; 2) справедливост и „честност“; 3) използване, което да не вреди; 4) отговорност; 5) цялост/защита на данните.

В същото време проучването установява, че има значителни разлики в това как тези принципи се тълкуват, защо се считат за важни, с какъв проблем, област или участници са свързани и как трябва да бъдат приложени.

Основа за успеха и приемането на новите технологии в обществото ни е безспорно доверието. Без доверие в сигурността, полезността и прозрачността на приложенията ИИ няма да може да се използва в ЕС в пълния си потенциал. Следователно генерирането на доверие чрез овластяване на гражданите трябва да бъде централна цел на политиката за следващото десетилетие на дигиталната трансформация. За да се доверят на технологичните решения, хората трябва да могат да ги разбират, контролират и да виждат ползите от тях. Европейските граждани се нуждаят от подходящ, ефективен, прозрачен и съгласуван регулаторен подход на общоевропейско ниво. Това е пътят за постигане на достатъчно ясни условия за дружествата, така че те да могат да разработват приложения и да планират своите бизнес модели, а същевременно Съюзът и държавите членки да упражняват ефективен контрол чрез установяването на регулаторни норми – без излишни рестрикции, но в защита на интересите на гражданите и предприятията.

Увеличеното използване на алгоритмично вземане на решения въз основа на ИИ в областите на финансовите услуги, банковото и наказателното правосъдие, без участието на човешка преценка или надлежен процес, може да утвърди



вредни социални стереотипи срещу определени малцинствени групи и да засили пристрастията, основани на полов и расов признак. Тази практика вече беше критикувана от няколко институционални участника в целия Европейски съюз като Съвета на ЕС, Европейския надзорен орган по защита на данните и Агенцията на ЕС за основните права.

Така нареченият „силен” ИИ прониква в правните категории и предизвиква преразглеждане на традиционни правни термини като „автономия“ и „неприкосновеност на личния живот“. Това от своя страна повдига въпроси относно способността на правото на ЕС да постигне правилния баланс между технологията като регулаторен обект или категория и технологията като регулаторна програма за определяне на дневен ред. Силното развитие на изкуствения интелект освен това поражда необходимостта да се признае, че кодовете кодифицират стойности и не могат да бъдат третираны просто като въпрос на инженерството. Етичните проблеми, свързани със социалната сила на алгоритмите, придружават въпросите за дигиталното разделение между поколенията, което може да повлияе на упражняването на правото на живот, правото на справедлив процес, правото на личен живот, свободата на изразяване, а също и на правата на работниците. Съществуват сериозни етични, психологически и правни опасения относно автономията на интелигентните роботи и тяхното въздействие върху взаимоотношенията лекар – пациент в здравеопазването (Андреева, А., Йолова, Г., 2018) или при евентуално използване на ИИ в правораздаването. Те все още не са разгледани по подходящ начин на равнище ЕС, особено по отношение на защитата на личните данни на пациентите, отговорността при решаване на съдебни спорове и произтичащите от това нови икономически и правни взаимоотношения.<sup>15</sup> Следователно е необходима строга и ефективна ръководна етична рамка за разработването, проектирането, производството и използването на ИИ. Тази рамка трябва да запазва човешкия надзор над автоматизираните и алгоритмичните вземания на решения, да се основава на принципите и ценностите, залегнали в Хартата на основните права на ЕС, като човешко достойнство, равенство, справедливост, недискриминация, информирано съгласие, личен и семеен живот и защита на данните, както и други основни принципи и ценности.

Остават открити въпросите каква трябва да бъде кодификацията на етиката на ИИ в ЕС и кои части от нея биха били най-подходящи за законодателството. Иначе казано, за да намерим ключа за създаването на надежден ИИ, който да отговаря на основни етични стандарти, трябва да сме наясно какви стандарти да бъдат включени в бъдещото законодателство.

Важно е да се разбере, че има два вида етични проблеми, които изкуственият интелект повдига. Можем да ги открием въз основа на разликата между установе-

ни права и възникващи права (Evas, 2020).

Вече установените етични принципи са заложили в настоящите европейски правни и политически рамки и обхващат много от идеите, развити в Хартата на основните права на ЕС.<sup>16</sup> Ще открие само няколко основни и добре познати примера като свободата на мисълта, съвестта и религията (член 10), свободата на събранията и сдружаванията (член 12), равенството пред закона (член 20) и недискриминацията (член 21). Например използването на софтуер за лицево разпознаване по време на митинги или протести не представлява нова заплаха за свободата на събиране, а по-скоро засилва вече съществуващ инструмент за проследяване на лица. Това обаче може да представлява проблем за някого, който иска да остане анонимен.

