



Съвет на
Европейския съюз

Брюксел, 9 април 2019 г.
(OR. en)

8396/19

JAI 386	TELECOM 169
CYBER 123	AUDIO 62
DATAPROTECT 119	DEVGEN 79
FREMP 52	COMPET 323
CSC 131	ECOFIN 381
CSCI 60	EMPL 217
DIGIT 76	CONSOM 136
JUSTCIV 102	MI 354
PI 64	ENER 224
JEUN 56	RECH 212
EDUC 198	IND 133

ПРИДРУЖИТЕЛНО ПИСМО

От: Генералния секретар на Европейската комисия,
подписано от г-н Jordi AYET PUIGARNAU, директор

Дата на получаване: 8 април 2019 г.

До: Г-н Jeppe TRANHOLM-MIKKELSEN, генерален секретар на Съвета на
Европейския съюз

№ док. Ком.: COM(2019) 168 final

Относно: СЪОБЩЕНИЕ НА КОМИСИЯТА ДО ЕВРОПЕЙСКИЯ ПАРЛАМЕНТ,
СЪВЕТА, ЕВРОПЕЙСКИЯ ИКОНОМИЧЕСКИ И СОЦИАЛЕН КОМИТЕТ
И КОМИТЕТА НА РЕГИОНИТЕ
**Изграждане на доверие в ориентирания към човека изкуствен
интелект**

Приложено се изпраща на делегациите документ COM(2019) 168 final.

Приложение: COM(2019) 168 final



Брюксел, 8.4.2019 г.
COM(2019) 168 final

**СЪОБЩЕНИЕ НА КОМИСИЯТА ДО ЕВРОПЕЙСКИЯ ПАРЛАМЕНТ, СЪВЕТА,
ЕВРОПЕЙСКИЯ ИКОНОМИЧЕСКИ И СОЦИАЛЕН КОМИТЕТ И КОМИТЕТА
НА РЕГИОНИТЕ**

Изграждане на доверие в ориентирания към човека изкуствен интелект

1. УВОД — ЕВРОПЕЙСКАТА СТРАТЕГИЯ ЗА ИЗКУСТВЕН ИНТЕЛЕКТ

Изкуственият интелект (ИИ) има потенциал да преобрази света ни — Благодарение на него може да се подобри здравното обслужване, да се намали потреблението на енергия, автомобилите да станат по-безопасни, а земеделските стопани да могат да използват по-ефикасно водата и природните ресурси. Изкуственият интелект може да се използва за прогнозиране на промените в околната среда и климата, подобряване на управлението на финансовия риск и производство на съобразени с нашите нужди продукти, при което се генерират по-малко отпадъци. Изкуственият интелект може да използва също за откриване на измами и заплахи за киберсигурността и позволява на органите на реда да се борят по-ефективно с престъпността.

Изкуственият интелект може да бъде от полза за цялото общество и за икономиката. Той е стратегическа технология, която понастоящем се разработва и използва с бързи темпове в цял свят. Същевременно изкуственият интелект е свързан с нови предизвикателства за бъдещето на труда и повдига правни и етични въпроси.

В отговор на тези предизвикателства и с цел извличане на максимална полза от възможностите, които изкуственият интелект предлага, през април 2018 г. Комисията публикува европейска стратегия¹. В стратегията хората са поставени в центъра на развитието на изкуствения интелект — става въпрос за **ориентиран към човека изкуствен интелект**. Това е тристранен подход за засилване на технологичния и промишлен капацитет на ЕС и въвеждане на изкуствения интелект в цялата икономика, подготвяне за социално-икономическите промени и осигуряване на подходяща етична и правна рамка.

В изпълнение на стратегията за изкуствения интелект през декември 2018 г. **Комисията представи координиран план², изготвен със съдействието на държавите членки**, за създаване на полезни взаимодействия, събиране на данни, които представляват основна суровина за множество приложения на изкуствения интелект, и увеличаване на съвместните инвестиции. Планът бе представен през декември 2018 г. Целта е да се стимулира трансграничното сътрудничество и да се мобилизират всички действащи лица с цел увеличаване на публичните и частните инвестиции до **поне 20 милиарда евро** годишно през идното десетилетие³. Комисията удвои своите инвестиции в изкуствения интелект в рамките на програма „Хоризонт 2020“ и планира да инвестира 1 милиард евро годишно по програма „Хоризонт Европа“ и програма „Цифрова Европа“ в подкрепа по-специално на общи пространства за данни в сферата на здравеопазването, транспорта и производството, големи експериментални съоръжения, като например интелигентни болници и инфраструктури за автоматизирани превозни средства, както и стратегическа програма за научни изследвания.

