

Изх. № 139/17.09.2020 г.

Становище от Фондация „ЛИБРе“
по
Концепция за развитието на изкуствения интелект в България
до 2030 г.

Уважаеми дами и господа,

В качеството ни на организация с нестопанска цел, работеща за развитието на държавните политики в дигиталното общество и внедряването на технологиите във всяка професионална област и в личния живот, в т.ч. нови и нововъзникващи приложения на изкуствения интелект (ИИ), представяме нашата позиция относно публикуванията на 19 август 2020 г. проект на Концепция за развитието на изкуствения интелект в България до 2030 г.¹ (Концепцията), както следва:

1. Статут на представения проект

1.1.В Концепцията е посочено, че същата се основава на разработените от екип на Българската академия на науките (БАН) и външни експерти „Рамка за Национална стратегия за развитието на изкуствения интелект в България“ и последващата я „Стратегия за развитието на изкуствения интелект в България до 2030 г. (предварителна визия)“. От публично достъпната информация не става ясно дали и как цитираните документи са разработени в съответствие с Методологията за стратегическо планиране на Република България²; как и защо национална стратегия от голяма обществена значимост се разработва от научна организация, а не

¹ <http://strategy.bg/PublicConsultations/View.aspx?@lang=bg-BG&Id=5396>

² <http://www.strategy.bg/Publications/View.aspx?lang=bg-BG&Id=90>

от ресорното министерство³; както и дали изготвените документи са създадени с участието на всички заинтересовани страни⁴ и дали са преминали обществено обсъждане⁵. Предвид това, че последващ документ от по-нисък ранг е основан на тези разработки, същите следва да са демократично легитимни по своята същност и създадени в съответствие с приетите национални правила.

- 1.2. Предвид това, че концепцията като документ не е призната за стратегически документ съгласно Методологията за стратегическо планиране на Република България, от предложението проект не става ясна необходимостта от изработването на Концепция за развитието на изкуствения интелект в България до 2030 г., предвид явната необходимост от национална стратегия.
- 1.3. В заглавието на Раздел 6 от Концепцията некоректно се използва думата „стратегия“, предвид изборият от авторите на документа подход за „концепция“. В случай, че под „Стратегията ИИ (2021-2023)“ се има предвид друг документ, това следва да бъде изрично уточнено.

2. Общи положения

- 2.1. В Концепцията липсва анализ на състоянието на сектора в България. Изготвен е преглед на стратегически документи на национално и европейско ниво, представено е общото ниво на цифровизация и цифрови умения, както и общо представяне на ИТ сектора в България (стр. 16-20). Макар това да са предпоставки за развитие на ИИ в България, предложената информация не представя реалното състояние на сектора, развиващ ИИ в България, респективно не може да бъде оценено как предложените в Концепцията мерки

³ Ресорните министерства са отговорни в рамките на предоставените им компетенции за областите на политиките. Разработването и изпълнението на политиките заедно с подчинените институции са изначално задължение на ресорните министерства. Ресорните министерства са отговорни за изпълнение на Програмата на правителството, за разработването на национални стратегии, разработването на стратегически планове (бюджетни програми), които по най-ефективен начин изпълняват приоритетите на Правителството.

Органът на изпълнителната власт отговаря за организацията на дейностите по разработването на стратегията. Обикновено за разработване и актуализация на стратегии се създава експертна работна група с членове от компетентните институции. В работните групи могат да се привличат външни за администрацията експерти.

⁴ Реалистичните стратегии могат да бъдат разработени само в партньорство с всички заинтересовани страни (университети, НПО, бизнес организации и граждани). Заинтересованите страни трябва да бъдат включени в процеса на разработването на стратегии. По всяко едно време гражданите, социалните и бизнес партньори трябва да се запознават с готвените промени в регулираните от държавата области. Те могат да заявят своята позиция и да дават предложения за промяна.

⁵ Видно от преглед на Портала за обществени консултации проекти на „Рамка за Национална стратегия за развитието на изкуствения интелект в България“ и последващата я „Стратегия за развитието на изкуствения интелект в България до 2030 г. (предварителна визия)“ не са били публикувани.