Във фокуса на възникващите етични проблеми е поверителността на данните, която се е превърнала в право само по себе си, а не в компонент на някакво друго право. В член 8 от Хартата на ЕС защитата на личните данни се разглежда като независимо право.

Много хора разбират идеята за недискриминацията, но идеята за поверителността на данните като право е по-малко приета от гледна точка на етичните съображения. В проучване на Евробарометър, проведено през 2019 г.<sup>17</sup> относно дискриминацията в ЕС, дори не се появява въпросът дали недискриминацията е нещо, което заслужава защита. И все пак през същата година Cisco прави глобално проучване на нагласите към поверителността на данните.<sup>18</sup> В него повечето въпроси са съсредоточени върху разбирането на хората, че поверителността на данните е важна и за нея трябва да се носи ясна отговорност – дали от правителствата или от индустрията. Въпреки че поверителността и защитата на данните имат статут, равен с правото на недискриминация съгласно Хартата, те очевидно нямат същия статут в съзнанието на обикновените граждани и нормите все още се разработват с политиките. Това разграничение е важно както за разбирането на начина, по който етичните стандарти се превръщат в законодателство, така и за това как експертите разбират ефектите от етичната рамка върху икономическото и социалното благосъстояние на европейските граждани.

От гледна точка на разработчика достъпът до по-малко данни влияе върху скоростта на развитие и точността на резултатите от приложенията на изкуствения интелект. Проучванията за въздействието на Общия регламент за защита на данните (GDPR) върху компаниите отчитат, че организациите отделят ресурс за изпълнение на задълженията си за поверителност на данните. Изследване от април 2020 г., анализиращо стартиращи компании с изкуствен интелект и връзката им с данните и GDPR, сочи, че европейските компании са по-склонни да развиват връзки с американски големи технологични фирми, за да получат достъп до

данни.

През април 2018 г. ЕС прие стратегия за изкуствен интелект и назначи Експертната група на високо ниво, която със своите 52-ма членове да предоставя съвети както за инвестициите, така и за етичните въпроси на управлението във връзка с ИИ в Европа. Самата експертна група е доста смесена – съставена е от изследователи и представители на университетите (в области като роботика, компютърни науки и философия), както и представители на индустрията (като Zalando, Bosch и Google) и организации на гражданското общество (като Access Now, ANEC и BEUC). На групата беше възложено да изготви проект на два документа: 1) етични насоки за ИИ<sup>19</sup> и 2) препоръки относно политиката и инвестициите в областта на ИИ.

В Насоките относно етичните аспекти за надежден ИИ Експертната група посочва, че надеждният изкуствен интелект има три основни характеристики: 1) той следва да е законосъобразен, като се гарантира спазването на всички приложими закони и подзаконови актове; 2) той следва да е етичен, като се гарантира зачитането на етичните принципи и ценности; 3) той следва да е стабилен от техническа и от социална гледна точка, така че да се гарантира, че при най-добри намерения системите с изкуствен интелект не причиняват неумишлено вреди.

Всяка от тези характеристики е необходима, но не е достатъчна за постигането на надежден изкуствен интелект. В идеалния случай всичките три характеристики работят в синхрон и се припокриват при своето функциониране.

Документът разглежда и четири етични принципа, които произтичат от основните права и които трябва да се спазват, за да се гарантира, че системите с ИИ са разработени, внедрени и използвани по надежден начин. Системите с ИИ следва да подобряват личното и колективното благосъстояние. Те са определени като етични императиви, така че професионалистите в областта на ИИ винаги да се стремят да се придържат към тях. Тези принципи са: 1) зачитане на автономността на хората; 2) предотвратяване на вреди; 3) справедливост; 4) обяснимост.

Според **принципа на зачитане на автономността на хората** основните права, на които се основава ЕС, трябва да са насочени към гарантиране именно на свободата и автономността на хората. Взаимодействайки си със системите с ИИ, те трябва да са в състояние да продължат цялостно и ефективно да се самоопределят и да могат да участват в демократичния процес. Системите с ИИ не следва необосновано да подчиняват, принуждават, заблуждават, манипулират, уговарят или подбират хората. Вместо това те трябва да бъдат проектирани така, че да разширяват, допълват и предоставят възможности за човешките познавателни, социални и културни умения. При разпределението на функциите между хората и системите с ИИ следва да се осигурява човешки надзор и контрол върху