За да изпълни тази обща стратегическа програма за научни изследвания, иновации и внедряване, Комисията засили своя **диалог с всички имащи отношение заинтересовани страни** от промишлеността, изследователските институти и публичните органи. Новата програма „Цифрова Европа“ ще бъде от решаващо значение за подпомагане на достъпа до изкуствен интелект за малките и средните предприятия

¹ COM(2018) 237.

² COM(2018) 795.

³ За да помогне за постигането на тази цел, Комисията предложи през следващия програмен период 2021—2027 г. Съюзът да отпуска финансиране от поне 1 милиард евро годишно по програма „Хоризонт Европа“ и програма „Цифрова Европа“ за инвестиции в изкуствен интелект.

във всички държави членки чрез центрове за цифрови иновации, съоръжения за изпитвания и експерименти, пространства на данни и програми за обучение.

В съответствие с репутацията си за безопасни и висококачествени продукти, Европа предприема етичен подход към изкуствения интелект и по този начин засилва доверието на гражданите в развитието на цифровите технологии и цели предоставяне на конкурентно предимство за европейските дружества в областта на изкуствения интелект. Целта на настоящото съобщение е да се постави началото на всеобхватен пилотен етап с участието на възможно най-широк кръг заинтересовани страни, за да се изпита прилагането на практика на етични насоки за разработване и използване на изкуствен интелект.

2. ИЗГРАЖДАНЕ НА ДОВЕРИЕ В ОРИЕНТИРАНИЯ КЪМ ЧОВЕКА ИЗКУСТВЕН ИНТЕЛЕКТ

От европейската стратегия за изкуствения интелект и координирания план става ясно, че **доверието е предпоставка за гарантиране на ориентиран към човека подход към изкуствения интелект** — той не е цел сам по себе си, а инструмент в услуга на хората, чиято крайна цел е да се подобри тяхното благосъстояние. За да се постигне това, трябва да се гарантира **надеждността на изкуствения интелект**. Ценностите, върху които се основават нашите общества, трябва да бъдат напълно интегрирани в начина, по който той се развива.

Съюзът се основава на **ценностите на зачитане на човешкото достойнство, свободата, демокрацията, равенството, правовата държава, както и на зачитането на правата на човека**, включително правата на лицата, които принадлежат към малцинства⁴. Тези ценности се споделят от обществата във всички държави членки, в които доминират плурализмът, недискриминацията, толерантността, справедливостта, солидарността и равенството. В допълнение към това в **Хартата на основните права на ЕС** са събрани в един-единствен текст личните, гражданските, политическите, икономическите и социалните права на хората в рамките на ЕС.

ЕС разполага със **солидна нормативна уредба**, която ще определи глобалния стандарт за ориентиран към човека изкуствен интелект. С Общия регламент относно защитата на данните се гарантира висок стандарт за защитата на личните данни и се изисква прилагането на мерки за гарантиране на защитата на данните по подразбиране и още на етапа на проектирането⁵. С Регламента за свободното движение на нелични данни се премахват пречките пред това движение и се осигурява обработката на всички категории данни навсякъде в Европа. Неотдавна приетият Акт за киберсигурността ще помогне за укрепване на доверието в онлайн пространството, а предложеният

⁴ Освен това ЕС е страна по Конвенцията на ООН за правата на хората с увреждания.

⁵ Регламент (ЕС) № 2016/679. Общият регламент относно защитата на данните (ОРЗД) гарантира свободното движение на лични данни в рамките на Съюза. В него се съдържат разпоредби относно вземането на решения, базирани изцяло на автоматично обработване, включително профилирането. Засегнатите лица имат право да бъдат информирани за съществуването на автоматизирано вземане на решения и да получават смислена информация за логиката, използвана при това вземане на решения, както и за значението и предвидените последици от обработването на данните за тях самите. В такива случаи те също така имат право на човешка намеса, да изразят гледната си точка и да оспорят решението.

регламент за неприкосновеност на личния живот и електронните комуникации⁶ също преследва тази цел.

Изкуственият интелект носи със себе си нови предизвикателства, тъй като позволява на машините да се „учат“ и да вземат и изпълняват решения без човешка намеса. Не след дълго този вид функционалност ще се превърне в стандарт при редица видове стоки и услуги — от смартфоните до автоматизираните превозни средства, роботите и онлайн приложенията. При все това решенията, взети от алгоритми, могат да се базират на непълни и следователно ненадеждни данни, да бъдат манипулирани от извършители на кибератаки, да съдържат предубеждения или да са просто погрешни. Безрезервното прилагане на технологиите, докато са в процес на развитие, може да доведе до проблематични резултати, както и до нежелание от страна на гражданите да ги приемат или използват.

