ЕВРОПЕЙСКА
КОМИСИЯ

Брюксел, 19.2.2020 г.
COM(2020) 65 final

БЯЛА КНИГА

за изкуствения интелект — Европа в търсене на високи постижения и атмосфера на доверие

Бяла книга за изкуствения интелект — Европа в търсене на високи постижения и атмосфера на доверие

Изкуственият интелект се развива бързо. Той ще има положително отражение върху живота ни чрез по-добро здравеопазване, например с по-точно диагностициране и здравна превенция, по-ефективно селско стопанство, нови идеи в борбата с ограничаването и адаптирането към изменението на климата, по-голяма ефективност на системите за производство чрез прогнозна поддръжка, повече сигурност за европейците и ред други придобивки, които едва започваме да осъзнаваме. Същевременно изкуственият интелект (ИИ) крие редица потенциални рискове, като непрозрачност на процеса на вземане на решения, полова или друг вид дискриминация, нарушаване на правото ни на личен живот или престъпна употреба на ИИ.

На фона на ожесточената конкуренция в световен мащаб е необходим солиден европейски подход, който се основава на европейската стратегия за ИИ, представена през април 2018 г.[1] За да се възползва от предимствата на ИИ и в отговор на предизвикателствата, които той поставя, ЕС се нуждае от единодушие в определянето на собствен подход, основан на европейските ценности, чрез който да се насърчи разработването и внедряването на ИИ.

Комисията се ангажира да способства за постигането на научни пробиви, за запазването на технологичното лидерство на ЕС и за осигуряване на гаранции, че новите технологии ще бъдат в услуга на всички европейци — за подобряване на техния живот при зачитане на правата им.

В своите политически насоки[2] председателят на Комисията Урсула фон дер Лайен представи координиран европейски подход към човешките и етичните аспекти на технологиите с ИИ, както и по-задълбочен анализ как да бъдат използвани по-ефективно големите информационни масиви с цел новаторство.

В този смисъл Комисията подкрепя приемането на регулаторен и инвестиционно ориентиран подход, който едновременно би насърчил навлизането на ИИ и би ограничил рисковете, които пораждат някои приложения на тази нова технология. Целта на настоящата Бяла книга е да представи варианти на политики за постигането на тези цели. В нея не се разглежда разработването и използването на ИИ за военни цели. Комисията приканва държавите членки, другите европейски институции и всички заинтересовани страни, включително промишлеността, социалните партньори, организациите на гражданското общество, изследователите, обществеността като цяло и всички заинтересовани страни, да се произнесат върху посочените по-долу варианти и да внесат своя принос в бъдещите решения на Комисията в тази област.

1. ВЪВЕДЕНИЕ

Цифровите технологии присъстват все повече във всяка част от живота ни и затова е необходимо да се радват на доверието ни. Внедряването им би било невъзможно без такова доверие. Изграждането и запазването му играе в полза на Европа, която държи на своите ценности и върховенството на закона, както и на своя доказан капацитет да изгражда безопасни,

[1] Изкуствен интелект за Европа, COM (2018) 237 final.

[2] https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission_bg_1.pdf

надеждни и комплексни продукти и услуги в най-разнообразни области като аеронавтиката, енергетиката, автомобилостроенето и медицинското оборудване.

Настоящият и бъдещият устойчив икономически растеж и обществено благосъстояние на Европа все повече се опират на стопанските ползи от обработката на данни. ИИ е едно от най-важните приложения на основаната на данни икономика. Днес повечето данни са потребителски и се съхраняват и обработват в централизирана инфраструктура „в облак“. За сметка на това голям дял от бъдещите и далеч по-богати данни, чийто източник ще бъдат промишлеността, предприятията и публичният сектор, ще се съхраняват в различни системи, по-специално изчислителни устройства в периферията на мрежата. Това разкрива нови възможности за Европа, която има силна позиция в цифровизираните промишлени и междуфирмени приложения, но донякъде отстъпва при потребителските платформи.

С две думи ИИ е набор от технологии, които съчетават данни, алгоритми и изчислителна мощ. Напредъкът при изчислителните технологии и все по-големият обем от налични данни са основна причина за засиления интерес към ИИ. В Европейската стратегия за данните се посочва[3], че Европа може да съчетае своите предимства в областта на технологиите и промишлеността с висококачествена цифрова инфраструктура и регулаторна рамка, основана на нейните основни ценности, **и да се превърне в световен лидер при иновациите в основаната на данни икономика и нейните приложения**. На тази основа тя може да разработи „екосистема“ за ИИ, която да пренесе ползите от тази нова технология в европейското общество и икономика като цяло:

- за **гражданите** — например по-добро здравеопазване, по-дълготрайни домакински уреди, по-безопасни и по-чисти транспортни системи, по-добри обществени услуги;
- за **развиването на стопанска дейност** — например продукти и услуги от ново поколение в области, в които Европа е доказала капацитета си (машиностроене, транспорт, киберсигурност, селско стопанство, зелена и кръгова икономика, здравеопазване и отраслите с висока добавена стойност, като модата и туризма); както и
- за услугите от **обществен интерес** — например намаляване на разходите за предоставяне на услуги (в транспорта, образованието, енергетиката и управлението на отпадъците), подобряване на устойчивостта и енергийната ефективност на продуктите[4] и осигуряване на точните инструменти за правоприлагащите органи с оглед на защитата на гражданите[5] и подходящи гаранции за техните права и свободи.

Предвид потенциалното значително въздействие на ИИ върху нашето общество и необходимостта от изграждане на доверие Европа е длъжна да обвърже разработките си в тази област със своите ценности и с основни човешки права като например човешкото достойнство и защитата на неприкосновеността на личния живот.

---

[3] COM(2020) 66 final.

[4] ИИ и цифровизацията като цяло са ключови фактори за постигане на амбициозните цели, заложени в европейския зелен пакт. Текущият екологичен отпечатък на сектора на ИКТ обаче се оценява на над 2 % от всички емисии в световен мащаб. В европейската стратегия в областта на цифровите технологии, придружаваща настоящата Бяла книга, се съдържат мерки за екологосъобразно преобразуване на сектора на цифровите технологии.

[5] Инструментите с ИИ имат потенциала да осигурят по-добра защита от престъпления и терористични актове за гражданите на ЕС. Те биха могли например да помогнат при засичането на терористична пропаганда онлайн, подозрителни трансакции при продажбата на опасни продукти, представляващи опасност скрити предмети или забранени вещества или продукти или при оказване на помощ на граждани в извънредни ситуации и насочване на екипите за оказване на първа помощ.

Въздействието на системите с ИИ следва да се разглежда не само от индивидуална гледна точка, но и от гледна точка на обществото като цяло. Използването на системи с ИИ може значително да спомогне за постигането на целите за устойчиво развитие и да подсили демократичния процес и социалните права. С неотдавнашните си предложения в европейския зелен пакт[6] Европа се превърна в пример за подражание при решаването на проблемите, свързани с промяната на климата и с екологията. Без цифрови технологии като ИИ е невъзможно да се постигнат целите на Зеления пакт. Предвид нарастващото значение на ИИ въздействието на системите с ИИ върху околната среда заслужава внимание през целия им жизнен цикъл и по цялата верига на доставки, например по отношение на използването на ресурси за обучението на алгоритми и съхранението на данни.

За постигането на мащабност и за да се избегне разпокъсването на единния пазар е необходим общоевропейски подход към ИИ. Въвеждането на национални инициативи може да застраши правната сигурност, да отслаби доверието на гражданите и да осуети появата на динамична европейска промишленост.

Настоящата Бяла книга представя политически варианти, които позволяват надеждно и сигурно разработване на ИИ в Европа при пълно зачитане на ценностите и правата на европейските граждани. Основните градивни елементи на Бялата книга са:

- рамката на политиката, в която се определят мерки за съгласуване на усилията на европейско, национално и регионално равнище. Тя има за цел, чрез партньорство между частния и публичния сектор, да мобилизира необходимия ресурс, с който да се изгради **„екосистема за високи постижения“** по цялата верига за създаване на стойност, като се започне от научните изследвания и иновациите, и да се създадат правилните стимули за ускореното внедряване на решения, основани на ИИ, включително в малките и средните предприятия (МСП);

- основните елементи на една бъдеща регулаторна рамка за ИИ в Европа, която ще създаде уникална **„екосистема на доверие“.** За тази цел тя трябва да гарантира спазването на правилата на ЕС, включително тези за защита на основните права и правата на потребителите, особено спрямо системите с ИИ, които работят в ЕС и представляват висок риск[7]. Изграждането на екосистема на доверие е само по себе си политическа цел, която следва да разсее притесненията на гражданите да използват приложения с ИИ и да осигури на предприятията и обществените организации необходимата правна сигурност за иновации, използващи ИИ. Комисията твърдо подкрепя антропоцентричния подход, основан на Съобщението относно изграждането на доверие в ориентирания към човека ИИ[8] и ще отчете отзивите от пилотната фаза на етичните насоки, изготвени от експертната група на високо равнище по въпросите на ИИ.

Европейската стратегия за данните, която придружава настоящата Бяла книга, има за цел да даде възможност на Европа да се превърне в най-привлекателната, сигурна и динамична икономика в света, която разчита на данните за своята адаптивност, като ѝ осигури данните, необходими за намиране на по-добри решения и подобряване на живота на всички нейни граждани. В стратегията се определят редица политически мерки, включително набирането на

[6] COM(2019) 640 final.

[7] Въпросът за възможното въвеждане на допълнителни мерки за предотвратяване и противодействие на злоупотребата с ИИ за престъпни цели е извън обхвата на настоящата Бяла книга.

[8] COM (2019) 168.

необходимите за целта частни и публични инвестиции. Освен всичко това, в доклада на Комисията към настоящата Бяла книга се анализира отражението на ИИ, интернет на предметите и други цифрови технологии върху законодателството в областта на безопасността и отговорностите.

2. ОПОЛЗОТВОРЯВАНЕ НА СИЛНИТЕ СТРАНИ НА ПРОМИШЛЕНИТЕ И ПРОФЕСИОНАЛНИТЕ ПАЗАРИ

Европа е в изгодна позиция и може да се възползва от потенциала на ИИ не само като потребител, но и като създател и производител на тази технология. Тя има отлични научноизследователски центрове, новаторски стартиращи предприятия, водеща позиция в световен мащаб в областта на роботиката и други конкурентоспособни отрасли в производството и услугите — от автомобилостроенето и здравеопазването до енергетиката, финансовите услуги и селското стопанство. Европа разви силни изчислителни мощности (например способност за високоскоростни изчисления), което е от съществено значение за функционирането на ИИ. Европа разполага също така с големи обеми от публични и промишлени данни, чиито потенциал не се използва понастоящем в достатъчна степен. Нейната промишленост се слави с безопасността и сигурността на своите цифрови системи с ниска консумация на енергия, които са от съществено значение за по-нататъшното развитие на ИИ.

Впрягането на капацитета на ЕС за инвестиции в технологии и инфраструктури от следващо поколение, както и в умения като грамотността в областта на цифровите технологии, ще укрепи технологичния суверенитет на Европа в ключови базови технологии и инфраструктури в основаната на данни икономика. Инфраструктурите следва да подкрепят създаването на европейски масиви от данни, позволяващи разработването на надежден ИИ, тоест ИИ, основан на европейските ценности и правила.

Европа следва да се възползва от своите силни страни и да заеме водещи позиции в „екосистемите“ и по цялата верига за създаване на стойност — от някои отрасли за производство на хардуер през софтуера до услугите. Това вече се случва в известна степен. Европа произвежда повече от една четвърт от всички промишлени и професионални роботи (например в прецизното земеделие, сигурността, здравеопазването, логистиката) и играе важна роля в разработването и използването на софтуерни приложения за дружества и организации (в междуфирмени приложения като софтуера за планиране на ресурсите на предприятията или този за проектиране и инженеринг), както и на приложения в услуга на електронното управление и „интелигентните предприятия“.

