



V Bruseli 8. 4. 2019
COM(2019) 168 final

**OZNÁMENIE KOMISIE EURÓPSKEMU PARLAMENTU, RADE, EURÓPSKEMU
HOSPODÁRSKEMU A SOCIÁLNEMU VÝBORU A VÝBORU REGIÓNOV**

Budovanie dôvery v umelú inteligenciu sústredenú na človeka

1. ÚVOD – EURÓPSKA STRATÉGIA PRE UMELÚ INTELEGENCIU

Umelá inteligencia má potenciál zmeniť náš svet k lepšiemu: môže zlepšiť zdravotnú starostlivosť, znížiť spotrebu energie, zvýšiť bezpečnosť automobilov a umožniť poľnohospodárom efektívnejšie využívať vodu a prírodné zdroje. Umelú inteligenciu možno použiť na predpovedanie zmien v životnom prostredí a zmeny klímy, na zlepšenie riadenia finančných rizík a na poskytovanie nástrojov na výrobu produktov, ktoré sú prispôsobené našim potrebám, pri menšom množstve odpadu. Môže takisto pomôcť odhaliť podvody a kybernetické hrozby a umožňuje orgánom presadzovania práva efektívnejšie bojovať proti trestnej činnosti.

Umelá inteligencia môže byť prospešná pre celú spoločnosť a hospodárstvo. Ide o strategickú technológiu, ktorá sa v súčasnosti rýchlym tempom vyvíja a používa na celom svete. Umelá inteligencia však so sebou prináša aj nové výzvy pre budúcnosť práce a nastoľuje právne a etické otázky.

V záujme riešenia týchto výziev a čo najlepšieho využitia príležitostí, ktoré poskytuje umelá inteligencia, Komisia v apríli 2018 uverejnila európsku stratégiu¹. Táto stratégia stavia ľudí do centra vývoja umelej inteligencie – **umelej inteligencie sústredenej na človeka**. Ide o trojstupeňový prístup na podporu technologických a priemyselných kapacít EÚ a uplatňovanie umelej inteligencie v hospodárstve, na prípravu na sociálno-ekonomické zmeny a na zabezpečenie vhodného etického a právneho rámca.

Komisia v záujme plnenia stratégie v oblasti umelej inteligencie vypracovala spolu s členskými štátmi koordinovaný plán v oblasti umelej inteligencie², ktorý predložila v decembri 2018, s cieľom vytvárať synergie, zhromažďovať údaje – základ mnohých aplikácií umelej inteligencie – a zvyšovať spoločné investície. Cieľom je podporiť cezhraničnú spoluprácu a mobilizovať všetkých aktérov na zvýšenie verejných a súkromných investícií na **najmenej 20 miliárd EUR** ročne v nasledujúcom desaťročí³. Komisia zdvojnásobila svoje investície do umelej inteligencie v programe Horizont 2020 a plánuje ročne investovať 1 miliardu EUR z programov Horizont Európa a Digitálna Európa na podporu najmä spoločného dátového priestoru v oblasti zdravotníctva, dopravy a výroby a veľkých experimentálnych zariadení, ako sú inteligentné nemocnice a infraštruktúry pre, ako aj strategického výskumného plánu.

Na vykonávanie takéhoto spoločného strategického programu v oblasti výskumu, inovácie a zavádzania Komisia zintenzívnila **dialóg so všetkými príslušnými zainteresovanými stranami** z priemyslu, výskumných ústavov a verejných orgánov. Nový program Digitálna Európa bude takisto veľmi dôležitý pri sprístupňovaní umelej inteligencie pre malé a stredné podniky vo všetkých členských štátoch prostredníctvom centier digitálnych inovácií, posilnených testovacích a experimentálnych zariadení, dátových priestorov a programov odbornej prípravy.