Вместо това технологиите за изкуствен интелект следва да се разработват по ориентиран към хората начин, който заслужава доверието на обществеността. Това означава, че приложенията на изкуствения интелект следва не само да бъдат в съответствие със закона, но и да не и да са съобразени с етични принципи и да не нанасят непреднамерени вреди. Многообразието по отношение на пола, расовия или етнически произход, религията или убежденията, уврежданията или възрастта, следва да се гарантира на всеки етап от развитието на изкуствения интелект. Приложенията на изкуствения интелект следва да оправомощават гражданите и да зачитат основните им права. Тяхната цел следва да бъде да подобряват способностите на хората, а не да ги заместват, както и да позволяват достъп на хората с увреждания.

Поради това е налице необходимост от **етични насоки**, които се основават на съществуващата нормативна уредба и се прилагат от разработчиците, доставчиците и ползвателите на изкуствен интелект на вътрешния пазар, за създаване на етични условия на равнопоставеност във всички държави членки. Ето защо Комисията създаде **експертна група на високо равнище по въпросите на изкуствения интелект**⁷, в която са представени широк кръг заинтересовани страни и която има за задача да изготви етични насоки за изкуствения интелект и да подготви набор от препоръки за общата политика в тази област. В същото време бе създаден **Европейският алианс за изкуствен интелект**⁸ — открита платформа за заинтересованите страни с над 2700 членове, целяща да предоставя по-широк принос към работата на експертната група на високо равнище.

Експертната група на високо равнище публикува първа версия на етичните насоки през декември 2018 г. След **консултация със заинтересованите страни**⁹ и **срещи с представители от държавите членки**¹⁰ експертната група за изкуствения интелект представи преработен документ на Комисията през март 2019 г. В своите коментари до

⁶ COM(2017) 10.

⁷ <https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>

⁸ <https://ec.europa.eu/digital-single-market/en/european-ai-alliance>

⁹ Тази консултация доведе до коментари от 511 организации, асоциации, предприятия, научноизследователски институти, отделни лица и други. Резюме на получените коментари е достъпно на адрес: https://ec.europa.eu/futurium/en/system/files/ged/consultation_feedback_on_draft_ai_ethics_guidelines_4.pdf

¹⁰ Работата на експертната група бе приета положително от държавите членки, като в заключенията на Съвета, приети на 18 февруари 2019 г., между другото бе отбелязано предстоящото публикуване на етичните насоки и бе изразена подкрепа за усилията на Комисията за извеждане на етичния подход на ЕС на световната сцена: <https://data.consilium.europa.eu/doc/document/ST-6177-2019-INIT/en/pdf>

момента заинтересованите страни като цяло приветстват практическото естество на насоките и конкретните напътствия, които те предлагат на разработчиците, доставчиците и ползвателите на изкуствен интелект относно начините за гарантиране на надеждност.

2.1. Насоки за надежден изкуствен интелект, изготвени от експертната група на високо равнище за изкуствения интелект

Насоките, изготвени от експертната група на високо равнище за изкуствения интелект, за които се отнася настоящото съобщение¹¹, са базирани по-конкретно върху работата на Европейската група по етика в науката и новите технологии и на Агенцията за основните права.

В насоките се изтъква, че за разработването на „надежден изкуствен интелект“ е необходимо той (1) да спазва закона, (2) да отговаря на етичните принципи и (3) да бъде стабилен.

Въз основа на тези три условия и на европейските ценности, посочени в раздел 2, в насоките се идентифицират седем ключови изисквания, които приложенията на изкуствения интелект трябва да спазват, за да се смятат за надеждни. Насоките включват и списък за оценка, който помага да се провери дали тези изисквания са изпълнени.

Седемте ключови изисквания са:

- Човешки фактор и надзор
- Техническа стабилност и безопасност
- Управление на данните и неприкосновеност на личния живот
- Прозрачност
- Многообразие, недискриминация и справедливост
- Благосъстояние на обществото и околната среда
- Отчетност

Тези изисквания са предназначени за всички системи с изкуствен интелект в различни среди и отрасли, но специфичният им контекст на прилагане следва да се взема предвид за конкретното им и пропорционално изпълнение, като се възприема подход, основан върху въздействието. Така например приложение на изкуствения интелект, което предлага неподходяща книга за четене, е много по-малко опасно от приложение, което поставя погрешна диагноза за рак, и следователно може да подлежи на не толкова строг надзор.

Насоките, изготвени от експертната група на високо равнище, не са обвързващи и не създават нови правни задължения. Въпреки това редица съществуващи (и често специфични за дадена употреба или област) разпоредби на правото на Съюза, разбира се, вече отразяват едно или няколко от тези ключови изисквания, например безопасността, защитата на личните данни, неприкосновеността на личния живот или правилата за защита на околната среда.