Европа е водеща и във внедряването на ИИ в производството. Над половината от водещите производители използват ИИ най-малко в един етап от производството[9].

Европа отчасти дължи силната си позиция в научните изследвания на програмата на ЕС за финансиране, която се доказа при съгласуването на действията, избягването на дублиране и привличането на публични и частни инвестиции в държавите членки. През последните три години финансирането от ЕС на научни изследвания и иновации в областта на ИИ достигна 1,5 милиарда евро, което представлява увеличение с 70 % в сравнение с предходния период.

Същевременно инвестициите в научни изследвания и иновации в Европа все още представляват само малка част от публичните и частни инвестиции в тази област в други части на света. През

[9] Следват Япония (30 %) и САЩ (28 %). Източник: CapGemini (2019).

2016 г. в Европа са инвестирани около 3,2 милиарда евро в ИИ, докато в Северна Америка и Азия тази сума възлиза съответно на 12,1 и 6,5 милиарда евро[10]. Съответно Европа трябва да увеличи значително своите инвестиции. Координираният план за ИИ[11], разработен заедно с държавите членки, се доказва като добро начало за изграждането на по-тясно сътрудничество в областта на ИИ в Европа и за създаването на синергии за привличане на максимални инвестиции във веригата за създаване на стойност при технологиите с ИИ.

3. НОВИ ВЪЗМОЖНОСТИ: СЛЕДВАЩАТА ВЪЛНА ОТ ДАННИ

Европа изостава при потребителските приложения и онлайн платформите, което спъва достъпа ѝ до повече данни, а същевременно протичат важни промени в оценката и повторното използване на данни в различни отрасли. Обемът на произведените в света данни нараства бързо — от 33 зетабайта през 2018 г. на прогнозирани 175 зетабайта през 2025 г.[12] Всяка нова вълна от данни отваря възможности за Европа да се позиционира в една световна икономика, която разчита на данните за своята адаптивност, и да се превърне в световен лидер в тази област. Освен това начинът, по който данните се съхраняват и обработват, ще се измени неимоверно през следващите пет години. Понастоящем 80 % от обработката и анализа на данни в облак се извършват в центрове за данни и централизирани изчислителни системи, а 20 % — в интелигентни свързани обекти, като автомобили, домакински уреди или роботи, и в компютърни системи в близост до потребителя (периферни изчисления). До 2025 г. се очакват съществени изменения в тези съотношения[13].

Европа е световен лидер в електрониката със слаба консумация на енергия, което е от ключово значение за следващото поколение специализирани процесори за ИИ. В момента този пазар се доминира от участници извън ЕС. Това може да се промени посредством инициативи като тази за европейския процесор, която поставя ударение върху разработването на изчислителни системи с ниска консумация на енергия както за периферни изчисления, така и за високопроизводителни изчислителни технологии от следващо поколение, и като работата по линия на съвместното предприятие за ключови цифрови технологии, чието начало е планувано за 2021 г. Европа също така е лидер при невроморфните решения[14], които идеално отговарят на нуждите за автоматизирането на промишлените процеси (Индустрия 4.0) и транспортните средства. Те могат да подобрят енергийната ефективност с няколко порядъка.

Неотдавнашният напредък в областта на квантовите изчисления ще генерира експоненциално нарастване на капацитета за обработка[15]. Европа може да застане начело при разработването на тази технология благодарение на своите академични постижения в областта на квантовите изчисления, както и на силните позиции на европейската промишленост в областта на квантовите симулатори и на програмната среда за квантови изчисления. Европейските инициативи, насочени към увеличаване броя на съоръженията за експериментиране и

[10] 10 imperatives for Europe in the age of AI and automation (10 императива за Европа в епохата на ИИ и автоматизацията) McKinsey, (2017).
[11] COM(2018) 795.
[12] IDC (2019).
[13] Gartner (2017).
[14] Невроморфни решения означава всяка широкомащабна система от интегрални схеми, които имитират невро-биологични структури.
[15] Квантовите компютри ще разполагат с капацитета да обработват в рамките на секунди много по-големи набори от данни, отколкото позволяват днешните високопроизводителни компютри. Това ще позволи разработването на нови приложения на ИИ в различни отрасли.

изпитване в областта на квантовите изчисления, ще помогнат за прилагането на тези нови квантови решения в ред промишлени и академични отрасли.

Успоредно с това Европа ще продължи да играе водеща роля в алгоритмичните основи на ИИ, доразвивайки собствените си високи научни постижения. Нужно е да се изградят мостове между понастоящем несвързани в разработките си отделни дисциплини, като машинното самообучение и задълбоченото машинно учене (характеризиращи се с ограничена тълкуваемост, необходимост от голям обем данни за обучение на моделите и обучаване чрез корелации) и символните подходи (където правилата се задават от човек). Съчетаването на подхода на символни съждения с дълбоките невронни мрежи може да ни помогне да подобрим обяснимостта на резултатите от ИИ.

4. ЕКОСИСТЕМА ЗА ВИСОКИ ПОСТИЖЕНИЯ

За изграждането на екосистема за високи постижения, която да подкрепи разработването и навлизането на ИИ в икономиката и публичната администрация на ЕС, е необходимо да се усили работата на няколко равнища.

А. СЪТРУДНИЧЕСТВО С ДЪРЖАВИТЕ ЧЛЕНКИ

В изпълнение на своята стратегия за ИИ[16], приета през април 2018 г., през декември 2018 г. Комисията представи изготвен съвместно с държавите членки координиран план, който насърчава разработването и използването на технологии с изкуствен интелект в Европа[17].

В него се предлагат около 70 съвместни действия за по-тясно и по-ефикасно сътрудничество между държавите членки и Комисията в ключови области като научните изследвания, инвестициите, внедряването на пазара, развитието на уменията и талантите, управлението на данни и международното сътрудничество. Предвижда се планът да продължи до 2027 г. с редовен мониторинг и преглед.

Целта е да се извлече допълнителна максимална полза от инвестициите в научни изследвания, иновации и внедряване, да се оценят националните стратегии в областта на ИИ и да се доразработи и развие координираният план в областта на ИИ заедно с държавите членки.

- *Действие 1: Комисията ще вземе предвид резултатите от обществената консултация по Бялата книга и ще предложи на държавите членки преразгледана версия на координирания план, която да бъде приета до края на 2020 г.*

Финансирането на равнище ЕС в областта на ИИ следва да привлече и обедини инвестициите в областите, в които самостоятелните действия на отделните държави членки не са достатъчни. Целта е всяка година да се привличат над 20 млрд. евро[18] общи инвестиции в ЕС в областта на ИИ през следващите десет години. За да стимулира частните и публичните инвестиции, ЕС ще осигури ресурси по програмите „Цифрова Европа“, „Хоризонт Европа“, както и по линия на европейските структурни и инвестиционни фондове, в отговор на нуждите на по-слабо развитите региони, както и на селските райони.

Координираният план би могъл също така да разглежда общественото и екологичното благосъстояние като основен принцип в областта на ИИ. Системите с ИИ крият потенциал за

[16] „Изкуствен интелект за Европа“, COM (2018) 237.
[17] Координиран план за изкуствения интелект, COM (2018) 795.
[18] COM(2018) 237.

решаването на най-неотложни проблеми, включително в областта на климатичните изменения и влошаването на състоянието на околната среда. Важно е също така тези решения да са екологосъобразни. ИИ може и е нужно самостоятелно да оценява критично потреблението на ресурси и енергия и да се обучава във вземането на решения, които са от полза за околната среда. Заедно с държавите членки Комисията ще разглежда възможностите за насърчаване и популяризиране на подобни решения в областта на ИИ.

Б. С ОБЩИ УСИЛИЯ ЗА НАУЧНИ ИЗСЛЕДВАНИЯ И ИНОВАЦИИ

Европа не може да си позволи да запази настоящата разпокъсаност на специализираните експертни центрове, тъй като нито един от тях не разполага с необходимия размах, за да се конкурира с водещите институти по света. Наложително е да се създадат повече полезни взаимодействия и връзки между различните европейски научноизследователски центрове в областта на ИИ и да се синхронизират усилията им за подобряване на техните постижения, за запазване и привличане на най-добрите изследователи и за разработване на върхови технологии. Европа трябва да разполага с водещ център за научни изследвания, иновации и експертен опит, който да координира тези усилия и да се превърне в световен еталон за високи постижения в областта на ИИ, както и да привлича инвестиции и най-добрите таланти в тази област.

Центровете и мрежите следва да се концентрират в отраслите, в които Европа има потенциала да се превърне в световен лидер, като промишлеността, здравеопазването, транспорта, веригите за създаване на стойност в хранително-вкусовата промишленост, енергетиката/околната среда, горското стопанство, наблюдението на Земята и космическите изследвания. Във всички тези области тече надпревара за световно лидерство, а Европа разполага със значителен потенциал, знания и експертен опит [19]. Еднакво важно е да се създадат обекти за изпитване и експериментиране в подкрепа на разработването и последващото внедряване на нови приложения с ИИ.

- *Действие 2: Комисията ще способства за създаването на центрове за високи постижения и за изпитвания, които да се ползват едновременно от европейски, национални и частни инвестиции, евентуално чрез нов правен инструмент. В рамките на програмата „Цифрова Европа“ и, според случая, посредством действия за научноизследователски и иновационни дейности по програмата „Хоризонт Европа“ от многогодишната финансова рамка за периода 2021—2027 г. Комисията предложи да се задели значителна сума специално в подкрепа на центровете за изпитване в Европа, представляващи световен еталон.*

В. УМЕНИЯ

В подкрепа на европейския подход към ИИ особено внимание трябва да се отдели на придобиването на умения, за да се преодолее недостигът на квалифицирани кадри[20]. Комисията скоро ще представи допълнение на програмата за умения, която има за цел да гарантира, че всички граждани на Европа могат да се възползват от екологичната и цифровата трансформация на икономиката на ЕС. Инициативите биха могли също така да включват

[19] Бъдещият Европейски фонд за отбрана и постоянното структурирано сътрудничество (ПСС) също ще осигурят възможности за научноизследователска и развойна дейност в областта на ИИ. Тези проекти следва да бъдат синхронизирани с по-всеобхватните граждански програми на ЕС, посветени на ИИ.

[20] https://ec.europa.eu/jrc/en/publication/academic-offer-and-demand-advanced-profiles-eu

подкрепа за секторните регулатори и усъвършенстване на уменията им в областта на ИИ с цел ефективно и ефикасно прилагане на съответните правила. Актуализираният план за действие в областта на цифровото образование ще спомогне за по-ефективното използване на данните и основаните на ИИ технологии, като машинното обучение и прогнозните анализи, с цел подобряване на системите за образование и обучение и съобразяването им с цифровата ера. Планът също така ще популяризира ИИ на всички равнища на образование, за да подготви гражданите за информирани решения, които все повече ще зависят от ИИ.

Развиването на уменията, необходими за работа в областта на ИИ, адаптирането на образователните системи и повишаването на квалификацията на работната сила, за да бъде тя в крак с обусловената от ИИ трансформация, ще бъдат заложени приоритетно в преразгледания координиран план за ИИ, който ще бъде разработен с държавите членки. За тази цел списъкът за оценка в етичните насоки може да залегне в един вид примерна „програма“ за разработчиците на ИИ, която да се предостави като спомагателен ресурс на учебните заведения. Специални усилия следва да се положат за увеличаване броя на жените, които се обучават и намират заетост в тази област.