Európsky etický prístup k umelej inteligencii vychádzajúci z dobrého mena Európy, pokiaľ ide o bezpečné a kvalitné výrobky, posilňuje dôveru občanov v digitálny rozvoj a zameriava sa na budovanie konkurenčnej výhody pre európske spoločnosti v oblasti umelej inteligencie. Účelom tohto oznámenia je začať komplexnú pilotnú fázu s čo najširšou účasťou

¹ COM(2018) 237.

² COM(2018) 795.

³ Komisia so zámerom pomôcť dosiahnuť tento cieľ v súvislosti s nasledujúcim programovým obdobím 2021 – 2027 navrhla, aby Únia vyčlenila aspoň 1 miliardu EUR ročne z financovania programov Horizont Európa a Digitálna Európa na investície do umelej inteligencie.

zainteresovaných strán s cieľom otestovať praktické vykonávanie etických usmernení pre vývoj a používanie umelej inteligencie.

2. BUDOVANIE DÔVERY V UMELÚ INTELIGENCIU SÚSTREDENÚ NA ČLOVEKA

Európska stratégia v oblasti umelej inteligencie a koordinovaný plán jasne ukazujú, že **dôvera je nevyhnutným predpokladom na zabezpečenie prístupu sústredeného na človeka v oblasti umelej inteligencie**: umelá inteligencia nepredstavuje samotný cieľ, ale je to nástroj, ktorý musí slúžiť ľuďom s konečným cieľom zvyšovania ľudského blahobytu. Na to **by sa mala zabezpečiť dôveryhodnosť umelej inteligencie**. Hodnoty, na ktorých sú založené naše spoločnosti, by mali byť v plnej miere začlenené do spôsobu vývoja umelej inteligencie.

Únia je založená na **hodnotách úcty k ľudskej dôstojnosti, slobody, demokracie, rovnosti, právneho štátu a rešpektovania ľudských práv** vrátane práv osôb patriacich k menšinám⁴. Tieto hodnoty sú spoločné pre spoločnosti všetkých členských štátov, v ktorých prevláda pluralizmus, nediskriminácia, tolerancia, spravodlivosť, solidarita a rovnosť. **Charta základných práv EÚ** okrem toho v jednom texte zoskupuje všetky osobné, občianske, politické, hospodárske a sociálne práva, ktoré požívajú ľudia v EÚ.

EÚ má **silný regulačný rámec**, ktorý stanoví globálny štandard pre umelú inteligenciu sústredenú na človeka. Všeobecné nariadenie o ochrane osobných údajov zabezpečuje vysoký štandard ochrany osobných údajov a vyžaduje sa v ňom vykonávanie opatrení na zabezpečenie špecificky navrhnuté a štandardnej ochrany údajov⁵. Nariadením o voľnom toku iných ako osobných údajov sa odstraňujú prekážky voľného pohybu iných ako osobných údajov a zabezpečuje sa spracovanie všetkých kategórií údajov kdekoľvek v Európe. Nedávno prijatý akt o kybernetickej bezpečnosti pomôže posilniť dôveru v internetový svet a na tento cieľ sa zameriava aj navrhované nariadenie o súkromí a elektronických komunikáciách⁶.

Umelá inteligencia však prináša nové výzvy, keďže umožňuje strojom „učiť sa“ a prijímať a vykonávať rozhodnutia bez ľudského zásahu. Už onedlho sa tento typ funkcií stane štandardom pri mnohých typoch tovaru a služieb od smartfónov až po automatizované autá, roboty a on-line aplikácie. Rozhodnutia prijaté algoritmami by však mohli prameniť z údajov, ktoré sú neúplné, a preto nie sú spoľahlivé, môžu byť zmanipulované pri kybernetických útokoch, môžu byť skreslené alebo jednoducho chybné. Nepremyslené uplatňovanie týchto technológií by v priebehu ich rozvoja preto viedlo k problematickým výsledkom, ako aj k neochote občanov ich prijímať alebo používať.