- биха намерили реално отражение в конкретни индикатори за успех.
- 2.2. Концепцията коректно посочва, че ключов сектор, развиващ ИИ, е секторът на научните изследвания. В същото време ИКТ секторът и сектор роботика са negliжирани непропорционално спрямо тяхното пазарно и/или индустриално влияние. Предвид развитието на ИИ през последните 10 години и възможността за технологично обезпечаване с изчислителна мощ на доскоро теоретични постановки, е погрешно да се счита, че секторът на научните изследвания е единственият, който има пряко влияние върху навлизането на ИИ в бизнеса и обществото, респективно – върху създаването на подходяща рамка за развитие и създаване на надежден ИИ.
- 2.3. В приложение са изброени проекти по ОП НОИР (2014-2020), свързани с инфраструктури и изследвания по ИИ. В същото време липсва анализ на технологични разработки, инициативи и/или действащи секторни проекти, които биха имали ефект и/или биха били движещ фактор за приложението на ИИ в съответния сектор, развиващ ИИ, сектор-консуматор или сектор, създаващ условия за развитието и внедряването на ИИ.
- 2.4. В Концепцията липсва цялостна визия за развитието на сектора на ИИ в България. Посочените цели/под-цели са твърде общи и не почиват на конкретни индикатори за успех. Предложените области на въздействие, специфични мерки и приоритетни сектори са до голяма степен пожелателни инициативи, като за съществена част от тях липсва конкретика на интервенциите, както и каква е тяхната добавена стойност в процеса по развитие и използване на системи с ИИ на национално ниво.
- 2.5. Така дефинираните цели/под-цели не дават ясно разграничение между краткосрочни (1-3 години), средносрочни (3-9 години) и дългосрочни (10+ години) цели на България в областта на ИИ. Концепцията предлага обща рамка на първия етап на изпълнение (2021-2023 г.), но не съдържа препратки и/или предложения за втори и следващи етапи (в какви времеви рамки ще се развиват, какви мерки ще включват, кои сектори биха попаднали в тях и т.н.).
- 2.6. Не става ясно и как предложеният обхват от цели/под-цели и първият етап на изпълнение на Концепцията са обвързани с текущите и предвидени етапи на развитие на сектора на европейско ниво (относно етична рамка, нормативна уредба и т.н.). Пример може да се даде с изведените специфични мерки за „[с]ъздаване на нормативна база за развитие и използване на надежден ИИ в съответствие с международните регулаторни и етични стандарти“ – единно признати международните регулаторни и етични стандарти към момента липсват.

2.7. Концепцията създава впечатление, че мерките за развитие на сектор ИИ ще се осъществяват приоритетно от академичната общност, научните среди и публичния сектор. Настоящото практико-приложно поле на ИИ съществено надхвърля ограниченията на научния или публичния домейн и неговото развитие следва да се търси в мултидисциплинарния подход и крос-секторното сътрудничество, включително в партньорство с всички заинтересовани страни (неправителствен сектор, бизнес организации, медии, граждани и т.н.).

2.8. Концепцията не предоставя мотиви как и въз основа на каква информация са подбрани описаните в Раздел 6 от Концепцията приоритетни сектори за първия етап на изпълнение (2021-2023 г.). Нещо повече, необходимо е пълноценно взаимодействие, съгласуваност и координация на всички институционални и организационни нива, както в процеса на разработване, така и в процеса на изпълнение на стратегическите документи. В този смисъл, Концепцията не предоставя и информация как отговорните институции по приоритетни сектори разпознават мерките и областите на интервенция и какви са конкретните инициативи и секторни приложения, които ще бъдат реализирани за периода 2021-2023 г.

2.9. Концепцията изброява редица сектори-консуматори на ИИ.

Недостатъчно е застъпено приложението на ИИ в производството (т.нар. индустриален ИИ), както и връзката на ИИ с цифровата трансформация в промишлеността (Индустрия 4.0). Във връзка с това вж. т. 6.3 по-долу.