Освен това един водещ център за научни изследвания и иновации в областта на ИИ в Европа може да привлече таланти от цял свят с възможностите, които може да им предложи. Той може също така да развива и разпространява върхови умения, които да намерят почва и да бъдат дообогатени в цяла Европа.

- *Действие 3: Създаване и подкрепа на мрежи от водещи университети и институти за висше образование посредством стълба за придобиване на задълбочени цифрови умения по програмата „Цифрова Европа“, с цел привличане на преподаватели и учени от най-високо равнище и предлагане на водещи в световен мащаб магистърски програми в областта на ИИ.*

Освен въпроса за повишаването на квалификацията, работниците и работодателите са пряко засегнати от проектирането и използването на системи с ИИ на работното място. Участието на социалните партньори ще бъде решаващ фактор за гарантиране на антропоцентричния подход към ИИ на работното място.

Г. АКЦЕНТ ВЪРХУ МСП

От значение е и МСП да разполагат с достъп до ИИ и да го използват. За тази цел центровете за цифрови иновации[21] и платформата за ИИ по заявка[22] следва да получат допълнителна подкрепа и да насърчават сътрудничеството между МСП. Програмата „Цифрова Европа“ ще способства за постигането на тази цел. Всички центрове за цифрови иновации следва да подкрепят МСП в разбирането и внедряването на ИИ, но е важно поне един иновационен център на държава членка да има висока степен на специализация в областта на ИИ.

МСП и стартиращите предприятия ще се нуждаят от финансиране за адаптирането на своите процеси или за да новаторстват, ползвайки ИИ. Посредством бъдещия пилотен инвестиционен фонд от 100 милиона евро за финансиране в областта на ИИ и блок-веригите Комисията планира да разшири допълнително достъпа до финансиране в областта на ИИ по линия на

[21] ec.europe.eu/digital-single-market/en/news/digital-innovation-hubs-helping-companies-across-economy-make-most-digital-opportunities

[22] www.Ai4eu.eu

InvestEU[23]. ИИ присъства изрично сред областите, отговарящи на условията за използване на гаранцията InvestEU.

- *Действие 4: Комисията ще работи с държавите членки, за да гарантира, че поне един център за цифрови иновации във всяка държава членка има висока степен на специализация в областта на ИИ. Центровете за цифрови иновации могат да бъдат подкрепяни по линия на програмата „Цифрова Европа“.*

- *Комисията и Европейският инвестиционен фонд ще стартират пилотна схема за 100 милиона евро през първото тримесечие на 2020 г. за осигуряване на капиталово финансиране за иновативни разработки в областта на ИИ. При условие че бъде постигнато окончателно споразумение за МФР, намерението на Комисията е схемата да се разрасне значително считано от 2021 г. посредством InvestEU.*

Д. ПАРТНЬОРСТВО С ЧАСТНИЯ СЕКТОР

От съществено значение е също така да се гарантира, че частният сектор участва пълноценно в планирането на приоритетите за научните изследвания и иновациите и осигурява необходимото равнище на съфинансиране. Това изисква създаването на широко публично-частно партньорство и осигуряването на ангажимент от страна на висшето ръководство на дружествата.

- *Действие 5: По линия на програмата „Хоризонт Европа“ Комисията ще създаде ново публично-частно партньорство в областта на ИИ, данните и роботиката, за съчетаване на усилията, осигуряване на координация на научните изследвания и иновациите в областта на ИИ, сътрудничество с други публично-частни партньорства в рамките на „Хоризонт Европа“ и съвместна работа с центровете за изпитване и посочените по-горе центрове за цифрови иновации.*

Е. НАСЪРЧАВАНЕ НА ВНЕДРЯВАНЕТО НА ИИ В ПУБЛИЧНИЯ СЕКТОР

От голямо значение е публичните администрации, болниците, комуналните и транспортните услуги, финансовите надзорни органи и други области от обществен интерес да започнат бързо да предлагат продукти и услуги, ползващи ИИ. Приоритетни ще бъдат здравеопазването и транспортът, при които технологията е вече добре развита и позволява широкомащабно внедряване.

- *Действие 6: Комисията ще инициира отворен и прозрачен диалог по отрасли, с предимство на здравеопазването, администрациите в селските райони и операторите на обществени услуги, за да представи план за действие с цел улесняване на разработването, експериментирането и внедряването. Резултатите от диалога по отрасли ще се използват при изготвянето на специална „Програма за внедряване на ИИ“, която ще бъде в услуга на обществените поръчки за набавяне на системи с ИИ и ще спомогне за преобразуването на самите процеси на възлагане на обществени поръчки.*

Ж. ОСИГУРЯВАНЕ НА ДОСТЪП ИНФРАСТРУКТУРИ, СВЪРЗАНИ С ДАННИ И ИЗЧИСЛЕНИЯ

[23] Europe.eu/investeu

Областите на действие, посочени в настоящата Бяла книга, допълват плана, който е представен успоредно в рамките на Европейската стратегия за данните. Подобряването на достъпа до данни и тяхното управление е ключов въпрос. Без данни развитието на ИИ и на други приложения в областта на цифровите технологии е невъзможно. Огромният обем нови данни, които предстои да бъдат генерирани, представлява възможност Европа да се позиционира начело на основаващата се на данни и ИИ трансформация. Насърчаването на отговорни практики в управлението на данни и съответствието на данните с принципите FAIR ще допринесе за изграждането на доверие и ще осигури възможност за многократно използване на данните[24]. Също толкова важни са инвестициите в ключови изчислителни технологии и инфраструктури.

Комисията предложи да бъдат заделени над 4 милиарда евро в подкрепа на високопроизводителните и квантовите изчислителни технологии, включително периферните изчисления, инфраструктурата, свързана с ИИ, както и на данните и инфраструктурата „в облак“. Европейската стратегия за данните доразвива тези приоритети.

3. МЕЖДУНАРОДНО ИЗМЕРЕНИЕ

Европа е в изгодна позиция и може да играе водеща роля в световен мащаб за привличането на съмишленици около общи ценности и за популяризиране на етичното използване на ИИ. Работата на ЕС в областта на ИИ вече има отражение върху темите, които се дискутират на международно ниво. При разработването на своите етични насоки експертната група на високо равнище се допита до редица организации извън ЕС и няколко правителствени наблюдатели. Успоредно с това ЕС участва активно в разработването на етичните принципи на ОИСР за ИИ[25]. Впоследствие Г-20 одобри тези принципи в Декларацията на министрите от юни 2019 г. относно търговията и цифровата икономика.

Наред с това ЕС отчита, че важна работа в областта на ИИ се осъществява и в рамките на други многостранни форуми, сред които са Съветът на Европа, Организацията за образование, наука и култура на ООН (ЮНЕСКО), Организацията за икономическо сътрудничество и развитие (ОИСР), Световната търговска организация и Международния съюз по далекосъобщения (МСД). В рамките на ООН ЕС следи отзивите на доклада на групата на високо равнище за сътрудничество в областта на цифровите технологии и нейната препоръка относно ИИ.

ЕС ще продължи да си сътрудничи приоритетно с държави със сходни възгледи, но и с глобални участници в областта на ИИ, залагайки на правилата и ценностите на ЕС (напр. подкрепа за възходящото регулаторно сближаване, достъп до ключови ресурси, включително данни, създаване на равни условия). Комисията ще наблюдава отблизо политиките на трети държави, които ограничават потоците от данни, и ще засегне необоснованите ограничения в своите двустранни търговски преговори и чрез действия в рамките на Световната търговска организация. Комисията е убедена, че международното сътрудничество при ИИ трябва да се основава на подход, който насърчава зачитането на основните права, включително човешкото достойнство, плурализма, интеграцията, недискриминацията и защитата на неприкосновеността

[24] FAIR — Findable, Accessible, Interoperable and Reusable, т.е. данните да са лесни за намиране, достъпни, оперативно съвместими и многократно използваеми, както е посочено в окончателния доклад и плана за действие на експертната група на Комисията относно данните „FAIR“, 2018 г., https://ec.europa.eu/info/sites/info/files/turning_fair_into_reality_1.pdf.

[25] https://www.oecd.org/going-digital/ai/principles/

на личния живот и личните данни[26], и ще се стреми да популяризира своите ценности по света[27]. Също така е ясно, че отговорното разработване и използване на ИИ може да бъде движеща сила за постигане на целите за устойчиво развитие и тези на Програмата до 2030 г.

5. ЕКОСИСТЕМА НА ДОВЕРИЕ: РЕГУЛАТОРНА РАМКА ЗА ИИ

Както при всяка нова технология, използването на ИИ крие както възможности, така и рискове. Гражданите се страхуват да не изгубят възможността да защитават правата и безопасността си, когато се сблъскат с информационната асиметрия на вземаните чрез алгоритми решения, а дружествата се притесняват за правната сигурност. ИИ може да спомогне за защитата на сигурността на гражданите и да им позволи да се ползват от основните си права, но те се опасяват, че ИИ може да има нежелани последствия или дори да бъде използван за зли цели. Тези опасения трябва да бъдат разсеяни. Нещо повече, по-широкото използване на ИИ е спъвано не само от липсата на инвестиции и умения — основен фактор в това отношение е отсъствието на доверие.

Ето защо на 25 април 2018 г. Комисията представи стратегия за ИИ[28], съобразена със социално-икономическите аспекти и разглеждаща допълнителните инвестиции в научни изследвания, иновации и капацитет за ИИ в целия ЕС. Тя постигна съгласие с държавите членки по координиран план[29] за съгласуване на стратегиите. Комисията създаде също така експертна група на високо равнище, която през април 2019 г. публикува насоки относно надеждния ИИ[30].

Комисията публикува съобщение[31], в което се приветстват седемте ключови изисквания, посочени в насоките на експертната група на високо равнище:

- Човешки фактор и надзор,
- Техническа стабилност и безопасност,
- Управление на данните и неприкосновеност на личния живот,
- Прозрачност,
- Многообразие, недискриминация и справедливост,
- Обществено и екологично благополучие,
- Отчетност.

Освен това насоките съдържат списък за оценка на практическото използване на ИИ от страна на дружествата. През втората половина на 2019 г. този списък за оценка беше изпробван от над 350 организации, които изпратиха отзивите си. Групата на високо равнище е в процес на преразглеждане на насоките в светлината на тази обратна информация и ще приключи работата си до юни 2020 г. Този процес на събиране на обратна информация откроi факта, че макар редица изисквания да присъстват вече в съществуващите правни или регулаторни режими, изискванията за прозрачност, проследимост и човешки надзор не са конкретно обхванати от настоящото законодателство в много икономически сектори.

---

[26] По линия на Инструмента за партньорство Комисията ще финансира проект на стойност 2,5 милиона евро в подкрепа на сътрудничеството с партньори със сходни възгледи за съобразяване с етичните насоки на ЕС в областта на ИИ и за приемане на общи принципи и оперативни заключения.

[27] Политически насоки на председателя Урсула фон дер Лайен, „Съюз с по-големи амбиции: моята програма за Европа“, стр. 17.

[28] COM(2018) 237.

[29] COM(2018) 795.

[30] https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top

[31] COM(2019) 168.

В допълнение към този набор от необвързващи насоки на експертната група на високо равнище и в съответствие с политическите насоки на председателя на Комисията, една ясна европейска регулаторна рамка ще създаде доверие в ИИ у потребителите и предприятията и следователно ще ускори възприемането на технологията. Подобна регулаторна рамка следва да е съобразена с други действия за насърчаване на иновационния капацитет и конкурентоспособността на Европа в тази област. Освен това тя трябва да гарантира оптимални в социално, екологично и икономическо отношение резултати, както и спазване на законодателството, принципите и ценностите на ЕС. Това е от особено значение в областите, в които правата на гражданите са потенциално най-уязвими, например когато ИИ се използва в правоприлагането и съдебната система.