Namiesto toho by sa technológie umelej inteligencie mali vyvíjať spôsobom, ktorý sa sústreďuje na ľudí, a preto si zaslúži dôveru verejnosti. To znamená, že aplikácie umelej inteligencie by mali byť nielen v súlade so zákonom, ale aj s etickými zásadami a malo by sa zabezpečiť, aby sa pri ich zavádzaní zabránilo neúmyselnej ujme. V každej fáze rozvoja

⁴ EÚ je zmluvnou stranou Dohovoru OSN o právach osôb so zdravotným postihnutím.

⁵ Nariadenie (EÚ) 2016/679. Všeobecným nariadením o ochrane údajov (GDPR) sa garantuje voľný tok osobných údajov v rámci Únie. Obsahuje ustanovenia o rozhodovaní založenom výlučne na automatizovanom spracovaní vrátane profilovania. Dotknutí jednotlivci majú právo byť informovaní na existencii automatizovaného rozhodovania a získať zmysluplné informácie o logike automatizovaného rozhodovania a o význame a predpokladaných dôsledkoch spracovania pre týchto jednotlivcov. V takýchto prípadoch majú aj právo na ľudský zásah, právo vyjadriť svoje stanovisko a napadnúť rozhodnutie.

⁶ COM(2017) 10.

umelej inteligencie by sa mala zabezpečiť rozmanitosť z hľadiska rodu, rasového alebo etnického pôvodu, náboženstva alebo presvedčenia, zdravotného postihnutia a veku. Aplikácie umelej inteligencie by mali posilniť postavenie občanov a rešpektovať ich základné práva. Mali by sa zamerať na posilnenie schopností ľudí, nie na ich nahrádzanie, ako aj na zabezpečenie prístupu pre osoby so zdravotným postihnutím.

Preto sú potrebné **etické usmernenia**, ktoré budú vychádzať z existujúceho regulačného rámca a ktoré by mali uplatňovať vývojári, dodávatelia a používatelia umelej inteligencie na vnútornom trhu, čím sa vytvoria rovnaké etické podmienky vo všetkých členských štátoch. Komisia práve preto zriadila **expertnú skupinu na vysokej úrovni pre umelú inteligenciu**⁷ zastupujúcu široké spektrum zainteresovaných strán, ktorú poverila vypracovaním etických usmernení v oblasti umelej inteligencie, ako aj prípravou súboru odporúčaní pre širšiu politiku v oblasti umelej inteligencie. Zároveň bola vytvorená **Európska aliancia pre umelú inteligenciu**⁸, otvorená platforma viacerých zainteresovaných strán s viac ako 2 700 členmi, aby sa mohla vo väčšej miere zúčastňovať na práci expertnej skupiny na vysokej úrovni pre umelú inteligenciu.

Expertná skupina na vysokej úrovni v oblasti umelej inteligencie uverejnila v decembri 2018 prvý návrh etických usmernení. Po **konzultácii so zainteresovanými stranami**⁹ a **stretnutiach so zástupcami z členských štátov**¹⁰ expertná skupina pre umelú inteligenciu predložila v marci 2019 Komisii revidovaný dokument. V doterajšej spätnej väzbe zainteresované strany celkovo uvítali praktickú povahu usmernení a konkrétne rady, ktoré ponúkajú vývojárom, dodávateľom a používateľom umelej inteligencie v súvislosti s tým, ako zabezpečiť dôveryhodnosť.

2.1. Usmernenia pre dôveryhodnú umelú inteligenciu vypracované expertnou skupinou na vysokej úrovni pre umelú inteligenciu

Usmernenia vypracované expertnou skupinou na vysokej úrovni pre umelú inteligenciu, na ktoré sa vzťahuje toto oznámenie¹¹, vychádzajú najmä z práce Európskej skupiny pre etiku vo vede a nových technológiách a Agentúry pre základné práva.