работните процеси в системите с ИИ.<sup>20</sup>

**Принципът на предотвратяване на вреди** гласи, че системите с ИИ не трябва нито да причиняват, нито да засилват вреди или да оказват по друг начин неблагоприятно въздействие върху хората. Това включва защитата на човешкото достойнство, както и психическата и физическата неприкосновеност. Системите с ИИ и средите, в които те функционират, трябва да бъдат безопасни и сигурни, да бъдат технически стабилни и следва да се гарантира, че не са отворени за злонамерена употреба. Уязвимите лица следва да получат по-голямо внимание и да бъдат включени при разработването и внедряването на системите с ИИ. Специално внимание трябва да се обърне и на ситуации, в които системите с ИИ могат да причинят или да засилят неблагоприятните въздействия поради несъразмерност на правомощията или на достъпа до информация, например между работодатели и служители, предприятия и потребители или правителства и граждани. Предотвратяването на вредите включва също така вземането под внимание на естествената среда и на всички живи същества.<sup>21</sup>

**Принципът на справедливост** засяга разработването, внедряването и използването на системи с ИИ, които трябва бъдат справедливи. Макар да признава, че има много различни тълкувания на справедливостта, експертната група отчита, че тя има както материалноправно, така и процесуално измерение. Материалноправното измерение предполага ангажимент към: гарантиране на равномерно и справедливо разпределение на ползите, но и на разходите и гарантиране, че по отношение на лицата и групите лица няма несправедливи предубеждения, дискриминация и заклеймяване.

Използването на системи с ИИ, освен това, никога не трябва да води до заблуждаване или нарушаване на свободата на избор на (крайните) ползватели. Справедливостта предполага също, че професионалистите в областта на ИИ трябва да зачитат принципа на пропорционалност между средствата и целите и да обмислят внимателно как да намерят баланс между конкуриращите се интереси и цели. Процесуалното измерение на справедливостта включва способността да се оспорва и да се търси ефективна правна защита срещу решения, взети от системите с ИИ и от хората, които ги експлоатират. За да е възможно това, трябва да бъде установен субектът, отговорен за вземането на решението, а процесите на вземане на решения следва да бъдат обясними.<sup>22</sup>

**Принципът на обяснимост** е от съществено значение за изграждането и поддържането на доверието на ползвателите в системите с ИИ. Това означава, че процесите трябва да бъдат прозрачни, възможностите и целта на системите с ИИ да бъдат открито съобщавани, а решенията да бъдат разяснявани до степенята, в която е възможно, на пряко или косвено засегнатите от тях. Без такава информа-

ция няма как решението да се разбере и да се оспори, ако е необходимо. Невинаги обаче е възможно да се даде обяснение за причината, поради която даден модел е довел до конкретен резултат или решение (и каква комбинация от входящи фактори е допринесла за това). В тези случаи алгоритмите се оприличават на „черна кутия“ и се изисква специално внимание.<sup>23</sup>

В документа са очертани и 7 основни предпоставки, които е важно да се спазват в тази област: човешки контрол и автономност, многообразие от данни, прозрачност, безопасност и сигурност на данните, поверителност и защита на личния живот, защита на социално-етичното и екологичното благосъстояние и отчетност. Те трябва да се оценяват и управляват непрекъснато през целия жизнен цикъл на системата с ИИ.

Поставя се въпросът дали тези насоки и принципи за създаване на надежден ИИ ще работят на практика. Етичните насоки съдържат списък за практическа оценка, който да се използва от компаниите. През втората половина на 2019 г. над 350 организации са тествали този списък за оценка и са изпратили обратна връзка. Групата на високо ниво е в процес на преразглеждане на насоките в светлината на тази обратна връзка и ще завърши работа през 2020 г.

Бялата книга за ИИ от февруари 2020 г. също изразява позицията за координиран европейски подход относно човешките и етични последици от ИИ като размисъл за по-доброто използване на големи данни за иновации. Комисията подкрепя регулаторен и инвестиционен подход с това, което нарича „двойна цел за насърчаване на възприемането на ИИ и за справяне с рисковете, свързани с определени употреби на тази нова технология“. Бялата книга се стреми да изложи варианти на политиката за това как да се постигнат тези цели.<sup>24</sup> Основното предложение е възприемането на специфичен секторен подход към регулирането на ИИ, при който високорисковите приложения се разграничават от останалите. Като високорискови сектори, т.е. сектори, където „могат да се очакват значителни рискове“, се разглеждат „здравеопазване; транспорт; енергетика и части от публичния сектор“.<sup>25</sup>