¹¹ <https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top>

Комисията приветства работата на експертната група на високо равнище за изкуствения интелект и я счита за ценен принос към своята работа по изготвяне на политики.

2.2. Ключови изисквания за надежден изкуствен интелект

Комисията подкрепя следните основни изисквания за надежден изкуствен интелект, базирани на европейските ценности. Тя насърчава заинтересованите страни да прилагат изискванията и да използват списъка за оценка при тяхното прилагане на практика, за да създадат подходяща среда на доверие за успешното развитие и използване на изкуствения интелект. Комисията ще се радва да получи мнения и коментари от заинтересованите страни, които ще ѝ помогнат да прецени дали списъкът за оценка, предоставен в насоките, се нуждае от корекции.

I. Човешки фактор и надзор

Системите с изкуствен интелект трябва да помагат на хората да вземат по-добри и по-информирани решения в съответствие с техните цели. Тези системи следва да помагат за изграждане на процъфтяващо и справедливо общество чрез подкрепа на човешката дейност и **основните права**, без да намаляват, ограничават или заблуждават човешката автономност. **Благосъстоянието на ползвателите** като цяло трябва да бъде основна цел на функционалността на системите.

Човешкият надзор помага да се гарантира, че дадена система с ИИ не накърнява автономността на човека и не причинява други неблагоприятни последици. В зависимост от специфичната базирана на изкуствен интелект система и нейната област на приложение трябва да се гарантират подходящи **мерки за контрол**, включително приспособимостта, точността и обяснимостта на съответната система¹². **Надзорът** може да бъде постигнат чрез механизми за управление, като например прилагане на подход, при който човекът участва в процеса (human-in-the-loop) или го наблюдава (human-on-the-loop) или контролира (human-in-command)¹³. Трябва да се гарантира, че публичните органи са в състояние да упражняват своите надзорни правомощия в съответствие със своя мандат. При равни други условия, колкото по-малко надзор върху дадена система с ИИ може да упражнява човекът, толкова по-задълбочено изпитване и по-строго управление се изисква.

¹² Общият регламент относно защитата на данните предоставя на отделните лица правото да не бъдат обект на решение, основаващо се единствено на автоматизирана обработка, когато това поражда правни последици за субектите на данните или по подобен начин сериозно ги засяга (член 22 от ОРЗД).

¹³ Човешкото участие (human-in-the-loop, HITL) се отнася до човешката намеса на всеки етап на вземане на решение от системата, което в много случаи не е нито възможно, нито желателно. Човешкото наблюдение (human-on-the-loop, HOTL) се отнася до възможността за човешка намеса по време на проектирането на системата и за мониторинг на функционирането на системата. Човешкият контрол (human-in-command, HIC) се отнася до способността за упражняване на надзор на цялостната дейност на системата с изкуствен интелект (включително на по-общото ѝ икономическо, социално, правно и етично въздействие) и възможността да се решава кога и как да се използва системата в дадена конкретна ситуация. Това може да включва решението да не се използва система с изкуствен интелект в дадена ситуация, да се установят степени на човешко участие по време на използването на системата или да се гарантира възможността за отмяна на решение, взето от системата.

II. Техническа стабилност и безопасност

За надежден изкуствен интелект се изискват алгоритми, които са достатъчно сигурни, надеждни и стабилни, за да се справят с грешки или несъответствия по време на всички етапи от жизнения цикъл на системата с изкуствен интелект, и които могат адекватно да се справят с грешни резултати. Системите с изкуствен интелект трябва да са **надеждни**, достатъчно **сигурни**, за да издържат както на явни атаки, така и на прикрити опити за манипулиране на данните или алгоритмите, и трябва да гарантират наличието на **резервен план** в случай на проблеми. Решенията на системите трябва да бъдат **точни** или поне да отразяват правилно тяхното ниво на точност, а резултатите следва да бъдат **възпроизводими**.

Освен това системите с изкуствен интелект следва да включват механизми за безопасност и сигурност от етапа на проектирането, за да се гарантира, че **безопасността им е проверяема** на всеки етап и че е взета под внимание физическата и психическата безопасност на всички засегнати страни. Тук влиза свеждането до минимум и, когато това е възможно, обратимостта на нежеланите последици или грешките при експлоатацията на системата. Следва да бъдат въведени процеси за изясняване и оценяване на потенциалните рискове, свързани с използването на системи с изкуствен интелект в различните области на приложение.