Липсва фокус върху приложението на ИИ в търговията на дребно, като ключово разклонение при електронната търговия, и използваните от търговците способности за влияние върху крайните потребители – от съществено значение за тяхното поведение онлайн и създаването на конкретни социално-икономически нагласи.

Съществен пропуск на Концепцията е липсата на фокус върху сектор „Правосъдие“, като сектор-консуматор на ИИ с изключително значение за правата и свободите на гражданите, при който са налице редица правни, етични и социални предизвикателства пред въвеждането на технологията.

Развлекателната индустрия, услугите в областта на маркетинга, рекламата и обслужването на клиенти, комуналните услуги, образованието и др. също не са застъпени от Концепцията, което отчитаме като грешка предвид възможностите за манипулация на поведението и/или отказ на достъп до услуги и/или други ограничения в социален/социално-икономически аспект, до които би могло да доведе приложението на системите с ИИ.

3. Неточности в понятийния апарат и/или при препратки към документи, изготвени от трети страни

- 3.1. Концепцията се позовава на дефиницията за интелигентни системи, предложена от Експертната група на високо равнище по изкуствен интелект и цитирана в доклада „Научно, научноизследователство и иновационно представяне на ЕС 2020: справедлива, зелена и цифрова Европа“ (стр. 4). Обръщаме внимание, че дефиницията не е нито формално приета, нито изрично одобрена от ЕК и понастоящем е широко критикувана от академичните среди и гражданския сектор.
- 3.2. Проектът на Концепция работи с понятието „потребителско доверие“ (стр. 5) и създава впечатлението, че засяга потребителските нагласи и очаквания. Такова разбиране, обаче, не е в унисон с виждането на Европейската комисия или на Експертната група на високо равнище по изкуствен интелект, а именно – че се цели изграждане на доверие у гражданите на ЕС, а не просто на потребителите.
- 3.3. В документа има и неточности, които рискуват създаването на объркване у обществеността. Така напр. на стр. 5 се твърди, че „[п]рез 2019 г. ЕК публикува седем ключови изисквания“, а именно изискванията, изброени в Насоките относно етичните аспекти за надежден изкуствен интелект. Насоките са продукт на работата на Експертната група, но те (все още) не са нито формално приети, нито изрично одобрени от страна на ЕК.⁶
- 3.4. Твърди се, че абстрактни „етични норми“ ще създадат предпоставки за развитие на научните изследвания (стр. 5). Тези норми не са категоризирани, нито разяснени където и да било по-нататък в проекта.
- 3.5. Проектът на Концепция твърди, че „данните са важна суровина [...] за развитието на нови алгоритми“, но не посочва ясно как точно данните служат за развитието на алгоритми (стр. 6).
- 3.6. Проектът на Концепция използва донякъде произволно понятия като „анонимизирани лични данни“, „фирмени данни“ и „публични данни“, като ги поставя в една и съща категория. Понятието „фирмени данни“ не е с ясно дефиниран обхват и би създавало

⁶ B Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions “Building Trust in Human-Centric Artificial Intelligence”, COM(2019) 168 final 4 (2019), достъпно на адрес <https://ec.europa.eu/transparency/readdoc/rep/1/2019/EN/COM-2019-168-F1-EN-MAIN-PART-1.PDF> (последно посетен на 16.09.2020), се отбелязва, че Насоките са ценни в процеса на нормотворчество, но не са формално одобрени като източник.

съществени затруднения при използването му в подобни документи от стратегическо значение.

4. Необходимост от мултидисциплинарност в образователните инициативи в областта на ИИ

4.1. В проекта за Концепция се отбелязва, че „висшето образование, което създава профилираните специалисти в областта“ е първият основен фактор за създаване на условия за развитие и внедряване на ИИ. Същевременно се коментира, че „в средното образование се изграждат базова компютърна грамотност и основни познания по ИКТ, в частност и използване на системи с ИИ“ (стр. 14). Концепцията обаче не акцентира върху необходимостта от развитие на образователни инициативи, които надхвърлят създаването на чисто технически научен потенциал.