V usmerneniach sa uvádza, že na dosiahnutie „dôveryhodnej umelej inteligencie“ sú potrebné tri zložky: 1. mala by byť v súlade s právnymi predpismi; 2. mala by spĺňať etické zásady a 3. mala by byť spoľahlivá.

Na základe týchto troch zložiek a európskych hodnôt uvedených v oddiele 2 sa v usmerneniach určuje sedem kľúčových požiadaviek, ktoré by mali aplikácie umelej inteligencie spĺňať, aby sa dali považovať za dôveryhodné. Usmernenia obsahujú aj zoznam posúdení, ktorý pomáha overiť, či sú tieto požiadavky splnené.

Ide to týchto sedem kľúčových požiadaviek:

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 Výsledkom tejto konzultácie boli pripomienky od 511 organizácií, združení, spoločností, výskumných ústavov, jednotlivcov a ďalších. Súhrn tejto prijatej spätnej väzby je dostupný na adrese: https://ec.europa.eu/futurium/en/system/files/ged/consultation_feedback_on_draft_ai_ethics_guidelines_4.pdf.

¹⁰ Prácu expertnej skupiny pozitívne prijali členské štáty, pričom v záveroch Rady prijatých 18. februára 2019 sa okrem iného zohľadnilo nadchádzajúce zverejnenie etických usmernení a podporilo úsilie Komisie presadiť etický prístup EÚ na celosvetovej úrovni:

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

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

- Ľudský faktor a dohľad
- Technická odolnosť a bezpečnosť
- Riadenie súkromia a údajov
- Transparentnosť
- Rozmanitosť, nediskriminácia a spravodlivosť
- Spoločenský a environmentálny blahobyt
- Zodpovednosť

Hoci sa majú tieto požiadavky vzťahovať na všetky systémy umelej inteligencie v rôznych prostrediach a odvetviach, pri ich konkrétnom a primeranom vykonávaní by sa mal zohľadniť osobitný kontext, v ktorom sa uplatňujú, pričom sa zaujme prístup založený na vplyve. Na ilustráciu, aplikácia umelej inteligencie, ktorá navrhne na čítanie nevhodnú knihu, je omnoho menej nebezpečná ako aplikácia, ktorá nesprávne diagnostikuje rakovinu, a preto by mohla podliehať menej prísnemu dohľadu.

Usmernenia vypracované expertnou skupinou na vysokej úrovni pre umelú inteligenciu sú nezáväzná a ako také nevytvárajú žiadne nové právne záväzky. Mnohé existujúce ustanovenia práva Únie (často aj špecifické pre konkrétne použitie či oblasť) už však zahŕňajú jednu alebo niekoľko z týchto kľúčových požiadaviek, napríklad pravidlá v oblasti bezpečnosti, ochrany osobných údajov, súkromia alebo ochrany životného prostredia.

Komisia víta prácu expertnej skupiny na vysokej úrovni pre umelú inteligenciu a považuje ju za hodnotný vstup pre tvorbu politiky.

2.2. Kľúčové požiadavky na dôveryhodnú umelú inteligenciu

Komisia podporuje nasledujúce kľúčové požiadavky na dôveryhodnú umelú inteligenciu, ktoré vychádzajú z európskych hodnôt. Vyzýva zainteresované strany, aby uplatňovali požiadavky a testovali zoznam posúdení, ktorým sa zavádzajú do praxe, s cieľom vytvoriť vhodné prostredie dôvery pre úspešný rozvoj a používanie umelej inteligencie. Komisia víta spätnú väzbu od zainteresovaných strán s cieľom vyhodnotiť, či si tento zoznam posúdení uvedený v usmerneniach vyžaduje ďalšiu úpravu.