В Бялата книга се отново се поставя акцент и върху концепцията за доверие: „Предвид голямото въздействие, което ИИ може да окаже върху нашето общество, и необходимостта от изграждане на доверие, жизненоважно е европейският ИИ да се основава на нашите ценности и основни права, като човешкото достойнство и защита на личния живот“.<sup>26</sup> Изразена цел на Политическата рамка на ЕС е да мобилизира ресурси за постигане на „екосистема на върховете постижения“ по „цялата верига на стойността“. Ключовите елементи на бъдещата регулаторна рамка за ИИ в Европа е да се създаде „уникална екосистема на доверие“, която се описва като политическа цел сама по себе си. От Комисията се надяват, че ясната

европейска регулаторна рамка би „изградила доверие сред потребителите и бизнеса в ИИ и следователно ще ускори внедряването на технологията“.<sup>27</sup>

Бялата книга ясно препраща към човекоцентричния подход. В нея се посочва, че Комисията ще вземе предвид приноса, постигнат по време на пилотната фаза на етичните насоки, изготвени от Експертна група на високо ниво по ИИ. Интересен е фактът, че според ЕК тези относно прозрачността, проследимостта и човешкия надзор не са специално обхванати от настоящото законодателство в много икономически сектори. Липсата на прозрачност, която Комисията изтъква, затруднява „идентифицирането и доказването на възможни нарушения на законите, включително законови разпоредби, които защитават основните права, приписват отговорност и отговарят на условията за искане на обезщетение“.<sup>28</sup> (Андреева, А., Йолова, Г., 2020)

Европейският парламент също работи активно в областта на установяване на обща етична рамка при употребата на технологии с ИИ. За него е изключително важно да бъдат включени следните етични принципи за разработването, внедряването и използването на изкуствен интелект, роботика и свързаните с тях технологии:

- ориентирани към човека, направени от човека и контролирани от човека изкуствен интелект, роботика и свързани с тях технологии;
- задължителна оценка на съответствието на високорисковия изкуствен интелект, роботика и свързаните с тях технологии;
- безопасност, прозрачност и отчетност;
- предпазни мерки и средства за правна защита срещу предубеденост и дискриминация;
- право на правна защита;
- социална отговорност и равенство между половете в изкуствения интелект, роботиката и свързаните с тях технологии;
- екологично устойчив изкуствен интелект, роботика и свързани с тях технологии;
- зачитане на неприкосновеността на личния живот и ограниченията за използването на биометрично разпознаване;
- добро управление на изкуствения интелект, роботиката и свързаните с тях технологии, включително данните, използвани или генерирани от такива технологии;
- способност да извличат поуки от натрупания опит и взаимодействие;
- формата на физическия носител на работа;
- способност да адаптират поведението и действията си към околната среда.<sup>29</sup>

Видно е, че има доста широко разбиране на това какви принципи да залегнат в бъдещата европейска регулация по темата. Европейският парламент предложи също въвеждането на Европейско сертифициране за етично съответствие. Целта е в контекста на координацията на равнище ЕС да се разработят общи критерии и процес на кандидатстване, свързани с предоставянето на европейски сертификат за съответствие, включително след искане от всеки разработчик, внедрител или ползвател на невисокорискови технологии да сертифицира положителната оценка на съответствието, извършена от национален надзорен орган. Според ЕП такъв европейски сертификат за етично съответствие би насърчил етичността още при проектирането по цялата верига на доставки на екосистемите с изкуствен интелект. Затова се предлага при високорисковите технологии сертифицирането да бъде задължителна предпоставка за допустимост до процедурите за възлагане на обществени поръчки в областта на изкуствения интелект, роботиката и свързаните с тях технологии.

## **Заклучение**

ЕС има потенциал да установява стандарти на глобално ниво по отношение на етиката на ИИ. Общите законодателни действия в тази сфера биха могли да засилят вътрешния пазар и да създадат важно стратегическо предимство. Въпреки че многобройни публични и частни участници по света са изготвили различни етични насоки, в момента няма изчерпателна правна рамка. Затова ЕС може да се възползва от отсъствието на конкуриращ се глобален модел на управление и да спечели пълните предимства на „първия двигател“. Въз основа на икономическите и регулаторни правомощия на ЕС общите законодателни действия ще осигурят на европейската индустрия конкурентно преимущество.

Приемането на стандарти в световен мащаб със сигурност няма как да се постигне чрез действия на отделни държави членки. Успехът зависи от наличието на общоевропейски, но и навременни законодателни решения със силен демократичен надзор, отчетност и ефективно прилагане.