III. Управление на данните и неприкосновеност на личния живот

Неприкосновеността на личния живот и **защитата на данните** трябва да се гарантират на **всички етапи** от жизнения цикъл на системата с изкуствен интелект. Цифровите записи на човешкото поведение може да позволяват на системите с изкуствен интелект да правят заключения не само за възрастта, пола и предпочитанията на отделни лица, но и за тяхната сексуална ориентация и религиозни или политически възгледи. За да могат лицата да се доверяват на обработването на данни, трябва да се гарантира, че те имат пълен контрол върху данните, които ги засягат, и че тези данни не се използват по вреден или дискриминационен за тях начин.

Освен защитата на неприкосновеността на личния живот и на личните данни трябва да бъдат изпълнени и други изисквания за гарантиране на високото качество на системите с изкуствен интелект. Качеството на използваните набори от данни е от първостепенно значение за ефективността на системите с изкуствен интелект. Когато се събират данни, те могат да отразяват социално изградени предубеждения или да съдържат неточности и грешки. Тези проблеми трябва да бъдат разрешени, преди дадена система да бъде обучавана с набор от данни. Освен това трябва да се гарантира **целостта** на данните. Процесите и наборите от данни трябва да се изпитват и документируют на всеки етап — планиране, обучение, тестване и внедряване. Това следва да се прилага и за системите с изкуствен интелект, които не се разработени в рамките на дадена организация, а са придобити от друго място. Не на последно място, **достъпът** до данни трябва да бъде управляван и контролиран по подходящ начин.

IV. Прозрачност

Проследимостта на системите с изкуствен интелект следва да бъде гарантирана. Важно е да се регистрират и документируют както решенията, взети от системите,

така и целият процес (включително описание на събирането и етиктирането на данни и описание на използвания алгоритъм), който е довел до тези решения. Във връзка с това и доколкото е възможно, следва да се осигури **обяснимост** на алгоритмичния процес на вземане на решения, адаптиран към съответните лица. Нужни са постоянни изследвания за разработване на механизми за обяснимост. Освен това следва да бъдат на разположение и обяснения за степента, до която дадена система с изкуствен интелект оказва влияние върху организационния процес на вземане на решения, за проектантските решения при създаване на системата и за мотивите за въвеждането ѝ, като по този начин ще се гарантира прозрачността не само на системата, но и на съответния бизнес модел.

Не на последно място е важно възможностите и ограниченията на системата да се **съобщават** по адекватен начин на различните заинтересовани страни, участващи в използването ѝ. Освен това системите с изкуствен интелект следва да бъдат лесно разпознаваеми, за да се гарантира, че ползвателите им знаят, че взаимодействат с такава система, и кои са отговорните за нея лица.

V. Многообразие, недискриминация и справедливост

Наборите от данни, използвани от системите с изкуствен интелект (както за обучение, така и при експлоатация), може да са засегнати от включването на непреднамерени исторически предубеждения, непълнота и лоши модели на управление. Продължаването на подобни предубеждения може доведе до пряка или непряка дискриминация. Преднамереното използване на предубеждения (на потребителите) или участието в нелоялна конкуренция също могат да причинят вреда. Освен това начинът, по който се разработват системите с изкуствен интелект (напр. начинът, по който е написан програмният алгоритъм), също може да е засегнат от предубеденост. Тези въпроси следва да се разглеждат от самото начало на разработването на системата.

За разрешаване на подобни проблеми може да помогнат създаването на **различни проектантски екипи** и установяването на механизми за гарантиране на **участието**, по-специално на гражданите, в разработването на изкуствения интелект. Препоръчително е да се провеждат консултации със заинтересованите страни, които могат пряко или косвено да бъдат засегнати от системата през целия ѝ жизнен цикъл. При разработването на системи с изкуствен интелект следва да се взема предвид целия спектър от човешки способности, умения и изисквания и да се гарантира достъпност чрез универсален дизайн, целящ постигане на равнопоставен достъп за лицата с увреждания.

VI. Благосъстояние на обществото и околната среда

За да бъде надежден изкуственият интелект, трябва да се има предвид неговото въздействие върху **околната среда и другите същества с усещания**. В идеалния случай всички хора, включително бъдещите поколения, следва да живеят в условия на биологично разнообразие и обитаема околна среда. Ето защо устойчивостта и **екологичната отговорност** на системите с изкуствен интелект следва да бъдат насърчавани. Същото важи за решенията с използване на изкуствен интелект в области от световно значение, като например целите на ООН за устойчиво развитие.

Въздействието на системите с изкуствен интелект следва да се разглежда не само от

индивидуална гледна точка, но и от гледна точка на **обществото като цяло**. Използването на системи с изкуствен интелект следва да се разглежда внимателно, особено в ситуации свързани с демократичния процес, включително оформянето на мнения, вземането на политически решения или провеждането на избори. Освен това трябва да се разглежда и **социалното въздействие** на изкуствения интелект. Системите с изкуствен интелект могат да се използват за подобряване на социалните умения, но могат да допринесат и за тяхното влошаване.