I. Ľudský faktor a dohľad

Systémy umelej inteligencie by mali podporovať jednotlivcov pri prijímaní lepších a informovanejších rozhodnutí v súlade s ich cieľmi. Mali by pomáhať zabezpečiť prosperujúcu a spravodlivú spoločnosť, a to podporovaním ľudského faktora a **základných práv**, a nie znižovať a obmedzovať ľudskú autonómiu či viesť v tejto súvislosti k omylom. Ústrednou súčasťou funkčnosti tohto systému by malo byť celkové **blaho používateľa**.

Ľudský dohľad pomáha zabezpečiť, aby systém umelej inteligencie neoslaboval ľudskú autonómiu ani nespôsoboval iné nepriaznivé účinky. V závislosti od špecifického systému založeného na umelej inteligencii a od jeho oblasti použitia by sa mali zabezpečiť vhodné stupne **kontrolných opatrení** vrátane adaptability, presnosti a vysvetliteľnosti systémov

založených na umelej inteligencii¹². **Dohľad** sa môže dosiahnuť prostredníctvom mechanizmov riadenia, ako je zabezpečenie prístupu human-in-the-loop, human-on-the-loop alebo human-in-command.¹³ Musí sa zabezpečiť, aby verejné orgány boli schopné vykonávať dohľad v súlade so svojimi mandátmi. Ak sa nezmenia ostatné okolnosti, platí, že čím menej dohľadu môže človek vykonávať nad systémom umelej inteligencie, tým viac je potrebné rozsiahlejšie testovanie a prísnejšie riadenie.

II. Technická odolnosť a bezpečnosť

Dôveryhodná umelá inteligencia si vyžaduje dostatočne bezpečné, spoľahlivé a odolné algoritmy, ktoré sú schopné riešiť chyby alebo nezrovnalosti počas všetkých fáz životného cyklu systému umelej inteligencie a primerane si poradiť s chybnými výsledkami. Systémy umelej inteligencie musia byť **spoľahlivé**, dostatočne bezpečné na to, aby boli **odolné** proti zjavným útokom a sofistikovanejším pokusom o manipuláciu s údajmi alebo samotnými algoritmami, a musia zabezpečovať **záložný plán** v prípade problémov. Ich rozhodnutia musia byť **presné** alebo aspoň správne odrážať ich úroveň presnosti a ich výsledky by mali byť **reprodukovateľné**.

Systémy umelej inteligencie by okrem toho mali mať integrované mechanizmy bezpečnosti a mechanizmy špecificky navrhutej ochrany údajov, aby sa zabezpečilo, že sú v každom kroku **preukázateľne bezpečné**, pričom sa zohľadní fyzická a duševná bezpečnosť všetkých zainteresovaných. To zahŕňa minimalizáciu a podľa možnosti zvrátenie neúmyselných dôsledkov alebo chýb v prevádzke systému. Mali by sa zaviesť postupy na objasnenie a posúdenie potenciálnych rizík spojených s používaním systémov umelej inteligencie v rôznych oblastiach použitia.

III. Riadenie súkromia a údajov

Ochrana súkromia a údajov musí byť zaručená vo **všetkých fázach** životného cyklu systému umelej inteligencie. Digitálne záznamy ľudského správania môžu systémom umelej inteligencie umožniť vyvodiť nielen preferencie jednotlivcov, vek a pohlavie, ale aj ich sexuálnu orientáciu, náboženské alebo politické názory. Aby jednotlivci mohli dôverovať procesu spracovania údajov, musí sa zabezpečiť, aby mali plnú kontrolu nad vlastnými údajmi a aby sa údaje o nich nepoužili s úmyslom uškodiť im alebo ich diskriminovať.