Неизбежно е да се налагат рестрикции, за да се осигурят механизми за защита на човешките права и ценностите на ЕС, но тези рестрикции не бива да спират развитието на високотехнологичния бизнес. Затова е важно да се намери правилния и балансиран подход, който ще гарантира повече ползи и по-малко рискове. Според анализите на европейската оценка на добавената стойност общата етична рамка има потенциала да донесе на ЕС 294,9 млрд. евро допълнителен брутен вътрешен продукт и 4,6 млн. допълнителни работни места до 2030 г.

## Бележки

- <sup>1</sup> <https://www.bbc.com/news/technology-30290540>
- <sup>2</sup> <https://www.cnbc.com/2019/03/26/bill-gates-artificial-intelligence-both-promising-and-dangerous.html>
- <sup>3</sup> OECD. (2017). The Next Production Revolution, Implications for Governments and Business.
- <sup>4</sup> <https://www.idc.com/getdoc.jsp?containerId=prUS42439617>
- <sup>5</sup> US NDAA, 2019.
- <sup>6</sup> PwC. (2018). The macroeconomic impact of artificial intelligence.
- <sup>7</sup> The federal government: Artificial Intelligence Strategy, Status: November 2018.
- <sup>8</sup> [https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051\\_BG.html](https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_BG.html).
- <sup>9</sup> Съобщение на Комисията до Европейския парламент, Европейския съвет, Съвета, Европейския икономически и социален комитет и Комитета на регионите – Изкуствен интелект за Европа“ COM(2018) 237 final.
- <sup>10</sup> Системите с ИИ се проектират директно от хората, но последните могат да използват и техники, свързани с изкуствен интелект, за да оптимизират своя проект, Експертната група на високо равнище на ЕС по въпросите на ИИ дава в своя документ „ Определение за ИИ: основни възможности и дисциплини“ с. 6.
- <sup>11</sup> <https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1/language-bg/format-PDF>.
- <sup>12</sup> HLEG, 2019 ; Европейска стратегия за ИИ, 2018.
- <sup>13</sup> Резолюция на Европейския парламент от 20 октомври 2020 г. с препоръки към Комисията относно рамка за етичните аспекти на изкуствения интелект, роботиката и свързаните с тях технологии и Резолюция на Европейския парламент от 20 октомври 2020 г. с препоръки към Комисията относно режим на гражданска отговорност за изкуствения интелект.
- <sup>14</sup> Определението заема включеното в съобщението на Европейската комисия COM(2018)0237, 25.4.2018 г., с. 1, като е адаптирано.
- <sup>15</sup> Artificial intelligence ante portas: Legal & ethical reflections (europa.eu).
- <sup>16</sup> Харта на основните права на Европейския съюз (2016/C 202/02).
- <sup>17</sup> Discrimination in the European Union. Special Eurobarometer 493.
- <sup>18</sup> Cisco (2019). Consumer Privacy Survey: The Growing Imperative of Getting Data Privacy Right.
- <sup>19</sup> НАСОКИ ОТНОСНО ЕТИЧНИТЕ АСПЕКТИ ЗА НАДЕЖДЕН ИИ, Експертна група на високо равнище по въпросите на изкуствения интелект, 8 април 2019 г.



<sup>20</sup> Параграф 50 от Насоките относно етичните аспекти за надежден ИИ.

<sup>21</sup> Параграф 51 от Насоките относно етичните аспекти за надежден ИИ.

<sup>22</sup> Параграф 52 от Насоките относно етичните аспекти за надежден ИИ.

<sup>23</sup> Параграф 53 от Насоките относно етичните аспекти за надежден ИИ.

<sup>24</sup> Бяла книга за изкуствения интелект, Европейска Комисия, февруари 2020 г.

<sup>25</sup> Бяла книга за изкуствения интелект, Европейска Комисия, февруари 2020 г., с. 17.

<sup>26</sup> Бяла книга за изкуствения интелект, Европейска Комисия, февруари 2020 г., с. 2.

<sup>27</sup> Бяла книга за изкуствения интелект, Европейска Комисия, февруари 2020 г., с. 9 – 10.

<sup>28</sup> Бяла книга за изкуствения интелект, Европейска Комисия, февруари 2020 г., с. 14.

<sup>29</sup> Резолюция на Европейския парламент от 20 октомври 2020 г. с препоръки към Комисията относно рамка за етичните аспекти на изкуствения интелект, роботиката и свързаните с тях технологии (2020/2012(INL)).