VII. Отчетност

Нужни са механизми за гарантиране на отговорност и отчетност за системите с изкуствен интелект и резултатите от тях, както преди, така и след тяхното въвеждане. В това отношение е изключително важна **възможността за одитиране** на тези системи, тъй като оценката от вътрешни и външни одитори и наличието на подобни доклади за оценка допринася значително за надеждността на технологиите. Възможността за външен одит следва да се осигурява особено за приложения, засягащи основните права, включително приложения от критично значение за безопасността.

Потенциалните отрицателни въздействия на системите с изкуствен интелект следва да се идентифицират, оценяват, документират и свеждат до минимум. Използването на оценки на въздействието улеснява този процес. Тези оценки следва да бъдат пропорционални на степента на рисковете, свързани с тези системи. **Компромисите между изискванията** — които често са неизбежни — следва да се разглеждат по рационален и методичен начин и да се отчитат. Не на последно място, при възникване на несправедливи неблагоприятни последици, следва да бъдат предвидени достъпни механизми за гарантиране на **подходяща правна защита**.

2.3. Следващи стъпки: пилотна фаза с участието на възможно най-широк кръг заинтересовани страни

Постигането на консенсус относно тези ключови изисквания за системите с изкуствен интелект е важна първа стъпка по пътя към насоките за етичен изкуствен интелект. Като следваща стъпка Комисията ще се погрижи тези насоки да могат да бъдат изпитвани и прилагани на практика.

За тази цел Комисията дава началото на целева пилотна фаза за получаване на структурирани отзиви от заинтересованите страни. Действията ще бъдат съсредоточени по-специално върху списъка за оценка, съставен от експертната група на високо равнище за всяко едно от ключовите изисквания.

Работата ще протече в две направления: 1) пилотна фаза за насоките с участието на заинтересованите страни, които разработват или използват изкуствен интелект, включително публичните администрации, и 2) продължаваща консултация със заинтересованите страни и повишаване на осведомеността във всички държави членки и сред различните групи от заинтересовани страни, включително от промишлеността и сектора на услугите:

- i) От юни 2019 г. нататък всички заинтересовани страни и отделни лица ще бъдат приканени да изпробват списъка за оценка и да споделят своите мнения и

коментари относно начините за усъвършенстването му. Освен това експертната група на високо равнище за изкуствения интелект ще изготви задълбочен преглед заедно със заинтересовани страни от частния и публичния сектор, за да се събере по-подробна обратна информация относно начините, по които насоките могат да бъдат прилагани в широк кръг от области. Всички мнения и коментари за практичността и използваемостта на насоките ще бъдат анализирани до края на 2019 г.

- ii) Успоредно с това Комисията ще организира допълнителни осведомителни дейности, даващи възможност на представителите на експертната група на високо равнище да представят насоките на съответните заинтересовани страни в държавите членки, включително от секторите на промишлеността и услугите, и да им предоставят допълнителна възможност да коментират и да дадат своят принос към насоките за изкуствения интелект.

Комисията ще вземе под внимание работата на групата от експерти по етика за свързаното и автоматизирано шофиране¹⁴ и работата с финансирани от ЕС проекти за научни изследвания в областта на изкуствения интелект, както и работата със съответните публично-частни партньорства по прилагането на ключовите изисквания¹⁵. Например Комисията ще подкрепи в сътрудничество с държавите членки изготвянето на обща база данни с медицински изображения, първоначално посветени на най-често срещаните форми на рак, за да може съответните алгоритми да бъдат обучени да диагностицират симптомите с много висока степен на точност. По подобен начин сътрудничеството между Комисията и държавите членки позволява увеличаване на броя на трансграничните коридори за изпитване на свързани и автоматизирани превозни средства. Насоките следва да се прилагат и тестват в рамките на тези проекти, а резултатите ще бъдат използвани в процеса на оценяване.

Европейският алианс за изкуствен интелект и AI4EU, платформата за изкуствен интелект по заявка, също ще дадат своя принос по време на пилотната фаза и консултацията със заинтересованите страни. Проектът AI4EU¹⁶, стартирал през януари 2019 г., обединява алгоритми, инструменти, набори от данни и услуги в помощ на организациите, и по-специално малки и средни предприятия, които желаят да прилагат решения с използване на изкуствен интелект. Европейският алианс за изкуствения интелект заедно с AI4EU ще продължи да мобилизира участниците в екосистемата на изкуствения интелект в Европа, включително с цел прокарване на етичните насоки за изкуствения интелект и насърчаване на зачитането на ориентирания към човека изкуствен интелект.