Разработчиците и внедрителите на технологии с ИИ вече са обект на европейското законодателство в областта на основните права, например защитата на данните, неприкосновеността на личния живот, недискриминацията, защитата на потребителите, както и на правилата относно безопасността на продуктите и отговорността. Потребителите очакват еднакви права и равнище на безопасност, независимо дали даден продукт или система използва ИИ или не. Някои присъщи на ИИ характеристики, например неговата непрозрачност, могат обаче да затруднят прилагането и изпълнението на въпросното законодателство. Затова е необходимо да се проучи дали действащото законодателство е в състояние да отговори на свързаните с ИИ рискове и може да се прилага ефективно, дали е необходимо законодателството да се адаптира или е необходимо ново законодателство.

Предвид темпа на развитие на свързаните с ИИ технологии регулаторната рамка трябва да се поддава и на доразработване. Всички промени следва да бъдат ограничени до ясно определени проблеми, за които съществуват приложими решения.

Държавите членки посочват отсъствието на общоевропейска рамка. Германската комисия по етика в областта на данните призова за регулаторна система с пет равнища, основана на оценка на риска, при която повечето безобидни системи с ИИ не подлежат на никакво регулиране, а най-опасните са напълно забранени. Дания наскоро въведе прототип на знак за етичност в областта на данните. Малта въведе доброволна система за сертифициране на ИИ. Ако ЕС не осигури общоевропейски подход, съществува реален риск от разпокъсване във вътрешния пазар, което би подкопало целите на доверието, правната сигурност и внедряването на пазара.

Една солидна европейска регулаторна рамка за надеждни технологии с ИИ би осигурила защита за всички европейски граждани и би спомогнала за създаването на безпроблемен вътрешен пазар за по-нататъшното разработване и внедряване на ИИ, както и за укрепването на промишлената база на Европа в областта на ИИ.

A. ОПРЕДЕЛЯНЕ НА ПРОБЛЕМА

ИИ носи много ползи, сред които е повишаването на безопасността на някои продукти и процеси, но той може да донесе и вреди. Те могат да са материални (за безопасността и здравето, включително загуба на човешки живот, имуществени щети) и нематериални (нарушаване на неприкосновеността на личния живот, ограничения на правото на свобода на изразяване, накърняване на човешкото достойнство, дискриминация, например при достъпа до заетост и др.) и могат да са свързани с най-разнообразни рискове. Всякаква регулаторна рамка би следвало да се съсредоточи върху начините за свеждане до минимум на различните рискове от потенциални вреди, особено на по-съществените.

Основните рискове, свързани с използването на ИИ, се отнасят до прилагането на правилата за защита на основните права (сред които са защитата на личните данни и неприкосновеността на

личния живот и недискриминацията), както и до въпроси, свързани с безопасността[32] и отговорността.

Рискове за основните права, включително неприкосновеността на личния живот, защитата на личните данни и недискриминацията

Използването на ИИ крие опасност от нарушаване на основните права[33], включително свободата на словото, свободата на събрания, човешкото достойнство, принципа на недискриминация на полова, расова или етническа основа, на основа на вероизповедание или религия, увреждания, възраст или сексуална ориентация, според случая в различните области, защитата на личните данни и личния живот[34] или правото на ефективна съдебна защита и справедлив съдебен процес, както и защитата на потребителите. Тези рискове могат да се дължат на недостатъци в цялостното проектиране на системите с ИИ (включително по отношение на човешкия надзор) или на използването на данни без коригиране на потенциална необективност (например системата се обучава само или предимно с данни за мъже, което води до неоптимални резултати по отношение на жените).

ИИ може да изпълнява много задачи, които в миналото можеха да извършват единствено хора. Тоест все по-често граждани и правни субекти ще стават обект на действия и решения, които се вземат от системи с ИИ или с помощта на такива системи и чието разбиране и, при необходимост, ефективно оспорване може понякога да се окаже трудно. Освен това ИИ открива повече възможности за проследяване и анализ на ежедневните навици на хората. Например, съществува потенциален риск ИИ да бъде използван, в нарушение на правилата на Съюза за защита на данните и други правила, от държавни органи или други органи за масово наблюдение, както и от работодатели, за да наблюдават поведението на своите служители. ИИ е способен да анализира големи количества данни и да разпознава взаимовръзки в тях, което може да се използва също за проследяване и свързване на определени данни с определени лица, което поражда нови рискове за защитата на личните данни дори по отношение на набори от данни, които не включват сами по себе си лични данни. ИИ се използва също така от онлайн посредници за приоритизиране на информацията, която предлагат на своите потребители, и за модериране на съдържанието. Обработваните данни, проектантският замисъл на приложенията и възможностите за човешка намеса могат да се отразят върху правото на свобода на словото, защита на личните данни и неприкосновеност на личния живот, както и върху политическите свободи.

[32] Това включва въпроси от областта на киберсигурността, приложенията на ИИ в жизненоважни инфраструктури или относно злонамереното използване на ИИ.

[33] Проучване на Съвета на Европа показва, че голям брой основни права могат да бъдат засегнати при използването на ИИ,https://rm.coe.int/algorithms-and-human-rights-en-rev/16807956b5.

[34] Общият регламент относно защитата на данните и Директивата за правото на неприкосновеност на личния живот и електронни комуникации (в момента се договаря нов Регламент за правото на неприкосновеност на личния живот и електронни комуникации) третират тези рискове, но може да е необходимо да се проучи дали системите с ИИ създават допълнителни рискове. Комисията неотклонно ще следи и оценява прилагането на ОРЗД.

Някои алгоритми за ИИ, използвани за предсказване на престъпни рецидиви, могат да бъдат необективни и дискриминиращи по полов или расов признак и да предсказват различна вероятност от рецидив при жените в сравнение с мъжете или при гражданите на съответната държава в сравнение с чужденците. *Източник: Tolan S., Miron M., Gomez E. and Castillo C. „Why Machine Learning May Lead to Unfairness: Evidence from Risk Assessment for Juvenile Justice in Catalonia" (Защо машинното самообучение може да доведе до неравноправие: доказателства от оценката на риска при правораздаването за непълнолетни лица в Каталония), награда за най-добра статия, Международна конференция по въпросите на изкуствения интелект и правото, 2019 г.*

При някои програми с ИИ за анализ на портретни снимки се наблюдава необективност на полова или расова основа. Например, системата разпознава по-правилно пола на мъже със светла кожа, но често греши при определянето на пола на жени с по-тъмна кожа. Източник: *Joy Buolamwini, Timnit Gebru; Proceedings of the 1st Conference on Fairness, Accountability and Transparency (Резултати от първата конференция за справедливост, отчетност и прозрачност), PMLR 81:77-91, 2018.*

Липсата на обективност и дискриминацията са рискове, присъщи за всяка обществена или стопанска дейност. Човешките решения също са податливи на грешки и предубеждения. Когато обаче имаме необективност на решенията вследствие използването на ИИ, това може да има далеч по-сериозни последствия, да засегне и дискриминира много хора, които няма да могат да се възползват от предпазните механизми за социален контрол, на които се подчиняват човешките действия[35]. Това може да се случи и когато системата с ИИ „се обучава", докато е в експлоатация. В такива случаи, когато резултатът не може да се предотврати или предвиди на етапа на проектиране, рисковете няма да произтичат от недостатък в първоначалния проект на системата, а по-скоро от практическите последици от корелациите или моделите, които системата разпознава в голям набор от данни.

Някои характеристики, присъщи на технологиите в областта на ИИ, сред които са непрозрачност („ефект на черната кутия"), сложността, непредвидимостта и донякъде самостоятелното поведение, могат да затруднят проверката на спазването на действащите законови правила на ЕС, предназначени да защитават основните права, и да попречат на ефективното им прилагане. Правоприлагащите органи и засегнатите лица може да не разполагат с начин да проверят как е било взето дадено решение с участието на ИИ и следователно дали са били спазени съответните правила. Когато подобни решения засягат отрицателно физически и юридически лица, същите могат да срещнат трудности да си осигурят действителен достъп до правосъдие.

Рискове за безопасността и ефективното функциониране на режима на отговорност

[35] Консултативният комитет на Комисията за равните възможности на жените и мъжете подготвя понастоящем „Становище относно изкуствения интелект", в което анализира, наред с другото, въздействието на ИИ върху равенството между половете. Очаква се то да бъде прието от Комисията в началото на 2020 г. В стратегията на ЕС за равенство между половете за периода 2020—2024 г. се разглежда и ролята на ИИ в сферата на равенството между половете; Европейската мрежа на националните органи за равно третиране (Equinet) се очаква да публикува в началото на 2020 г. доклад (от Robin Allen и Dee Masters) на тема „Регулиране на ИИ: новата роля на органите за равно третиране — посрещане на новите предизвикателства пред равенството и недискриминацията вследствие на нарастващата цифровизация и използването на ИИ).

Технологиите в областта на ИИ могат да породят нови рискове за безопасността на потребителите, когато те са вградени в продукти и услуги. Например, при недостатък в технологията за разпознаване на предмети, един автономен автомобил може погрешно да идентифицира определен обект на пътя и да причини злополука с наранявания и материални щети. Тези рискове могат, както рисковете за основните права, да бъдат породени от грешки в проектирането на технологиите за ИИ, да се дължат на недостатъчна наличност и качество на данните или на други проблеми, произтичащи от машинното самообучение. Някои от тези рискове не се ограничават до продукти и услуги, които разчитат на ИИ, но използването на ИИ може да ги увеличи или утежни.

Освен опасностите, които поражда за физическите лица, липсата на ясни разпоредби за ограничаването на рисковете може да създаде правна несигурност за предприятията, които предлагат на пазара на ЕС продукти с ИИ. Органите за пазарен надзор и правоприлагащите органи могат да се окажат в положение, в което не са наясно дали могат да се намесят, тъй като не са оправомощени да действат и/или не разполагат с подходящи технически възможности за контрол на системите[36]. Следователно правната несигурност може да намали общите равнища на безопасност и да постави под въпрос конкурентоспособността на европейските дружества.

Ако се появят рискове за безопасността, липсата на ясни изисквания и посочените по-горе характеристики на технологиите за ИИ ще затруднят проследяването на потенциално проблемните решения, взети с участието на системи с ИИ. Това от своя страна може да затрудни обезщетяването на ощетените съгласно действащото законодателство на ЕС и националните законодателства, които регулират отговорността[37].

Съгласно Директивата относно отговорността за вреди от стоки производителят носи отговорност за вредите, причинени от дефектни стоки. Същевременно, при системи с ИИ, като например автономните автомобили, може да бъде трудно да се докаже дефект в продукта, настъпилата вреда и причинно-следствената връзка между тях. Освен това съществува известна несигурност относно това как и до каква степен Директивата относно отговорността за вреди от стоки се прилага за някои видове дефекти, например ако те са резултат от слабости в киберсигурността на продукта.

Затова трудностите в проследяването на потенциално проблемните решения, свързани с основните права и взети от системи с ИИ, важат в еднаква степен и за въпросите, свързани с безопасността и отговорността. Например, ощетените лица може да не разполагат с ефективен достъп до доказателствата, които са необходими за образуване на дело в съда, и да разполагат с по-малко действителни възможности за правна защита в сравнение със ситуации, при които вредите са причинени посредством традиционни технологии. Тези рискове ще нарастват с повсеместното използване на ИИ.