### Използвана литература

1. Андреева, А. Йолова, Г., 2020. Трансформацията на правната връзка работодател-работник в резултат на влиянието на дигитализацията. 11(2020), pp. 11-18.

Andreeva, A. Yolova, G., 2020. Transformatsiyata na pravната vrazka rabotodatel-rabotnik v rezultat na vliyanieto na digitalizatsiyata. 11(2020), pp. 11-18.

2. Андреева, А., Йолова, Г., 2018. *За субекта на правото на труд и предизвикателствата на технологичното общество*. Варна, Сборник научни трудове, Largo Siti, pp. 320-328.

Andreeva, A., Yolova, G., 2018. *Za subekta na pravoto na trud i predizvikatelstvata na tehnologichното obshtestvo*. Varna, Sbornik nauchni trudove, Largo Siti, pp. 320-328.

3. Андреева, А., Йолова, Г., 2018. *Тенденции при антидискриминационните политики в сферата на здравното осигуряване в условията на дигиталното общество*. Варна, СТЕНО, pp. 255-261.

Andreeva, A., Yolova, G., 2018. *Tendentsii pri antidiskriminatsionните политики в sферата на здравното osiguryavane v usloviyata na digitalното obshtestvo*. Varna, STENO, pp. 255-261.

4. Андреева, А., Йолова, Г., 2020. Отговорността при инциденти на работното място, свързани с употреба на изкуствен интелект – съвременни предизвика-

телства пред законодателя. *Юридически сборник*, 27(2020), pp. 230-239.

Andreeva, A., Yolova, G., 2020. Otvornostta pri intsidenti na rabotnoto myasto, svarzani s upotreba na izkustven intelekt – savremenni predizvikatelstva pred zakonodatelya. *Yuridicheski sbornik*, 27(2020), pp. 230-239.

5. Albus, J. S. (1991). Outline for a theory of intelligence. *IEEE Transactions on Systems, Man, and Cybernetics*, 21, 473-509. doi:10.1109/21.97471, Published in: *IEEE Transactions on Systems, Man, and Cybernetics* (Volume: 21, Issue: 3, May/Jun 1991) Page(s): 473 - 509.

6. Andreeva, A., Yolova, G., Dimitrova, D., 2019. *Artificial Intellect: Regulatory Framework and Challenges Facing the Labour Market*. Ruse, New York: ACM, Digital Library, pp. 74-77.

7. Andreeva, A., Yolova, G., 2019a. *The Challenges of the Fourth Industrial Revolution Faced by the Labour Market: European and National Processes and Trends*. Varna, Science and Economics, University of Economics, pp. 178-188.

8. Andreeva, A., Yolova, G., 2019b. *The Right to Work in a Digital Society: Evolution and Trends*. Varna, Science and Economics, University of Economics - Varna, pp. 123-134.

9. Asimov, I., 1950. *I, Robot*. н.м.:н.а.

10. Bellman, R. (1978), *An introduction to Artificial Intelligence: Can Computers Think?*, Boyd&Fraser Publishing Company

11. Coeckelbergh, M. (2019) “Artificial Intelligence: Some ethical issues and regulatory challenges.” *Technology and Regulation*, 31-34.

12. Corinne Cath, Sandra Wachter, Brent Mittelstadt, Mariarosaria Taddeo & Luciano Floridi *Artificial Intelligence and the ‘Good Society’: the US, EU, and UK approach*, Springer, 28 March 2017

13. Dignum, V. (2019) *Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way*. Springer International Publishing.

14. Evas Tatjana, *Civil liability regime for artificial intelligence, European added value assessment, STUDY EPRS, European Parliamentary Research Service, September 2020*

15. Evas, T., 2020. *European framework on ethical aspects of artificial intelligence, robotics and related technologies European added value assessment*.

16. Evas Tatjana, *European framework on ethical aspects of artificial intelligence, robotics and related technologies, European added value assessment, STUDY EPRS | European Parliamentary Research Service, – September 2020*

17. Fast, E., & Horvitz, E. (2017). “Long-term trends in the public perception of artificial intelligence.”

18. Fogel D.B., Fogel L.J. (1996) *An introduction to evolutionary programming*.

In: Alliot JM., Lutton E., Ronald E., Schoenauer M., Snyers D. (eds) *Artificial Evolution*. AE 1995. Lecture Notes in Computer Science, vol 1063. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/3-540-61108-8\\_28](https://doi.org/10.1007/3-540-61108-8_28)

19. Gardner, H., 1983. *Multiple intelligences*. New York: Basic Books.

20. Gardner, H. (1987), The theory of multiple intelligences. *Annals of Dyslexia* 37, 19–35. <https://doi.org/10.1007/BF02648057>