В началото на 2020 г., въз основа на оценката на получените по време на пилотната фаза отзиви, експертната група на високо равнище ще направи преглед на насоките и ще ги актуализира. Въз основа на прегледа и на натрупания опит Комисията ще оцени резултатите и ще предложи следващи стъпки.

Етичният изкуствен интелект е от полза за всички. Гарантирането на зачитането на основните ценности и права е от първостепенно значение само по себе си, но също така

¹⁴ Вж. съобщението на Комисията относно свързаната и автоматизираната мобилност, COM(2018) 283.

¹⁵ В рамките на Европейския фонд за отбрана Европейската комисията ще разработи също така специфични етични насоки за оценката на предложенията за проекти за използване на изкуствен интелект за целите на отбраната.

¹⁶ <https://ec.europa.eu/digital-single-market/en/news/artificial-intelligence-ai4eu-project-launches-1-january-2019>

улеснява приемането от страна на обществеността и увеличава конкурентното предимство на европейските дружества в тази област, като утвърждава определен вид изкуствен интелект, който е надежден и ориентиран към човека и спомага за създаване на етични и сигурни продукти. Това се основава в по-общ план върху добрата репутация на европейските дружества при предоставяне на безопасни и сигурни продукти с високо качество. Пилотната фаза ще спомогне да се гарантира, че продуктите с изкуствен интелект изпълняват това обещание.

2.4. Към международни етични насоки за изкуствения интелект

Международните дискусии относно етичните аспекти на изкуствения интелект набраха скорост, след като японското председателството на Г-7 постави този въпрос на челно място в дневния ред през 2016 г. Като се имат предвид международните взаимовръзки при разработването на изкуствен интелект от гледна точка на движението на данните, разработването на алгоритми и инвестициите в научни изследвания, **Комисията ще продължи да полага усилия да даде гласност на подхода на ЕС в световен план и да изгради консенсус относно ориентирания към човека изкуствен интелект**¹⁷.

Работата на експертната група на високо равнище в областта на изкуствения интелект, и по-специално списъкът с изисквания и процесът на ангажиране на заинтересованите страни, предоставят на Комисията ценен допълнителен материал, който тя може да използва в международните дискусии. Европейският съюз може да играе водеща роля в разработването на международни насоки за изкуствения интелект и, ако това е възможно, на свързан с тях механизъм за оценка.

Поради това Комисията ще:

Засили сътрудничеството с партньори със сходни възгледи:

- като проучи до каква степен може да се постигне хармонизация с етичните насоки на други държави (напр. Япония, Канада, Сингапур) и като обедини усилията си с тези единомислещи страни за подготвяне на по-широка дискусия, подпомогната от действия за изпълнение на Инструмента за партньорство за сътрудничество с трети държави¹⁸; както и

¹⁷ Върховният представител на Съюза по въпросите на външните работи и политиката на сигурност, с подкрепата на Комисията, ще използва консултациите в рамките на ООН, световната експертна група по технологии (Global Tech Panel) и други многостранни форуми, по-специално за да координира предложенията за справяне с комплексните предизвикателства, свързани със сигурността, в тази област.

¹⁸ Регламент (ЕС) № 234/2014 на Европейския парламент и на Съвета от 11 март 2014 г. за създаване на Инструмент за партньорство за сътрудничество с трети държави (ОВ L 77, 15.3.2014 г., стр. 77). Например планираният проект за „Международен алианс за ориентиран към човека подход към изкуствения интелект“ ще улесни съвместните инициативи с единомислещи партньори с цел насърчаване на етичните насоки и приемане на общи принципи и оперативни заключения. Проектът ще даде възможност на ЕС и държавите със сходни възгледи да обсъждат оперативните заключения, произтичащи от етичните насоки за изкуствения интелект, предложени от експертната група, за да се достигне до общ подход. Освен това проектът ще позволи мониторинга на въвеждането на технологиите за изкуствен интелект в световен мащаб. Не на последно място, в рамките на проекта ще се организират дейности по публична дипломация, придружаващи международни събития, като например срещите на Г-7, Г-20 и Организацията за икономическо сътрудничество и развитие.

- като проучи как предприятията от държави извън ЕС и международните организации могат да допринесат за „пилотната фаза“ на насоките чрез тестване и валидиране.

Продължи да играе активна роля в международните дискусии и инициативи:

- като участва в международни форуми като Г-7 и Г-20;
- като участва в диалози с държави извън ЕС и организира двустранни и многостранни срещи за изграждане на консенсус относно ориентирания към човека изкуствен интелект;
- като допринася за съответните дейности по стандартизация на международните организации за разработване на стандарти с цел насърчаване на тази визия; както и
- като засилва събирането и разпространението на информация за публичните политики и работи съвместно със съответните международни организации.