Не е обърнато внимание на новосъздадени локални инициативи в България, като например дисциплината „Отговорен изкуствен интелект: етични, правни, социални и икономически аспекти“, създадена към Факултета по математика и информатика към Софийски университет „Св. Климент Охридски“, която би могла да служи като добра база за подход към създаването на умения за анализ на социално-етични и правни проблеми сред специалистите по ИИ. Нещо повече, бидейки мултидисциплинарна област, необходимо е ИИ да „надскочи“ чисто техническите си аспекти и образователните инициативи да бъдат насочени и към специалисти от други области – психология и когнитивна наука, право, философия и социална наука. В проекта не става ясно какъв би бил „българският“ подход към насърчаването на подобни инициативи, нито как той се вписва в общия европейски подход.

Наред със сътрудничеството с бизнеса е необходим и конкретен подход за идентифициране на най-атраktivните и социално полезни области на интердисциплинарно сътрудничество, в чиито рамки могат да се реализират интердисциплинарни научни проекти, като напр. етика и право, компютърни науки и право и др.

4.2. Проектът на Концепция поставя като мярка в средното образование „[у]величаване на компетентностите на учениците в областта на етичните проблеми, свързани с използването на информационните технологии и техните права в условията на цифровия свят, в който живеят“ (стр. 25). Същевременно се отбелязва, че и във висшето образование е необходимо „[с]пециално внимание върху изучаване на въздействието на ИИ върху обществото, както и върху стандартите за изграждане на надежден ИИ. Това предполага, от една страна, включване в университетските образователни програми по технически науки на учебни дисциплини, насочени към правните, етичните и социалните аспекти на ИИ, а от друга

страна, включване на дисциплини за изследване на въздействието на ИИ във факултетите по социални, правни и хуманитарни науки“ (стр. 26). Натрупаният преподавателски опит на екипа на фондация „ЛИБРе“ в споменатата по-горе дисциплина „Отговорен изкуствен интелект: етични, правни, социални и икономически аспекти“ към магистърската програма по изкуствен интелект във Факултета по математика и информатика на СУ „Св. Климент Охридски“ показва ясна необходимост от подобряване на обучението по т.нар. „меки умения“ и на етапа на висшето образование. Нещо повече, понастоящем или напълно липсват, или са практически трудни за изпълнение конкретни указания за социално отговорна, правна и етична развойна и научноизследователска дейност в областта на ИИ. Острата нужда от засилване на този аспект е очевиден и от главната цел, която преследва общоевропейският подход, основан на „надеждност“ на системите с ИИ. Бъдещите специалисти имат нужда не просто от знания за тези проблеми, но и от реални умения за участие в интердисциплинарен диалог със специалисти от други области. Това следва да бъде постигнато не само чрез насърчаване на интер- и интрадисциплинарността във висшето образование, но и чрез въвеждането на задължителни учебни курсове, които да подготвят бъдещите специалисти за справяне не само с техническите, но и с още по-трудните социални, етични и правни въпроси на ИИ.

5. Етични насоки за развитие и използване на надежден ИИ

Концепцията неколкократно прави препратки към необходимостта от конкретни етични норми, етичните принципи и свързаните с тях ценности, които трябва да се зачитат при разработването, внедряването и използването на системи с ИИ, развитие и използване на ИИ да се извършва съгласно с установените нормативни и етични принципи и сходни. В същото време, така цитираните нужди остават абстрактни и не носят добавена стойност за разработчиците, дистрибуторите, интеграторите и ползвателите на системи с ИИ.