21. Gardner, H. (1999). *Intelligence reframed: Multiple Intelligences for the 21st century*. New York: Basic Books.

22. Hagedorff, T. (2020). “The Ethics of AI Ethics: An Evaluation of Guidelines.” *Minds and Machines*, 1-22.

23. Jobin, A., Lenca, M., Vayena, E., (2019), The Global Landscape of AI Ethics Guidelines, *Nature Machine Intelligence* volume 1, pages389–399

24. Kaplan, 2016. *Artificial Intelligence: What Everyone Needs to Know.*

25. Koulu, R. (2020). “Human control over automation: EU Policy and AI Ethics.” *European Journal of Legal Studies*, 12(1): 9-46.

26. Kritikos Mihalis, *Artificial Intelligence ante portas: Legal & ethical reflections*, EPRS | European Parliamentary Research Service, Scientific Foresight Unit (STOA), March 2019

27. Luger, G.F., Stubblefield, W.A. (1993), *Artificial Intelligence*, Publisher: Benjamin-Cummings Publishing Co., Inc., Subs. of Addison-Wesley Longman Publ. Co390 Bridge Pkwy. Redwood City, CA, United States, ISBN:978-0-8053-4780-7

28. Madiega Tambiama, *EU guidelines on ethics in artificial intelligence: Context and implementation*, EPRS European Parliamentary Research Service, Members’ Research Service, September 2019

29. McCarthy, J., (2007), *What is Artificial Intelligence?*, Stanford University, <http://jmc.stanford.edu/articles/whatisai/whatisai.pdf>

30. McCarthy, J., Minsky, M.L., Rochester, N., Claude, E., (1955), *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence* August 31, 1955 Shannon AI Magazine Volume 27 Number 4 (2006) (© AAI).

31. Minsky, M., Papert, S., (1969); Review of ‘Perceptrons: An Introduction to Computational Geometry’, December 1969, *IEEE Transactions on Information Theory* 15(6):738- 739, DOI:10.1109/TIT.1969.1054388

32. Nathalie A. Smuha (2019), “The EU Approach to Ethics Guidelines for Trustworthy Artificial Intelligence”, *Computer Law Review International* Vol. 20; iss. 4; pp. 97 – 106

33. Neisser, U., Boodoo, G., (1996), *Intelligence: Knowns and Unknowns*, February 1996, *American Psychologist* 51(2):77-101, DOI:10.1037/0003-066X.51.2.77

34. Newell, A., Simon, H.A., (1976), *Computer Science as Empirical Inquiry:*

Symbols and Search, Communications of the ACM Volume 19 Issue 3 March 1976 pp 113–126, <https://doi.org/10.1145/360018.360022>

35. Nilsson, N., (1998), *Artificial Intelligence*, Stanford University, Morgan Kaufmann Publishers, Inc. San Francisco, California

36. Russell, S., Norvig, P., (1995), *Artificial Intelligence – A modern approach*, Pearson Education, Inc. New Jersey

37. Russel, Norvig, 2010. *Artificial Intelligence: A Modern Approach*. .

38. Samoili, S., 2020. *Defining artificial intelligence*. н.м.:European Commission.

39. Winston, P. H., (1992), *Artificial Intelligence*, Addison Wesley Publishing Company Massachusetts

40. Бяла книга за изкуствения интелект, Европейска Комисия, февруари 2020 г.

Byala kniga za izkustveniya intelekt, Evropeyska Komisiya, fevruari 2020 g.

41. HLEG, 2019; Европейска стратегия за ИИ, 2018

HLEG, 2019; Evropeyska strategiya za II, 2018

42. Насоки относно етичните аспекти за надежден ИИ, Експертна група на високо равнище по въпросите на изкуствения интелект 8 април 2019 г.

Nasoki odnosno etichnite aspekti za nadezhden II, Ekspertna grupa na visoko ravnishte po vaprosite na izkustveniya intelekt 8 april 2019 g.