Б. ВЪЗМОЖНИ КОРЕКЦИИ В СЪЩЕСТВУВАЩАТА ЗАКОНОДАТЕЛНА РАМКА НА ЕС, КОЯТО СЕ ОТНАСЯ ДО ИИ

[36] Пример за това са интелигентните часовници за деца. Продуктът не може да навреди пряко на детето, което го носи, но без минимално ниво на сигурност, той може лесно да се използва за достигане до детето. Намесата на органите за пазарен надзор може да е затруднена в случаи, когато рискът не е свързан с продукта като такъв.

[37] В доклада на Комисията към настоящата Бяла книга се анализира отражението на ИИ, интернет на предметите и други цифрови технологии върху законодателството в областта на безопасността и отговорностите.

Голяма част от съществуващото законодателство на ЕС в областта на безопасността на продуктите и отговорността[38], включително специфичните отраслови правила, допълнени от национални закони, засяга редица нововъзникващи приложения на ИИ и е възможно да се прилага за тях.

Що се отнася до защитата на основните права и правата на потребителите, законодателната рамка на ЕС включва например Директивата за расовото равенство[39], Директивата за равно третиране в областта на заетостта и професиите[40], директивите относно равното третиране на мъжете и жените по отношение на заетостта и достъпа до стоки и услуги[41], редица правила за защита на потребителите[42], както и правилата относно защитата на личните данни и неприкосновеността на личния живот, по-специално Общия регламент относно защитата на данните и друго секторно законодателство относно защитата на личните данни, като например Директивата относно правоприлагането в областта на защитата на данните[43]. Освен това от 2025 г. ще се прилагат правилата относно изискванията за достъпност на стоки и услуги, предвидени в Европейския акт за достъпността[44]. Също така, при прилагането на други законодателни актове на ЕС, включително в областта на финансовите услуги, миграцията или отговорността на онлайн посредниците, трябва да се зачитат основните права.

Законодателството на ЕС по принцип остава напълно приложимо, независимо от включването на ИИ, но е важно да се прецени дали то може да се прилага адекватно в отговор на рисковете, които пораждат системите с ИИ, или е необходимо някои правни инструменти да бъдат адаптирани.

Например стопанските субекти продължават да носят пълна отговорност за това ИИ да отговаря на съществуващите правила за защита на потребителите. Не се допуска каквото и да е алгоритмично използване на потребителските навици, което е в нарушение на съществуващите правила.

Комисията е на мнение, че законодателната рамка би могла да бъде подобрена в отговор на следните рискове и ситуации:

- *Ефективно прилагане и изпълнение на съществуващото законодателство на ЕС и националното законодателство*: някои основни характеристики на ИИ създават предизвикателства пред гарантирането на правилното прилагане и изпълнение на законодателството на ЕС и националното законодателство. Липсата на прозрачност при ИИ затруднява разкриването и доказването на възможни закононарушения, включително на правните разпоредби, които защитават основните права, както и чия е

[38] Правната рамка на ЕС за безопасността на продуктите се състои от Директивата относно общата безопасност на продуктите (Директива 2001/95/ЕО), в ролята на предпазна мрежа, и от редица специфични отраслови правила, обхващащи различни категории продукти — от машини, самолети и леки автомобили до детски играчки и медицински изделия — с цел осигуряване на високо равнище на здравеопазване и безопасност. Законодателството относно отговорността за вреди от стоки се допълва от различни системи за гражданска отговорност за вреди, причинени от продукти или услуги.

[39] Директива 2000/43/ЕО.

[40] Директива 2000/78/ЕО.

[41] Директива 2004/113/ЕО; Директива 2006/54/ЕО.

[42] Например Директивата за нелоялните търговски практики (Директива 2005/29/ЕО) и Директивата за правата на потребителите (Директива 2011/83/ЕО).

[43] Директива (ЕС) 2016/680 на Европейския парламент и на Съвета от 27 април 2016 г. относно защитата на физическите лица във връзка с обработването на лични данни от компетентните органи за целите на предотвратяването, разследването, разкриването или наказателното преследване на престъпления или изпълнението на наказания и относно свободното движение на такива данни, и за отмяна на Рамково решение 2008/977/ПВР на Съвета.

[44] Директива (ЕС) 2019/882 за изискванията за достъпност на продукти и услуги.

отговорността и дали ощетените отговарят на условията за предявяване на иск за обезщетение. Поради това, за да се гарантира ефективно правоприлагане, може да се наложи коригиране или изясняване на съществуващото законодателство в някои области, например в областта на отговорността, както е описано по-подробно в доклада, който придружава настоящата Бяла книга.

- *Ограничения в обхвата на съществуващото законодателство на ЕС*: законодателството на ЕС в областта на безопасността на продуктите поставя особено ударение върху пускането на пазара на продукти. Според законодателството на ЕС в областта на безопасността на продуктите, когато един софтуер е част от краен продукт, той трябва да е съобразен със съответните правила за безопасност на продуктите. Остава открит обаче въпросът дали самостоятелният софтуер е обхванат от законодателството на ЕС в областта на безопасността на продуктите, освен в някои сектори с изрични правила[45]. Общото текущо законодателство на ЕС в областта на безопасността се прилага за продукти, а не за услуги, и следователно по принцип не се отнася и до услугите, основани на ИИ (например здравни услуги, финансови услуги, транспортни услуги).

- *Промяна на функционалността на системите с ИИ*: интегрирането в продуктите на софтуер, включително ИИ, може да изменя функционирането на тези продукти и системи по време на жизнения им цикъл. Това се отнася особено за системите, които изискват често актуализиране на софтуера или които разчитат на машинно самообучение. Всичко това може да породи нови рискове, които не са съществували при пускането на системата на пазара. Въпросните рискове не са разгледани в достатъчна степен в съществуващото законодателство, което поставя ударение предимно на рисковете за безопасността, които съществуват към момента на пускането на пазара на даден продукт.

- *Несигурност по отношение на разпределянето на отговорностите между различните икономически оператори по веригата на доставки*: като цяло законодателството на ЕС относно безопасността на продуктите държи отговорен производителя на продукта, който се пуска на пазара, заедно с неговите компоненти, например системи с ИИ. Правилата обаче може да станат неясни, ако страна, която не е производител, добави ИИ към продукта след неговото пускане на пазара. Освен това законодателството на ЕС относно отговорността за продуктите държи отговорни производителите и оставя националните правила в областта да уреждат отговорността на другите лица по веригата за доставки.

- *Промени в концепцията за безопасност*: използването на ИИ в продукти и услуги може да доведе до рискове, които понастоящем не са изрично посочени в законодателството на ЕС. Те могат да са свързани с киберзаплахи, рискове за личната безопасност (свързани с нови приложения на ИИ, например в домакинските уреди), рискове които произтичат от загуба на връзка с интернет и т.н. Тези рискове могат да са налице в момента на пускане на продуктите на пазара или да възникнат в резултат на

[45] Например софтуер, предназначен от производителя за използване за медицински цели, се счита за медицинско изделие съгласно Регламента за медицинските изделия (Регламент (ЕС) 2017/745).

актуализиране на софтуера или машинно самообучение при използването на продукта. ЕС следва да използва пълноценно инструментите, с които разполага, за да обогати натрупаните до момента свидетелства за потенциални рискове, свързани с използването на ИИ, включително опита на Агенцията на ЕС за киберсигурност (ENISA) при оценяването на разнообразието от рискове, свързани с ИИ.

Както бе посочено по-горе, няколко държави членки вече проучват възможности за национално законодателство за справяне с предизвикателствата, породени от ИИ. Това поражда риск от

Доклад относно отражението на изкуствения интелект, интернет на предметите и роботиката върху безопасността и отговорностите

Докладът, който придружава настоящата Бяла книга, анализира съответната правна рамка. В него се набелязват несигурните аспекти на прилагането на тази рамка спрямо специфичните рискове, породени от системите с ИИ и други цифрови технологии.

В доклада се стига до заключението, че действащото законодателство в областта на безопасността на продуктите вече съдържа широко разбиране на понятието за защита срещу всички видове рискове, произтичащи от продукта, в зависимост от употребата му. Същевременно, за да се осигури по-голяма правна сигурност, може да се въведат разпоредби, които да обхващат изрично новите рискове, свързани с новопоявяващите се цифрови технологии.

- Автономното поведение на някои системи с ИИ през техния жизнен цикъл може да доведе до значителни промени в продуктите с отражение върху безопасността, което може да наложи преоценка на риска. Освен това, като предпазна мярка, може да се наложи човешки надзор от самото проектиране и през целия жизнен цикъл на продуктите и системите с ИИ.
- Възможно е да бъдат разгледани изрични задължения за производителите и по отношение на рисковете за психичното здраве на потребителите, ако е целесъобразно (напр. при работа с човекоподобни роботи).
- Законодателството на Съюза в областта на безопасността на продуктите може да съдържа специфични изисквания в отговор на рисковете за безопасността вследствие използването на ненадеждни данни на етапа на проектирането, както и механизми за гарантиране на запазване на качеството на данните по време на използването на продуктите и системите с ИИ.
- Непрозрачността на системите, основани на алгоритми, може да се преодолее чрез изисквания за прозрачност.
- Може да се наложи адаптиране и изясняване на съществуващите правила за използвания самостоятелно софтуер, предлаган на пазара отделно от продукти или изтеглян и инсталиран в продукти след пускането им на пазара, когато това се отразява на безопасността.
- Предвид нарастващата сложност на веригите на доставки по отношение на новите технологии, въвеждането на специални изисквания за сътрудничество между стопанските субекти във веригата на доставки и потребителите би могло да осигури правна сигурност.

Характеристиките на някои нови цифрови технологии, като ИИ, интернет на предметите и роботиката, могат да поставят на изпитание рамката от правила, свързани с отговорността, и да намалят тяхната ефективност. Някои от тези характеристики могат да затруднят проследяването на вредата до дадено лице, което би било необходимо в случай на иск на основание виновна отговорност в съответствие с повечето национални правила. Това може значително да увеличи разходите за ощетените и да затрудни предявяването на искове за отговорност срещу лица, различни от производителите.

- Лицата, които са претърпели вреди вследствие действия на системи с ИИ трябва да се ползват със същото равнище на защита, както при вреди, причинени от други технологии, но това не бива да спъва развитието на технологичните нововъведения.
- Нужно е внимателно да се преценят всички варианти за гарантиране на тази цел, включително евентуалните изменения на Директивата относно отговорността за вреди от стоки и евентуалното допълнително целенасочено хармонизиране на националните правила за отговорността. Например Комисията събира мнения за това дали и до каква степен би имало нужда от ограничаване на последствията от сложността, като се адаптира доказателствената тежест, налагана по националните правила за отговорност за вреди, причинени от приложения с ИИ.

разпокъсване на единния пазар. Различните национални правила могат да създадат пречки за предприятията, които искат да продават и управляват системи с ИИ в рамките на единния пазар. Един общ подход на равнище ЕС ще позволи на европейските дружества да се възползват от безпрепятствен достъп до единния пазар и ще подсили тяхната конкурентоспособност на световните пазари.

От обсъждането по-горе Комисията заключава, че — освен възможните корекции в съществуващото законодателство — може да се наложи приемането на ново законодателство, специално за ИИ, за да може правната рамка на ЕС да бъде съобразена с настоящото и очакваното технологично и търговско развитие.

B. ОБХВАТ НА БЪДЕЩАТА РЕГУЛАТОРНА РАМКА НА ЕС

От основно значение за бъдещата специална регулаторна рамка относно ИИ е определянето на нейния обхват. Работната хипотеза е, че регулаторната рамка ще се прилага за продукти и услуги, които разчитат на ИИ. Тоест ИИ следва да притежава ясно определение за целите на настоящата Бяла книга, както и за всяка бъдеща инициатива за разработване на политики.