Фокус на Концепцията би следвало да се търси относно:

- институционален ангажимент в разработването списък(-ци) за оценяване на надеждния ИИ на наднационално ниво и/или адаптирането на такива към условията за развитие на сектора в България (с оглед ниво на текущо развитие спрямо желано състояние, социално-икономически фактори и др.);
- разработване на секторни етични кодекси за поведение – за сектори, развиващи ИИ, или сектори-консуматори, при които използването на системи с ИИ се характеризира с висок риск за правата и свободите на гражданите и/или специфични

правни, етични и социални предизвикателства пред въвеждането на технологията;

- насърчаване създаване на етични (секторно- или организационно-ориентирани) комисии при провеждането на изследвания или внедряването на конкретни решения, основани на ИИ;
- разработване на етични насоки за конкретни бизнес направления, бизнес модели, основани на ИИ, и/или крос-индустриални приложения като ключово допълнение към стратегически реформи, законодателни изменения и/или конкретни държавни политики.

6. Нормативно обезпечаване развитието и използването на надежден ИИ

6.1. Обхват на правната рамка

Законодателната дейност правилно е разпозната като водещ фактор за създаване на условия за развитие и внедряване на ИИ. За съжаление документът показва фундаментално неразбиране на основната концепция, която стои зад едно от трите ключови изисквания, изведени от работата на Експертната група, а именно – създаването на ИИ, който е съобразен с действащата правна рамка в своята цялост.

Проектът на Концепция лаконично и необяснимо ограничава обхвата на правната рамка, относима към ИИ, до „гарантира[не] запазване на основните права на потребителя и защитата на личните данни, както и осигуряване на безопасността на продуктите и определяне на отговорността“ (стр. 15). Подобно ограничаване до конкретни правни отрасли и специфични режими (напр. защита на личните данни или безопасност на продуктите) създава грешното впечатление, че обезпечаването на правната сигурност при създаване и внедряване на ИИ може да бъде редуцирана до няколко области на правно-техническо законотворчество и последваща регулация. Напротив, европейският подход започва от сърцевината на режима на правото на ЕС, а именно – основните права на гражданите.

Споменаването на основните права се ограничава до ноторното твърдение, че „националните правила в сферата на основните права и защитата на личните данни са свързани с транспонирането на релевантните европейски правни разпоредби.“ (стр. 15). В документа отсъстват каквито и да било препратки към Хартата на основните права на ЕС, Европейската конвенция за правата на човека или дори конституционно гарантираните права на гражданите като първи и основен стълб за

закрила. Вместо това са подбрани правни режими с подчертано технически характер, които са механично и безкритично представени в светлината, в която са разгледани и в Бялата книга за ИИ, а именно – като несъздаващи особени проблеми. Липсват амбиции или насоки за конкретна холистична законодателна програма, която да обхване целия спектър от юридически проблеми, които възникват както йерархично, така и хоризонтално в правната система.

6.2. Актуализиране на действащо законодателство

Заявената цел на проекта на Концепция е „да фокусира усилията по създаване на научен и експертен капацитет, разработка и внедряване на системи с ИИ като подпомогне координацията на мерки и дейности, поставени като приоритет в релевантни Национални стратегически документи“ в шест основни направления (стр. 21). Като една от под-целите е посочено „актуализиране на съществуващата регулаторна рамка относно разработването и приложението на надеждни ИИ технологии, съобразено с установените правни и етични принципи в рамките на ЕС“. За съжаление документът не посочва конкретни мерки в тази посока, като оставя съществена празнота, която следва да бъде запълнена именно от такъв документ. Липсва дори базово посочване на конкретни юридически проблеми, които подобно „актуализиране“ на „съществуващата регулаторна рамка“ би следвало да разреши. Със съжаление и тук може да се отбележи терминологичната непоследователност на документа. Така напр. не изглежда в проекта да се провежда разлика между право и регулация, а такава съществува и има конкретни и различни последици, а освен това е насочена и към различни субекти.