43. Резолюция на Европейския парламент от 20 октомври 2020 г. с препоръки към Комисията относно рамка за етичните аспекти на изкуствения интелект, роботиката и свързаните с тях технологии (2020/2012(INL))

Rezolyutsiya na Evropeyskiya parlament ot 20 oktomvri 2020 g. s preporaki kam Komisiyata odnosno ramka za etichnite aspekti na izkustveniya intelekt, robotikata i svarzanite s tyah tehnologii (2020/2012(INL))

44. Съобщение на Комисията до Европейския парламент, Европейския съвет, Съвета, Европейския икономически и социален комитет и Комитета на регионите — Изкуствен интелект за Европа“ COM(2018) 237 final

Saobshtenie na Komisiyata do Evropeyskiya parlament, Evropeyskiya savet, Saveta, Evropeyskiya ikonomicheski i sotsialen komitet i Komiteta na regionite — Izkustven intelekt za Evropa“ COM(2018) 237 final

45. Харта на основните права на Европейския съюз (2016/C 202/02)

Harta na osnovnite prava na Evropeyskiya sayuz (2016/C 202/02)

46. Accenture (2018). *Realising the Economic and Societal Potential of Responsible AI in Europe*.

47. Accenture (2018). *Responsible AI: With Opportunity Comes Responsibility*.

48. Accenture (2019), *Building Data and AI Ethics Committees*.

49. *Artificial Intelligence and Civil Liability*, Legal Affairs, Policy Department for

Citizens' Rights and Constitutional Affairs, Directorate-General for Internal Policies

50. Artificial intelligence: From ethics to policy, STUDY, Panel for the Future of Science and Technology, EPRS European Parliamentary Research Service, Scientific Foresight Unit (STOA) – June 2020

51. Artificial intelligence: How does it work, why does it matter, and what can we do about it?, STUDY Panel for the Future of Science and Technology, EPRS European Parliamentary Research Service, Philip Boucher, Scientific Foresight Unit (STOA) – June 2020

52. Charter of Fundamental Rights of the European Union, OJ C 326 of 26.10.2012.

53. Cisco (2019). Consumer Privacy Survey: The Growing Imperative of Getting Data Privacy Right

54. Committee on Legal Affairs (2020). Draft report with recommendations to the Commission on a Civil liability regime for artificial intelligence (2020/2014(INL)).

55. Committee on Legal Affairs of the European Parliament. Report of 27.01.2017 with recommendations to the Commission on Civil Law Rules on Robotics, (2015/2103(INL)).

56. Conclusions of the European Council meeting of 19.10.2017 (EUCO 14/17).

57. Convention of the Council of Europe for the Protection of Human Rights and Fundamental Freedoms of 4 November 1950 (with protocols).

58. Discrimination in the European Union. Special Eurobarometer 493

59. EU Commission (2018) Artificial Intelligence for Europe. Brussels, 25.4.2018 COM(2018) 237 final.

60. EU Commission (2020) White Paper on Artificial Intelligence. A European approach to excellence and trust. COM(2020) 65 final.

61. European Commission. Building Trust in Human-Centric Artificial Intelligence, COM(2019) 168 of 08.04.2019.

62. European Commission. Coordinated Plan on Artificial Intelligence, COM(2018) 795 of 07.12.2018.

63. European Commission. Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics, COM(2020) 64 of 19.02.2020.

64. European Commission. Shaping Europe's Digital Future, February 2020.

65. OECD (2017), The Next Production Revolution: Implications for Governments and Business

66. PwC. (2018). The macroeconomic impact of artificial intelligence

67. The federal government: Artificial Intelligence Strategy, Status: November 2018

68. The High-Level Expert Group on Artificial Intelligence (2019b) A Definition of AI: Main Capabilities and Disciplines. Definition developed for the purpose of the

AI HLEG's deliverables. 8 April 2019.

69. The High-Level Expert Group on Artificial Intelligence (2019c) Policy and Investment Recommendations for Trustworthy Artificial Intelligence.

70. US NDAA, 2019.

<https://www.bbc.com/news/technology-30290540>

<https://www.cnbc.com/2019/03/26/bill-gates-artificial-intelligence-both-promising-and-dangerous.html>

<https://www.idc.com/getdoc.jsp?containerId=prUS42439617>

[https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051\\_BG.html](https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_BG.html)

<https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1/language-bg/format-PDF>

Artificial intelligence ante portas: Legal & ethical reflections (europa.eu)

## **THE ETHICAL CHALLENGES FACING THE EUROPEAN UNION USING THE ARTIFICIAL INTELLIGENCE**

**Emil RADEV**

### **Abstract**

The present article examines some aspects in application of the artificial intelligence and the related to it ethical challenges facing the European Union. The essence of the artificial intelligence is examined – definition and concepts of its application, as well as the European legal framework in the definition of the concept. The focus is put on the ethical challenges in the legal regulation of the artificial intelligence. In conclusion, based on the analysis summaries are formulated and proposals are made.

**Key words:** artificial intelligence, ethical challenges, ethical principles, definition for AI