В своето съобщение „Изкуствен интелект за Европа“ Комисията представи едно първо определение за ИИ[46]. То беше доуточнено от експертната група на високо равнище[47].

Във всеки нов правен инструмент определението за ИИ ще трябва да бъде достатъчно гъвкаво, за да отразява техническия напредък, но и достатъчно точно, за да осигурява необходимата правна сигурност.

За целите на настоящата Бяла книга, както и на евентуални бъдещи дискусии относно политическите инициативи, изглежда важно да се изяснят основните елементи на ИИ, тоест „данните“ и „алгоритмите“. ИИ може да се интегрира в хардуер. При техниките за машинно самообучение, които са тясно свързани с ИИ, алгоритмите се обучават да правят заключения за определени модели въз основа на набор от данни, за да определят действията,

> Например, при автономните превозни средства алгоритъмът използва в реално време данните от автомобила (скорост, разход на гориво, амортисьори и др.) и от датчиците, които сканират околната среда на автомобила (пътно платно, знаци, други превозни средства, пешеходци и т.н.), за да изчисли посоката, ускорението и скоростта, които трябва да възприеме автомобилът, за да достигне определена дестинация. Въз основа на наблюдаваните данни алгоритъмът се адаптира към ситуацията на пътя и към външните условия, включително към поведението на другите водачи, за да избере най-удобния и безопасен пробег.

[46] COM(2018) 237 final, стр. 1: „Наименованието „изкуствен интелект“ (ИИ) се използва за системи, които показват интелигентно поведение, като анализират своята среда и — с известна степен на самостоятелност — предприемат действия за постигане на конкретни цели.
Базираните на ИИ системи могат да бъдат изцяло софтуерни — действащи във виртуалния свят (напр. гласови асистенти, софтуер за анализ на изображения, търсачки, системи за разпознаване на глас и лица), а могат и да бъдат внедрени в хардуерни устройства (напр. усъвършенствани роботи, автономни автомобили, дронове или приложения за интернет на предметите).“

[47] Експертна група на високо равнище, определение за ИИ, стр. 8: „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 вземане 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.“

които са необходими за постигането на дадена цел. Алгоритмите могат да продължат да се обучават в процеса на използването им. Макар продуктите с ИИ да притежават автономност, като се ориентират в заобикалящата ги среда и без да следват предварително определен набор от инструкции, тяхното поведение е до голяма степен определено и ограничено от техните разработчици. Хората са тези, които определят и програмират целите, за които системата с ИИ следва да се оптимизира.

ЕС разполага със строга правна рамка, за да гарантира, наред с останалото, защитата на потребителите, борбата с нелоялните търговски практики и защитата на личните данни и неприкосновеността на личния живот. Освен това достиженията на правото на ЕС съдържат специфични правила за някои отрасли (напр. здравеопазване, транспорт). Тези съществуващи разпоредби на правото на ЕС ще продължат да се прилагат спрямо ИИ, въпреки че може да са необходими известни актуализации на тази рамка, за да бъде отразена цифровата трансформация и използването на ИИ (вж. раздел Б). Тоест аспектите, които са вече уредени в съществуващото хоризонтално или секторно законодателство (например в областта на медицинските изделия[48], в транспортните системи) и занапред ще се регулират от него.

Принципно новата регулаторна рамка за ИИ следва да бъде ефективна и да постига своите цели без да бъде прекалено предписателна, за да не се създаде непропорционална тежест, особено за МСП. За да се постигне този баланс, Комисията е на мнение, че следва да се придържа към един основан на риска подход.

Той ще гарантира, че регулаторната намеса е пропорционална. Този подход обаче изисква ясни критерии, за да се разграничават различните приложения на ИИ, по-специално по линия на това дали те са „високорискови“[49]. Определението за високорисково приложение на ИИ следва да е ясно, лесно разбираемо и приложимо за всички заинтересовани страни. Същевременно, дори дадено приложение на ИИ да не се приема за високорисково, то остава изцяло в обхвата на вече съществуващите правила на ЕС.

Комисията е на мнение, че дадено приложение на ИИ по принцип следва да се приема за високорисково в зависимост от залозите, тоест дали едновременно секторът и предвижданата употреба предполагат значителни рискове, по-специално за защитата на безопасността, правата на потребителите и основните права. По-конкретно, дадено приложение на ИИ следва да се приема за рисково, когато отговаря на следните два кумулативни критерия:

- първо, приложението на ИИ е в сектор, в който преобладаващите дейности и техните характеристики предполагат наличието на значителни рискове. Този първи критерий гарантира, че регулаторна намеса ще има в областите, където като цяло се приема, че е най-вероятно да възникнат рискове. Новата регулаторна рамка следва да съдържа изричен и изчерпателен списък на обхванатите сектори. Например здравеопазване; транспорт; енергетика и части от публичния сектор[50]. Списъкът следва да се преразглежда периодично и, когато е необходимо, да се изменя в зависимост от промените в практиката;

[48] Например съображенията за безопасност и правните последици се различават според това дали системите с ИИ предоставят специализирана медицинска информация на лекари, медицинска информация директно на пациента или изпълняват медицински задачи пряко върху пациента. Комисията разглежда тези характерни за сферата на здравните грижи предизвикателства пред безопасността и определянето на носещите отговорност.

[49] Законодателството на ЕС може да категоризира „рисковете“ по начин, който е различен от описания тук, в зависимост от областта, като например безопасността на продуктите.

[50] Публичният сектор би могъл да включва области като правото на убежище, миграцията, граничния контрол и съдебната система, социалната сигурност и службите по заетостта.

- второ, ИИ се прилага във въпросния сектор така, че има вероятност да възникнат значителни рискове. Този втори критерий отразява допускането, че не всяко използване на ИИ в избраните сектори неизбежно предполага значителни рискове. Например, макар здравеопазването да е типичен пример за принципно рисков сектор, неправилното функциониране на системата за насрочване на прегледи в някоя болница в общия случай няма да породи значителни рискове, които да оправдаят законодателна намеса. Оценката на равнището на риска за дадено приложение може да се основава на последствията за засегнатите страни. Например приложения на ИИ, които водят до правни или други подобни съществени последствия за правата на дадено физическо лице или дружество; които пораждат риск от нараняване, смърт или значителни материални или нематериални щети; които пораждат последствия, които физическите или юридическите лица не могат да избегнат освен с извънредни усилия.

Прилагането на двата кумулативни критерия би гарантирало, че обхватът на регулаторната рамка е целенасочен и осигурява правна сигурност. Задължителните изисквания в новата регулаторна рамка относно ИИ (вж. раздел Г по-долу) по принцип ще се прилагат само за заявленията, които са набелязани като високорискови в съответствие с двата кумулативни критерия.

Независимо от горепосоченото, може да има и изключителни случаи, при които поради възможните рискове използването на приложения с ИИ за определени цели следва да се счита за високорисково само по себе си — тоест независимо от въпросния сектор и ако посочените по-долу изисквания все още са в сила[51]. Може да се посочи по-специално следното:

- с оглед на неговото значение за физическите лица и за достиженията на правото на ЕС относно равенството в областта на заетостта, използването на приложения с ИИ в процеса на набиране на персонал, както и в ситуации с отражение върху правата на работниците, винаги ще се счита за „високорисково“ и поради това посочените по-долу изисквания ще се прилагат по подразбиране. Могат да бъдат разгледани и други специфични приложения, засягащи правата на потребителите;

- използването на приложения с ИИ за целите на дистанционната биометрична идентификация [52] и на други технологии за наблюдение, които нарушават неприкосновеността на личния живот, винаги ще се счита за „високорисково“ и поради това посочените по-долу изисквания ще се прилагат по подразбиране.

Г. ВИДОВЕ ИЗИСКВАНИЯ

При разработването на бъдещата регулаторна рамка за ИИ ще бъде необходимо да се вземат решения кои видове задължителни правни изисквания да бъдат наложени на различните

[51] Важно е да се подчертае, че други законодателни актове на ЕС също могат да са приложими. Например, когато ИИ се интегрира в потребителски продукти, Директивата относно общата безопасност на продуктите може да се прилага към безопасността на приложенията с ИИ.

[52] Дистанционната биометрична идентификация следва да се разграничава от биометричната проверка (втората е мярка за сигурност, която се основава на уникалните биологични характеристики на лицето и има за цел удостоверяване истинността на неговите/нейните твърдения относно неговата/нейната самоличност). За дистанционна биометрична идентификация говорим, когато от разстояние се установява самоличността на множество лица с помощта на биометрични идентификатори (пръстови отпечатъци, портретна снимка, ирис, съдови модели и т.н.), на обществено място и по постоянен или непрекъснат начин, като данните се сверяват с база данни.

участници. Тези изисквания могат да се доуточняват впоследствие чрез стандарти. Както е отбелязано в раздел В по-горе и в допълнение към вече съществуващото законодателство, тези изисквания ще се прилагат само за високорискови приложения с ИИ, като по този начин се гарантира, че всяка регулаторна намеса е целенасочена и пропорционална.

Предвид насоките на експертната група на високо равнище и гореизложеното, изискванията за високорискови приложения с ИИ биха могли да се състоят от следните основни елементи, които са разгледани по-подробно в следващите подраздели:

- данни за обучение;
- данни и документация;
- информация за предоставяне;
- надеждност и точност;
- човешки надзор;
- специфични изисквания за някои конкретни приложения на ИИ, като използваните за дистанционна биометрична идентификация.

За да се гарантира правна сигурност, тези изисквания ще бъдат доуточнени, за да се осигури ясен ориентир за всички участници, които трябва да ги спазват.

а) Данни за обучение

Повече от всякога е необходимо да се насърчават, укрепват и защитават ценностите и правилата на ЕС, и по-специално правата, които гражданите притежават благодарение на правото на ЕС. Тези усилия несъмнено обхващат и разглежданите тук високорискови приложения с ИИ, предлагани на пазара и използвани в ЕС.

Както бе посочено по-горе, ИИ е невъзможен без данни. Функционирането на много системи с ИИ, както и действията и решенията, които те подпомагат, зависят в голяма степен от набора от данни, с който са били обучени системите. Следователно са необходими мерки за гарантиране, че данните, използвани за обучението на системите с ИИ, са съобразени с ценностите и правилата на ЕС, по-специално във връзка с безопасността и съществуващите законодателни правила за защита на основните права. Ето някои примерни изисквания спрямо набора от данни, използвани за обучение на системите с ИИ:

- изисквания за осигуряване на разумни гаранции, че последващото използване на продуктите или услугите, които произтичат от дейността на системата с ИИ, е безопасно, в смисъл, че е съобразено със стандартите в приложимите правила за безопасност на ЕС (както съществуващи, така и потенциални бъдещи допълнителни). Например изисквания, които гарантират, че системите с ИИ са обучени с набори от данни, които са достатъчно всеобхватни и покриват всички необходими сценарии с цел избягване на опасни ситуации;

- изисквания за предприемане на разумни мерки за гарантиране, че това последващо използване на системите на ИИ не води до резултати, които пораждат забранена дискриминация. Тези изисквания биха могли да включват по-специално задължения за използване на набори от данни, които са достатъчно представителни, най-вече за да се гарантира, че всички важни характеристики (пол, етнос и др.), чието отсъствие в данните може да породи забранена дискриминация, присъстват адекватно във въпросните набори от данни;

- изисквания, които имат за цел да се гарантира, че неприкосновеността на личния живот и личните данни са защитени адекватно при използването на продукти и услуги с ИИ. Общият регламент относно защитата на данните и Директивата относно правоприлагането уреждат въпросите, които попадат в техния обхват.