6.3. Създаване на нормативна уредба за развитие и използване на надежден ИИ

Друга конкретна мярка на проекта е създаването на нормативна база за развитие и използване на надежден ИИ в съответствие с международните регулаторни и етични стандарти. Отново тук се забелязва смесване не само на различни дисциплини, но и на инструменти, които решават различни проблеми. В тази критично важна част от документа яснотата и конкретиката са повече от необходими, те са задължителни. Положително е отчитането на факта, че „[п]ринципите за съблюдаване на основните права, недискриминация и защита на личните данни следва да се разглеждат като неизменна част от изискванията, които гарантират безопасността на ИИ технологиите“ (стр. 29). Липсват обаче конкретни мерки или набелязване на проблеми, които ИИ поражда. В съответствие с Бялата книга са отбелязани като ключови прилагането на съществуващото законодателство и неговите

ограничения, промяната във функционалността на системите с ИИ, разпределението на юридическите задължения и договорната и/или деликтна отговорност за причинени вреди, както и промените в концепцията за безопасност (стр. 30). Следването на подхода за разделяне на приложенията с ИИ на такива с висок и такива с нисък риск е разбираемо предвид формулировките, заложи в Бялата книга. Авторите на проекта на Концепция обаче не следва да забравят, че този политически документ все още подлежи на обсъждане и заложените в него положения не са окончателни.

Оригинален подход към въпроса би включвал не само законодателни мерки, но и конкретни насоки за бизнеса и други участници в развойния процес и използването на ИИ. Такива насоки могат да бъдат основани на Ръководните принципи на ООН за бизнеса и човешките права, които предоставят солидна транснационална рамка за управление на рисковете за човешките права в глобални вериги на доставки и на стойност. Разработването на такива насоки конкретно за бизнес модели, основани на ИИ, е ключово допълнение към каквито и да било законодателни изменения (вкл. предстоящия общ хоризонтален европейски регламент за налагане на задължения на бизнеса за надлежна проверка (due diligence) на веригата на доставки).

6.4. Експериментални регулаторни практики

Сред специфичните мерки, заложи в Концепцията, е посочено изграждането на надеждна инфраструктура за развитие на ИИ. В конкретните мерки е посочено, че „наборът от инструменти следва да включва комбинация от процедури и гъвкави подходи, които насърчават информираната преценка в процеса на вземане на решение.“ (стр. 30). За съжаление водещи регулаторни инструменти, които да подпомогнат това изграждане, напълно отсъстват от концепцията. Така напр. няма дори загатване за необходимостта от и предимствата на т.нар. регулаторни инкубатори (т.нар. „пясъчници“) и/или други експериментални регулаторни практики, които могат да подпомогнат диалога между регулаторите, компетентни да регулират различни аспекти и приложения на ИИ, и разработчиците, дистрибуторите, интеграторите и ползвателите на тези системи. Регулаторните инкубатори, възникнали първоначално във финтех индустрията, дават възможност на (финансовите) институции и участници да експериментират с иновативни (финансови) продукти или услуги в „реална“ среда, но в рамките на предварително дефинирано пространство и продължителност. В зависимост от експеримента, регулаторът предоставя подходящата регулаторна подкрепа, като облекчава специфични законови и регулаторни изисквания, с които иначе субектът трябва да се съобразява. Инкубаторът включва подходящи мерки, които да поддържат общата стабилност на

(финансовата) система. След успешен експеримент и при излизане от инкубатора, субектът трябва изцяло да спазва съответните законови и регулаторни изисквания.

В нормативното регулиране на ИИ Концепцията би следвало да акцентира и на прехода от патерналистичните институции, които извършват дейности за и спрямо хората, към институции, които проактивно създават и улесняват възникването на нови модели и разрешения. Пример за това е т.нар. изпреварваща регулация (*anticipatory regulation*), която описва инструментите, които държавите могат да използват, за да работят с нови и възникващи технологии, напр. дроне и изкуствен интелект. Изпреварваща регулация е нововъзникващ подход за регулиране от страна на правителството. Тя представлява проактивен и итеративен метод, който използва експерименти, регулаторни инкубатори и данни за по-добро напътстване на развиващите се технологии.

6.5. Секторни особености

В документа не е отчетен като проблем от интердисциплинарен правно-етичен характер въпросът за споделянето на данни от сферата на здравеопазването. Докато Общият регламент относно защита на данните предписва забранителен режим с конкретни изключения, етичното измерение на този въпрос изисква отчитането не само на индивидуалните интереси на субектите на данни (което е основен постулат на регламента) и колективните интереси (като напр. откриването на ефективно лечение за ендемично заболяване, което изисква обработването на големи обеми от потенциално лични данни), които закриля общественото здравеопазване.