Okrem zabezpečenia súkromia a osobných údajov musia byť splnené požiadavky na zabezpečenie kvalitných systémov umelej inteligencie. Kvalita používaných súborov údajov je pre výkon systémov umelej inteligencie rozhodujúca. Údaje môžu pri ich získavaní odrážať spoločenskú predpojatosť alebo môžu obsahovať nepresnosti, chyby a omyly. To je potrebné riešiť pred výcvikom systému umelej inteligencie s daným

¹² Všeobecné nariadenie o ochrane údajov poskytuje jednotlivcom právo, aby sa na nich nevzťahovalo rozhodnutie založené výlučne na automatizovanom spracovaní údajov, z ktorého vyplývajú právne účinky na používateľov alebo ich ovplyvňuje podobne významným spôsobom (článok 22 GDPR).

¹³ Human-in-the-loop (HITL) znamená ľudský zásah v každom rozhodovacom cykle systému, ktorý v mnohých prípadoch nie je možný ani žiaduci. Human-on-the-loop (HOTL) znamená možnosť ľudského zásahu počas cyklu navrhovania systému a monitorovania prevádzky systému. Human-in-command (HIC) znamená schopnosť dohliadať na celkovú činnosť systému umelej inteligencie (vrátane jeho širšieho hospodárskeho, spoločenského, právneho a etického vplyvu) a schopnosť rozhodnúť, kedy a ako používať systém v akejkoľvek konkrétnej situácii. To môže zahŕňať rozhodnutie nepoužívať systém umelej inteligencie v konkrétnej situácii, stanoviť úroveň ľudského rozhodovania počas používania systému alebo zabezpečiť schopnosť zrušiť rozhodnutie systému.

súborom údajov. Okrem toho sa musí zabezpečiť **integrita** údajov. Použité procesy a súbory údajov sa musia testovať a zdokumentovať pri každom kroku, ako je plánovanie, výcvik, testovanie a využívanie. To by sa malo vzťahovať aj na systémy umelej inteligencie, ktoré neboli vyvinuté interne, ale boli získané inde. V neposlednom rade sa musí primerane riadiť a kontrolovať **prístup** k údajom.

IV. Transparentnosť

Mala by sa zabezpečiť **vysledovateľnosť** systémov umelej inteligencie; je dôležité zaznamenávať a dokumentovať rozhodnutia prijaté systémami, ako aj celý proces (vrátane opisu získavania a označovania údajov a opisu použitého algoritmu), ktorý viedol k rozhodnutiam. V tejto súvislosti by sa mala v čo najväčšej miere zabezpečovať **vysvetliteľnosť** algoritmického procesu rozhodovania prispôsobeného dotknutým osobám. Malo by sa pokračovať v prebiehajúcom výskume zameranom na rozvoj mechanizmov vysvetliteľnosti. Okrem toho by mali byť k dispozícii vysvetlenia o tom, do akej miery systém umelej inteligencie ovplyvňuje a formuje organizačný proces rozhodovania, možnosti výberu systému, ako aj dôvody jeho zavedenia (čím by sa zabezpečila nielen transparentnosť údajov a systému, ale aj transparentnosť obchodného modelu).

Nakoniec je dôležité primerane **oznámiť** schopnosti a obmedzenia systému umelej inteligencie rôznym zainteresovaným stranám takým spôsobom, ktorý je vhodný pre daný prípad použitia. Okrem toho by systémy umelej inteligencie mali byť identifikovateľné ako také, čím by sa zabezpečilo, aby používatelia vždy vedeli, že komunikujú so systémom umelej inteligencie a ktoré osoby sú zaň zodpovedné.

V. Rozmanitosť, nediskriminácia a spravodlivosť

Súbory údajov používané systémami umelej inteligencie (pri výcviku aj prevádzke) môžu byť negatívne ovplyvnené zahrnutím neúmyselnej historickej predpojatosti, neúplnosťou a zlými modelmi riadenia. Pretrvávajúca takejto predpojatosti by mohlo viesť k (ne)priamej diskriminácii. Ujma môže vyplynúť aj z úmyselného zneužívania (spotrebiteľskej) predpojatosti alebo zapojenia sa do nekalej súťaže. Okrem toho môže predpojatosť negatívne ovplyvniť aj spôsob, akým sa systémy umelej inteligencie vyvíjajú (napr. spôsob, akým je napísaný programovací kód algoritmu). Takéto obavy by sa mali riešiť od začiatku vývoja systému.