3. ЗАКЛЮЧЕНИЯ

ЕС е изграден върху набор от основни ценности и е създал силна и балансирана нормативна уредба, стъпила върху тях. Въз основа на тази съществуваща уредба има нужда от изграждане на етични насоки за развитието и използването на изкуствения интелект поради неговата новост и специфичните технологични предизвикателства, характерни за него. Изкуственият интелект може да се счита за надежден само ако се разработва и използва при зачитане на утвърдените етични ценности.

С оглед на постигането на тази цел Комисията приветства материалите, подготвени от експертната група на високо равнище за изкуствения интелект. Въз основа на ключовите изисквания за надеждност на изкуствения интелект Комисията ще даде начало на пилотна фаза, за да се гарантира, че изготвените етични насоки за разработване и използване на изкуствения интелект могат да бъдат прилагани на практика. Комисията ще работи също така за постигане на широк обществен консенсус по отношение на ориентирания към човека изкуствен интелект заедно с всички участващи заинтересовани страни и нашите международни партньори.

Етичното измерение на изкуствения интелект не е луксозна опция или добавка — то трябва да бъде неразделна част от разработването му. Стремехът към ориентиран към човека и основан на доверие изкуствен интелект ни позволява да опазим фундаменталните ценности на нашето общество и дава възможност на Европа и нейната промишленост да се утвърдят като лидер в областта на свръхмодерния изкуствен интелект, на който може да се разчита в цял свят.

За да се гарантира етичното развитие на изкуствения интелект в Европа в по-широк план, Комисията се стреми към всеобхватен подход, който включва по-специално следните насоки на действие, които трябва да бъдат изпълнени до третото тримесечие на 2019 г.:

- Тя ще даде начало на редица центрове за **високи постижения в научните изследвания на изкуствения интелект** по програма „Хоризонт 2020“. Ще бъдат избрани до четири мрежи, съсредоточени върху важни научни или технологични предизвикателства, като например обяснимостта и напредналото взаимодействие между човека и машината, които са ключови съставки на надеждния изкуствен интелект.

- Комисията ще започне създаването на **мрежи от центрове за цифрови иновации**¹⁹, съсредоточени върху изкуствения интелект в производството и върху големите информационни масиви.
- Заедно с държавите членки и заинтересованите страни Комисията ще започне подготвителни обсъждания за разработване и прилагане на **модел за обмен на данни и оптимално използване на общите пространства на данни**, с акцент върху транспорта, здравеопазването и промишленото производство²⁰.

В допълнение към това Комисията работи върху доклад относно свързаните с изкуствения интелект предизвикателства за безопасността и отговорността и върху документ с насоки относно прилагането на Директивата относно отговорността за вреди от стоки²¹. Същевременно Съвместно предприятие за европейски високопроизводителни изчислителни технологии (EuroHPC)²² ще разработи следващото поколение суперкомпютри, тъй като изчислителният капацитет е от съществено значение за обработването на данни и обучаването на изкуствен интелект и Европа трябва да овладее цялата цифрова верига за създаване на стойност. Текущото партньорство с държавите членки и промишлеността в сферата на микроелектронните компоненти и системи (ECSEL)²³, както и инициативата за европейски процесор²⁴ ще допринесат за разработването на процесорна технология с ниска консумация на енергия за надеждни и сигурни високопроизводителни и модерни изчисления.

Подобно на работата по етични насоки за изкуствения интелект всички тези инициативи се основават на **тясно сътрудничество между всички заинтересовани страни**, държавите членки, промишлеността, обществените участници и гражданите. Като цяло подходът на Европа към изкуствения интелект показва как икономическата конкурентоспособност и общественото доверие трябва да се базират на едни и същи основни ценности и взаимно да се подкрепят.

¹⁹ <http://s3platform.jrc.ec.europa.eu/digital-innovation-hubs>

²⁰ Необходимите средства ще бъдат мобилизирани по линия на програма „Хоризонт 2020“ (по която близо 1,5 милиарда евро са предназначени за изкуствен интелект за периода 2018—2020 г.) и нейния наследник „Хоризонт Европа“, цифровата част на Механизма за свързване на Европа и особено бъдещата програма „Цифрова Европа“. Проектите ще разчитат на средства също така от частния сектор и програмите на държавите членки.

²¹ Вж. съобщението на Комисията „Изкуствен интелект за Европа“, COM(2018) 237.

²² <https://eurohpc-ju.europa.eu>

²³ www.ecsel.eu

²⁴ www.european-processor-initiative.eu