6.6. Национални мерки

В Раздел 5.6 от проекта на Концепция е посочено разработването на комплекс от широкоспектърни мерки за обезпечаване на разработването на надежден ИИ, в т.ч. „[с]здаване на национална рамка за оценка на рисковете, свързани с развитието на ИИ технологиите“, „[с]здаване на специализиран механизъм за мониторинг и оценка на въздействията на ИИ технологиите“ и „[с]здаване на набор от инструменти за стимулиране на възприемането на принципите за безопасност и юридическа отговорност сред участниците в разработването, внедряването и използването на ИИ технологии“. Макар и коректни, посочените мерки имат наднационален характер – не само предвид приложимостта на технологията и произхода на съществена част от успешните ИИ решения, но и с оглед общата политика на ЕК в тази посока. Погрешно е да се смята, че България трябва и би могла да разработи посочените механизми и инструменти единствено и само на национално ниво. Описаните мерки следва да бъдат

съобразени с текущите и предвидени бъдещи етапи на развитие на сектора на европейско ниво (относно етична рамка, нормативна уредба и т.н.), като Концепцията не дава конкретика как и на какъв етап от изпълнението на Концепцията тези мерки ще бъдат реализирани.

7. Повишаване на обществената ангажираност относно ролята на ИИ

Проектът следва да бъде по-конкретен и амбициозен и по отношение на „създаване на възможности за повишаване на обществената ангажираност относно ролята на ИИ в социалния живот“ (стр. 30). По-конкретно могат да бъдат посочени мерки като анализ на данни, колективно създаване на знания, технологии за демократично приобщаване, разработване на прототипи, експериментирание, целеви изпитателни лаборатории (testbeds) за конкретни приложения, иновационни лаборатории и др.

8. Социални възприятия за рисковете от ИИ

Считаме също, че съществен пропуск на този и много други стратегически документи, насочени към ИИ, е мястото и представянето на ИИ в медийния дискурс. Начинът, по който ИИ се изобразява в политическия и правния дебат, и особено в медиите, играе ключова роля за формирането на общественото мнение. Социалните възприятия за риска от приложенията на ИИ и ролята на средствата за масова информация в процеса на общественото „конструиране“ на риска биха могли да окажат значително влияние и върху нормотворческия, а впоследствие - и правоприложен процес. Във времена на несигурност „масмедииите намират почва да влияят на социалното изграждане на реалността чрез посяване на съмнение и скептицизъм относно един или друг подход към управление на риска“.⁷ Това, разбира се, е валидно не само за разпределението на отговорността в случай на причинени вреди, но също така и за другите субекти, участващи в процеса на разработване и използване на ИИ. Добър пример в тази посока е авиацията, като „индустрия, която е много осъзната по отношение на имиджа си и изисква интензивно управление на риска, както и (...) исторически свързана с високи нива на грижа за потребителите“.⁸ Въвеждането на технологии с ИИ, способни да подкопаят доверието, прави още по-важно участието в процеси за изграждане на доверие по въпроси, свързани с „оперативната

⁷George Leloudas, *Risk and Liability in Air Law* 56-67 (2013).

⁸ibid, 239.

надеждност, прозрачността, потребителските услуги, връзките с медиите и опазването на околната среда“.⁹ Следователно е от съществено значение не само за регулаторите, но и за цялата общност от заинтересовани страни да разгледат социалните възприятия за рисковете от ИИ в по-широк и цялостен план за управление на социалния риск. Такъв план трябва да бъде продукт на координация и взаимно разбиране на всички заинтересовани страни в „екосистемата на доверие“.

Оставаме на разположение в случай на необходимост от разяснения по представеното становище.

С уважение,
Александра Цветкова
Управител на Фондация „ЛИБРе“

⁹ibid.