Vytvorenie **rozmanitých projektových tímov** a zriadenie mechanizmov na zabezpečenie **účasti**, najmä občanov, pri rozvoji umelej inteligencie takisto môže pomôcť riešiť tieto obavy. Odporúča sa viesť konzultácie so zainteresovanými stranami, ktoré môže systém priamo alebo nepriamo ovplyvniť počas jeho životného cyklu. Systémy umelej inteligencie by mali zohľadňovať celú škálu ľudských schopností, zručností a požiadaviek a mala by sa zabezpečiť ich prístupnosť prostredníctvom prístupu dizajnu pre všetkých s cieľom dosiahnuť rovnaký prístup pre osoby so zdravotným postihnutím.

VI. Spoločenský a environmentálny blahobyť

Aby bola umelá inteligencia dôveryhodná, mal by sa zohľadniť jej vplyv na **životné prostredie a ostatné cítiace bytosti**. V ideálnom prípade by všetci ľudia vrátane budúcich generácií mali mať možnosť žiť v obývateľnom prostredí a biodiverzite. Preto

by sa mala podporovať udržateľnosť a **ekologická zodpovednosť** systémov umelej inteligencie. To isté platí aj pre riešenia umelej inteligencie, ktoré sa týkajú oblastí globálneho záujmu, ako sú napríklad ciele OSN v oblasti trvalo udržateľného rozvoja.

Okrem toho by sa mal vplyv systémov umelej inteligencie posudzovať nielen z pohľadu jednotlivca, ale aj z hľadiska **spoločnosti ako celku**. Využívanie systémov umelej inteligencie by sa malo starostlivo zväžiť najmä v situáciách týkajúcich sa demokratického procesu vrátane formovania názorov, politického rozhodovania alebo v súvislosti s voľbami. Okrem toho by sa mal zväžiť **sociálny vplyv** umelej inteligencie. Hoci sa systémy umelej inteligencie môžu použiť na zlepšenie sociálnych zručností, môžu rovnako prispieť k ich zhoršeniu.

VII. Zodpovednosť

Mali by sa zaviesť mechanizmy na zabezpečenie zodpovednosti systémov umelej inteligencie a ich výsledkov, a to pred ich zavedením aj po ňom. V tejto súvislosti je kľúčová **auditovateľnosť** systémov umelej inteligencie, keďže posúdenie systémov umelej inteligencie internými a externými audítormi a dostupnosť takýchto hodnotiacich správ výrazne prispieva k dôveryhodnosti technológií. Externá auditovateľnosť by sa mala zabezpečiť najmä pri aplikáciách, ktoré majú vplyv na základné práva, vrátane aplikácií dôležitých z hľadiska bezpečnosti.

Potenciálne negatívne vplyvy systémov umelej inteligencie by sa mali identifikovať, vyhodnotiť, zdokumentovať a minimalizovať. Tento proces uľahčuje využívanie posúdení vplyvu. Tieto posúdenia by mali zodpovedať rozsahu rizík, ktoré systémy umelej inteligencie predstavujú. Často nevyhnutné **kompromisy** medzi požiadavkami by sa mali riešiť racionálnym a metodickým spôsobom, pričom by sa mali zdôvodniť. Napokon, keď dôjde k nespravodlivému nepriaznivému vplyvu, mali by sa stanoviť dostupné mechanizmy, ktorými sa zabezpečí **primeraná náprava**.