б) Съхранение на документация и данни

Предвид сложността и непрозрачността на много системи с ИИ, която може да затрудни ефективното проверяване на съответствието с приложимите правила и тяхното прилагане, наложително е да се приемат изисквания за водене на документация относно разработването на алгоритъма, данните, използвани за обучение на високорискови системи с ИИ, и в някои случаи — за съхраняването на самите данни. Тези изисквания по същество позволяват да бъдат проследявани и проверявани потенциално проблемни действия или решения на системите с ИИ. Това не само ще улесни надзора и правоприлагането, но може също така да стимулира допълнително съответните стопански субекти да вземат предвид на ранен етап необходимостта от спазване на въпросните правила.

За тази цел регулаторната рамка може да предписва следното:

- воденето на точна документация относно набора от данни, използвани за обучение и изпитване на системите с ИИ, включително описание на основните характеристики и начина на подбор на данните;

- в някои обосновани случаи, съхранение на самите набори от данни;

- съхраняване на документация относно методите, процесите и техниките на програмиране[53] и обучение, използвани за създаването, изпитването и валидирането на системите с ИИ, включително, когато това е от значение за безопасността и за избягване на необективност, която може да доведе до забранена дискриминация.

Записите, документацията и, според случая, наборите от данни трябва да се съхраняват за ограничен и разумен период от време, за да се гарантира ефективното прилагане на съответното законодателство. Необходими са мерки, които да гарантират, че същите могат да се предоставят при поискване, по-специално за изпитване или проверка от компетентните органи. Когато е необходимо, следва да бъдат взети мерки, за да се гарантира, че е защитена поверителната информация, като например търговските тайни.

в) Предоставяне на информация

Прозрачност е необходима не само при воденето на документация, споменато в буква в) по-горе. За да се постигнат поставените цели — по-специално, за да се насърчи отговорното използване на ИИ, да се изгради доверие и да се внесат корекции, където е необходимо — е важно да се предоставя по проактивен начин подходяща информация относно използването на високорискови системи с ИИ.

Следните изисквания заслужават внимание:

- предоставяне на ясна информация относно възможностите и ограниченията на системата с ИИ, по-специално нейното предназначение, условията, при които тя би трябвало да функционира според предвиденото и очакваната степен на точност в

[53] Например документация относно алгоритъма, включително за какви цели трябва да оптимизира моделът, какви са първоначално зададените тежести за определени параметри и т.н.

изпълнение на предназначението ѝ. Тази информация е важна, особено за внедряващите системите, но тя може да е значима и за компетентните органи и засегнатите страни;

- отделно, гражданите следва да са ясно осведомени, че взаимодействат със система с ИИ, а не с човек. Законодателството на ЕС за защита на данните вече съдържа някои правила в тази насока[54], но могат да се наложат допълнителни изисквания за постигането на горепосочените цели. В такъв случай следва да се избягва ненужното отежняване. За тази цел не е необходимо потребителите да бъдат информирани, че си взаимодействат със системи с ИИ, когато това е съвсем очевидно за тях. Освен това е важно предоставяната информация да бъде обективна, кратка и лесно разбираема. Начинът на предоставяне на информацията следва да бъде съобразен с непосредствения контекст.

г) Изправност и точност

Системите с ИИ — и със сигурност високорисковите приложения с ИИ — трябва да бъдат технически изправни и точни, за да бъдат надеждни. Това означава, че те трябва да се разработват отговорно и с предварително и подходящо отчитане на рисковете, които те могат да генерират. Тяхното разработване и функциониране трябва да бъде такова, че да гарантира, че системите с ИИ функционират надеждно, съгласно с предвиденото. Следва да се предприемат всички мерки в рамките на разумното, за да се сведе до минимум рискът от причиняване на вреда.

Следните аспекти заслужават внимание:

- изисквания, гарантиращи, че системите с ИИ са изправни и точни или най-малкото правилно отразяват своята степен на точност на всички етапи от жизнения си цикъл;

- изисквания, гарантиращи, че резултатите са възпроизводими;

- изисквания, гарантиращи, че системите с ИИ могат адекватно да се справят с грешки или отклонения на всички етапи от жизнения си цикъл;

- изисквания, гарантиращи, че системите с ИИ са устойчиви на явни атаки и прикрити опити за манипулиране на самите данни или алгоритми и че в такива случаи се предприемат мерки за ограничаване на риска.

д) Човешки надзор

Човешкият надзор спомага да се гарантира, че дадена система с ИИ не накърнява автономността на хората и не причинява други неблагоприятни последици. Надежден, етичен и антропоцентричен ИИ може да се постигне единствено с подобаващ човешки надзор на високорисковите приложения с ИИ.

Макар всички приложения с ИИ, разгледани в настоящата Бяла книга от гледна точка на специфичен правен режим, да се считат за високорискови, човешкият надзор може да варира по

[54] По-специално, съгласно член 13, параграф 2, буква е) от ОРЗД при получаването на личните данни администраторите трябва да предоставят на субектите на данните допълнителна информация, която е необходима за гарантиране на добросъвестно и прозрачно обработване на данните, а именно наличието на автоматизирано вземане на решения, както и определена допълнителна информация.

вид и степен в отделните случаи. Той ще зависи по-специално от предназначението на системите и от това как използването им се отразява на засегнатите граждани и правни субекти. Човешкият надзор няма да засяга и законовите права, установени с ОРЗД, когато системата с ИИ обработва лични данни. Например, той би могъл да се изразява в следното (списъкът е неизчерпателен):

- резултатите от системата с ИИ не се използват, докато не бъдат проверени и одобрени от човек (например заявление за социалноосигурителни обезщетения може да бъде отхвърлено само от човек);

- резултатите от системата с ИИ се използват непосредствено, но впоследствие се осигурява човешка намеса (например система с ИИ може да отхвърли заявление за кредитна карта, но трябва да е възможно решението да бъде разгледано от човек впоследствие);

- наблюдение на системата с ИИ в работен режим и възможност за намеса в реално време и изключване (напр. наличие на бутон или начин за спиране в автомобил без водач, когато човек установи, че превозното средство не работи безопасно);

- на етапа на проектиране, налагане на оперативни ограничения на системата с ИИ (напр. автомобил без водач спира при условия на ниска видимост, когато датчиците губят надеждност, или се движи във всички случаи на определено разстояние от предходното превозно средство).

е) Специфични изисквания за дистанционна биометрична идентификация

Събирането и използването на биометрични данни[55] с цел дистанционна идентификация[56], например чрез инсталиране на обществени места на системи за разпознаване на лица, носи особени рискове за основните права[57]. Използването на ИИ в системи за дистанционна биометрична идентификация може да засяга различно основните права в зависимост от целта, контекста и обхвата на използването.

---

[55] Биометрични данни означава „лични данни, получени в резултат на специфично техническо обработване, които са свързани с физическите, физиологичните или поведенческите характеристики на дадено физическо лице и които позволяват или потвърждават уникалната идентификация на това физическо лице, като лицеви изображения или дактилоскопични данни [пръстови отпечатъци]“. (Директива относно правоприлагането, член 3, параграф 13; ОРЗД, член 4, параграф 14; Регламент (ЕС) 2018/1725, член 3, параграф 18).

[56] При разпознаването на лица, идентифицирането се извършва чрез сравнение на образеца на портретна снимка на лицето с много други образци, съхранявани в база данни, за да се установи дали изображението на това лице присъства в тази база данни. За сметка на това проверката (на самоличността) често се разглежда като единично съпоставяне. То позволява сравняване на два биометрични образеца, за които принципно се предполага, че принадлежат на едно и също лице. Сравняват се два биометрични образеца, за да се определи дали лицето, което фигурира на двете изображения, е едно и също лице. Такава процедура се използва например на вратите за автоматизиран граничен контрол при гранични проверки на летищата.

[57] Например за човешкото достойнство. В тази връзка технологията за разпознаване на лица буди загриженост особено що се отнася до основни права като зачитането на личния живот и защитата на личните данни. Възможно е да бъде засегнато и правото на недискриминация, както и правата на отделни групи, като децата, възрастните хора и хората с увреждания. Освен това тази технология не бива да ограничава свободата на словото, свободата на сдружаване и на събрания. Виж: Технология за разпознаване на лица: съображения във връзка с основните права в контекста на правоприлагането — https://fra.europa.eu/en/publication/2019/facial-recognition.

Нормите на ЕС за защита на данните забраняват, освен при специални условия, обработването на биометрични данни единствено с цел да се разпознае дадено физическо лице[58]. По-специално съгласно ОРЗД такова обработване може да се извършва само на ограничен брой основания, като основното е по причини, свързани със значим обществен интерес. В такъв случай обработването трябва да се извършва въз основа на правото на ЕС или националното законодателство, при спазване на изискванията за пропорционалност, зачитане на същността на правото на защита на данните и с подходящи гаранции. Съгласно Директивата относно правоприлагането трябва да съществува строга необходимост от подобно обработване, то да е принципно разрешено от правото на ЕС или националното законодателство и да са налице подходящи гаранции. Тъй като всяко обработване на биометрични данни с единствена цел разпознаването на физическо лице би означавало изключение от забраната, предвидена в правото на ЕС, то ще бъде предмет на Хартата на основните права на ЕС.

От това следва, че в съответствие с действащите правила на ЕС за защита на данните и Хартата на основните права ИИ може да се използва за дистанционна биометрична идентификация само когато това е надлежно обосновано, пропорционално и е обект на подходящи гаранции.

В отговор на евентуално безпокойство сред обществеността, свързано с използването на ИИ за такива цели на обществени места, и за да се избегне фрагментирането на вътрешния пазар, Комисията ще инициира широк европейски дебат относно конкретните случаи, ако има такива, в които такова използване би било обосновано, както и относно общите гаранции.

Д. Адресати

По отношение на адресатите на правните изисквания, които биха били приложими спрямо посочените по-горе високорискови приложения с ИИ, трябва да бъдат разгледани два основни въпроса.

На първо място стои въпросът как трябва да се разпределят задълженията между участващите стопански субекти. Жизненият цикъл на една система с ИИ предполага много участници. Сред тях са разработчикът, внедрителят (лицето, което използва продукт или услуга с ИИ) и евентуално други (производител, дистрибутор или вносител, доставчик на услуги, професионален или частен ползвател).

Комисията е на мнение, че в една бъдеща регулаторна рамка всяко задължение следва да е отправено към участника или участниците, които са в състояние най-добре да отговорят на всички потенциални рискове. Например, разработчиците на системи с ИИ могат най-добре да се справят с рисковете на етапа на разработване, но техните възможности за контрол на риска са най-вероятно ограничени на етапа на използване. В този случай внедрителят е този, на когото следва да се вмени съответното задължение. Това не засяга въпроса дали — за целите на установяване на отговорността спрямо крайните ползватели или други ощетени лица, както и за гарантирането на ефективен достъп до правосъдие — някоя страна (и коя точно) следва да носи отговорност за някаква причинена вреда. Съгласно законодателството на ЕС относно отговорността за вреди от стоки, отговорността за вреди, причинени от дефектни стоки, се носи от производителя, без да се засягат националните закони, които също могат да позволяват обезщетения от други страни.

[58] Член 9 от ОРЗД, член 10 от Директивата относно прилагането. Вж. също член 10 от Регламент (ЕС) № 2018/1725 (приложим за институциите и органите на ЕС).

Второ, налице е въпросът относно географския обхват на законодателната намеса. Комисията е на мнение, че е от първостепенно значение изискванията да са приложими към всички замесени стопански субекти, предоставящи продукти или услуги с ИИ в ЕС, независимо дали са установени в ЕС или не. В противен случай посочената по-горе законодателна намеса не би могла да постигне напълно целите си.

Е. СПАЗВАНЕ И ПРАВОПРИЛАГАНЕ

За да се гарантира надеждността и сигурността на системите с ИИ при спазване на европейските ценности и правила, приложимите правни изисквания трябва да се спазват на практика и да се прилагат ефективно както от компетентните национални и европейски органи, така и от засегнатите страни. Компетентните органи следва да са в състояние да разследват отделни случаи, но също така да оценяват въздействието върху обществото.

С оглед на високия риск, който пораждат за гражданите и за обществото ни някои приложения на ИИ (вж. раздел А по-горе), на този етап Комисията смята, че ще е необходима предварителна оценка на съответствието, за да се проверява и гарантира спазването на някои от гореспоменатите задължителни изисквания спрямо високорисковите приложения (вж. раздел Г по-горе). Предварителната оценка на съответствието може да включва процедури за изпитване, инспектиране или сертифициране [59]. Възможно е също така тя да включва проверки на алгоритмите и на наборите от данни, използвани на етапа на разработване.

Оценките на съответствието по отношение на високорисковите приложения с ИИ следва да бъдат част от вече съществуващите механизми за оценка на съответствието за голям брой продукти, които се пускат на вътрешния пазар на ЕС. Когато съществуващите механизми не са надеждни, може да се наложи създаването на подобни механизми въз основа на най-добрите практики и евентуалните отзиви на заинтересованите страни и европейските организации по стандартизация. Всеки подобен нов механизъм следва да бъде пропорционален и недискриминационен и да използва прозрачни и обективни критерии в съответствие с международните задължения.

При разработването и прилагането на система, основаваща се на предварителни оценки на съответствието, следва да се обърне особено внимание на следното:

- не всички изисквания, посочени по-горе, се поддават на проверка чрез предварителна оценка на съответствието. Например изискването за предоставяне на информация по принцип не се поддава на подобна проверка чрез такава оценка;

- специално внимание заслужава възможността някои системи с ИИ да се развиват и обучават сами от опит, което може да наложи многократни оценки през целия жизнен цикъл на въпросните системи с ИИ;

- необходимостта да се проверяват данните, използвани за обучение, и съответните методологии, процеси и техники за програмиране и обучение, които се използват за създаване, изпитване и валидиране на системи с ИИ;

[59] Системата ще се основава на процедури за оценяване на съответствието в ЕС, вж. Решение № 768/2008/ЕО или Регламент (ЕС) 2019/881 (Акт за киберсигурността), като се вземат предвид особеностите на ИИ. Виж „Синьо ръководство“ за прилагането на правилата относно продуктите на ЕС, 2014 г.

- в случай че оценката на съответствието покаже, че дадена система с ИИ не отговаря на изискванията, например по отношение на данните, използвани за нейното обучение, установените недостатъци трябва да се отстранят, например чрез повторно обучение на системата в ЕС по начин, който гарантира изпълнението на всички приложими изисквания.

Оценките на съответствието ще бъдат задължителни за всички стопански субекти, за които се отнасят изискванията, независимо от мястото им на установяване[60]. За да се облекчат МСП, може да се предвиди някаква подкрепяща структура, включително чрез центровете за цифрови иновации. Освен това, спазването на изискванията може да се улесни чрез стандарти и специални онлайн инструменти.

Всякаква предварителна оценка на съответствието следва да не засяга наблюдението на съответствието и последващото правоприлагане от страна на компетентните национални органи. Това важи за високорисковите приложения с ИИ, но и за други приложения с ИИ, които са обект на правни изисквания, макар да е вероятно по-високият риск при първите да наложи компетентните национални органи да им обръщат по-голямо внимание. Последващите проверки следва да бъдат улеснявани от подходяща документация за съответното приложение с ИИ (вж. раздел Д по-горе) и, когато е целесъобразно, следва да се осигури възможност трети страни, например компетентни органи, да изпитват тези приложения. Това може да е особено важно при възникване на рискове за основните права, зависещи от контекста. Подобно наблюдение на съответствието следва да бъде част от схема за непрекъснат надзор на пазара. Аспектите, свързани с управлението, са разгледани по-подробно в раздел З по-долу.

Освен това както за високорисковите приложения с ИИ, така и за другите приложения с ИИ, следва да се осигури ефективна съдебна защита за страните, които са ощетени от системи с ИИ. Въпросите, свързани с отговорността, са допълнително обсъдени в доклада относно безопасността и отговорностите, който придружава настоящата Бяла книга.

Ж. ДОБРОВОЛЕН ЗНАК ЗА ПРИЛОЖЕНИЯ БЕЗ ВИСОКОРИСКОВ ИИ

За приложенията с изкуствен интелект, които не се определят като „високорискови“ (вж. раздел В по-горе) и които следователно не подлежат на разгледаните по-горе задължителни изисквания (вж. раздели Г, Д и Е по-горе), един от вариантите е към приложимото законодателство да се добави схема за присъждане на доброволен знак.

В рамките на тази схема заинтересованите стопански субекти, които не са обхванати от задължителните изисквания, могат да решат да се обвържат доброволно с тях или със специфичен набор от подобни изисквания, изготвени специално за целите на доброволната схема. В такъв случай съответните стопански субекти ще получават знак за качество за своите приложения с ИИ.

Доброволният знак ще позволява на съответните стопански субекти да означават своите продукти и услуги с ИИ като надеждни. Така потребителите ще могат лесно да научават, че въпросните продукти и услуги отговарят на определени обективни и стандартизирани за целия ЕС референтни показатели, които надскачат рамките на обичайните правни задължения. Това ще заздрави доверието на потребителите в системите с ИИ и ще спомогне за възприемането на технологията като цяло.

[60] За съответната структура на управление, включително органите, определени да извършват оценките на съответствието, вж. раздел З по-долу.

Този вариант би довел до изготвянето на нов правен инструмент, който да определи рамката за присъждане на доброволен знак, за създателите и/или внедрителите на системи с ИИ, които не се считат за високорискови. Участието в схемата ще бъде доброволно, но от момента, в който разработчиците или внедрителите изберат да използват знака, те се задължават да изпълняват изискванията. Комбинацията от *ex ante* и *ex post* правоприлагане трябва да гарантира спазването на всички изисквания.

3. УПРАВЛЕНИЕ

Необходима е европейска структура на управление в областта на ИИ под формата на рамка за сътрудничество между националните компетентни органи, за да се избегне разпокъсването на отговорностите, да се увеличи капацитетът в държавите членки и да се гарантира, че Европа постепенно натрупва капацитета, който ѝ е необходим за изпитване и сертифициране на продукти и услуги с ИИ. В този контекст би било от полза да се подкрепят компетентните национални органи, за да могат те да изпълняват мандата си в случаите, в които се използва ИИ.

Една европейска структура на управление може да има различни задачи в ролята си на форум за редовен обмен на информация и най-добри практики, където се разпознават възникващите тенденции, предоставят се съвети относно дейностите по стандартизация и относно сертифицирането. Тя следва също така да играе ключова роля за улесняване прилагането на правната рамка, например чрез предоставяне на насоки, становища и експертен опит. За тази цел тя следва да разчита на мрежа от национални органи, както и на секторни мрежи и регулаторни органи на национално равнище и на равнище ЕС. Освен това Комисията може да бъде подпомогната от експертен комитет.

Управленската структура следва да гарантира максимално участие на заинтересованите страни. Заинтересованите страни — организации на потребителите и социални партньори, предприятия, изследователи и организации на гражданското общество — следва да бъдат консултирани относно прилагането и по-нататъшното развитие на рамката.

Предвид вече съществуващите структури, например в областта на финансите, фармацевтичните продукти, въздухоплаването, медицинските изделия, защитата на потребителите и защитата на данните, предложената управленска структура не трябва да дублира вече съществуващи функции. Вместо това тя следва да установи тесни връзки с други европейски и национални компетентни органи в различните сектори, за да допълни наличния експертен опит и да подпомогне съществуващите органи в наблюдението и надзора на дейностите на стопанските субекти, включващи системи с ИИ и продукти и услуги с ИИ.

Най-сетне, ако бъде избрана тази възможност, извършването на оценки на съответствието може да бъде възложено на осведомени органи, определени от държавите членки. Центровете за изпитване следва да позволяват независим одит и оценка на системите с ИИ в съответствие с изискванията, посочени по-горе. Независимото оценяване ще заздрави доверието и ще гарантира обективност. То би могло също така да улесни работата на съответните компетентни органи.

ЕС разполага с отлични центрове за изпитване и оценка и следва да развие капацитета си и в областта на ИИ. Стопанските субекти, установени в трети държави, които искат да навлязат на вътрешния пазар, биха могли да се обърнат към упълномощените органи, установени в ЕС или, при спазване на споразуменията за взаимно признаване с трети държави, да се обърнат към органите на трета държава, упълномощени да извършват подобна оценка.

Управленската структура, свързана с ИИ, и разглежданите тук възможни оценки на съответствието няма да се отразят на правомощията и отговорностите по действащото законодателство на ЕС на съответните компетентни органи в конкретни сектори или по конкретни въпроси (финанси, лекарствени средства, въздухоплаване, медицински изделия, защита на потребителите, защита на данните и т.н.).

6. ЗАКЛЮЧЕНИЕ

ИИ е стратегическа технология, която е многообещаваща за гражданите, дружествата и обществото като цяло, при условие че бъде антропоцентрична, етична, устойчива и зачитаща основните права и ценности. ИИ предлага съществени подобрения на ефективността и производителността, които могат да повишат конкурентоспособността на европейската промишленост и да подобрят благосъстоянието на гражданите. Тя може също така да допринесе за намиране на решения на някои от най-неотложните предизвикателства пред обществото, сред които са борбата с изменението на климата и влошаването на околната среда, както и на предизвикателствата, свързани с устойчивото развитие, демографските промени, защитата на нашите демокрации и, когато е необходимо и съобразно, борбата с престъпността.

За да може Европа да се възползва пълноценно от възможностите, които предлага ИИ, тя трябва да развие и заздрави необходимия промишлен и технологичен капацитет. Както е посочено в придружаващата европейска стратегия за данните, това изисква също мерки, които ще позволят на ЕС да се превърне в център за данни със световна величина.

Европейският подход към ИИ има за цел да подсили иновационния капацитет на Европа в областта на ИИ и едновременно с това да подкрепи разработването и внедряването на етичен и надежден ИИ в икономиката на ЕС. ИИ следва да работи за хората и да бъде движеща сила за благото на обществото.

С настоящата Бяла книга и придружаващия я доклад относно рамката за безопасност и отговорност Комисията започва широки консултации с представители на гражданското общество, промишлеността и академичните среди на държавите членки по конкретни предложения за европейски подход към ИИ. Ще бъдат разисквани както политическите средства за насърчаване на инвестициите в научни изследвания и иновации, за подобряване на уменията и в подкрепа за внедряването на ИИ от МСП, така и предложенията за основните елементи на една бъдеща регулаторна рамка. Консултацията ще позволи да се проведе всеобхватен диалог с всички заинтересовани страни и резултатите от него ще се използват в следващите стъпки на Комисията.

Комисията очаква отзиви по предложенията, изложени в Бялата книга, чрез открита обществена консултация на адрес: https://ec.europa.eu/info/consultations_en. Тя ще продължи до 19 май 2020 г.

За Комисията е стандартна практика да публикува отзивите, получавани при обществена консултация. Въпреки това може да бъде поискано тези предложения или част от тях да останат поверителни. Ако желаете те да не бъдат публикувани, моля посочете ясно това на първата страница, като изпратите на Комисията и неповерителна версия за публикуване.