2.3. Ďalšie kroky: pilotná fáza s čo najväčšou účasťou zainteresovaných strán

Dosiahnutie konsenzu o týchto kľúčových požiadavkách na systémy umelej inteligencie je prvým dôležitým míľnikom, pokiaľ ide o vypracovanie usmernení pre etickú umelú inteligenciu. Komisia v ďalšom kroku zabezpečí, aby sa tieto usmernenia mohli testovať a uplatňovať v praxi.

Na tento účel teraz Komisia začne cielenú pilotnú fázu, ktorej cieľom je získať štruktúrovanú spätnú väzbu od zainteresovaných strán. Táto analýza sa zameria najmä na zoznam posúdení, ktorý vypracovala expertná skupina na vysokej úrovni pre každú z kľúčových požiadaviek.

Táto práca sa zameria na dve oblasti: i) pilotnú fázu pre usmernenia s účasťou zainteresovaných strán, ktoré vyvíjajú alebo používajú umelú inteligenciu, vrátane verejných správ; a ii) pokračovanie konzultácií so zainteresovanými stranami a zvyšovania povedomia v členských štátoch a medzi rôznymi skupinami zainteresovaných strán vrátane priemyslu a sektora služieb:

- i) od júna 2019 budú všetky zainteresované strany a jednotlivci vyzvaní, aby otestovali zoznam posúdení a poskytli spätnú väzbu o tom, ako ho zlepšiť. Okrem toho expertná skupina na vysokej úrovni pre umelú inteligenciu vykoná hĺbkové preskúmanie so zainteresovanými stranami zo súkromného a verejného sektora s cieľom získať podrobnejšiu spätnú väzbu o tom, ako možno usmernenia vykonávať v

širokej škále oblastí využitia. Celá spätná väzba o funkčnosti a uskutočniteľnosti usmernení sa vyhodnotí do konca roku 2019.

- ii) Komisia zároveň zorganizuje ďalšie osvetové činnosti, pričom predstaviteľom expertnej skupiny na vysokej úrovni pre umelú inteligenciu poskytne príležitosť predložiť usmernenia príslušným zainteresovaným stranám v členských štátoch vrátane priemyslu a sektora služieb a poskytne im aj ďalšiu príležitosť vyjadriť sa k usmerneniam pre umelú inteligenciu a prispieť k nim.

Komisia zohľadní prácu expertnej skupiny pre etiku v súvislosti s prepojeným a automatizovaným riadením motorových vozidiel¹⁴ a pri vykonávaní kľúčových požiadaviek bude spolupracovať s výskumnými projektmi financovanými zo zdrojov EÚ v oblasti umelej inteligencie a s príslušnými verejno-súkromnými partnerstvami¹⁵. Komisia v spolupráci s členskými štátmi podporí napríklad vytvorenie spoločnej databázy zdravotných snímok pôvodne zameraných na najbežnejšie formy rakoviny, aby sa algoritmy mohli zamerať na diagnostiku symptómov s veľmi vysokou presnosťou. Spolupráca Komisie a členských štátov takisto umožňuje zvyšujúci sa počet cezhraničných koridorov na testovanie prepojených a automatizovaných vozidiel. V týchto projektoch by sa mali uplatňovať a testovať usmernenia, pričom výsledky budú súčasťou procesu hodnotenia.

Pilotná fáza a konzultácie so zainteresovanými stranami budú využívať prínos Európskej aliancie pre umelú inteligenciu a platformy na požiadanie pre umelú inteligenciu AI4EU. Projekt AI4EU¹⁶, ktorý sa začal v januári 2019, spája algoritmy, nástroje, súbory údajov a služby na pomoc organizáciám, najmä MSP, zavádzať riešenia v oblasti umelej inteligencie. Európska aliancia pre umelú inteligenciu spolu s AI4EU bude aj naďalej mobilizovať ekosystém umelej inteligencie v Európe, a to aj vzhľadom na pilotné testovanie etických usmernení v oblasti umelej inteligencie a podporu rešpektovania umelej inteligencie sústredenej na ľudí.

Na začiatku roka 2020 expertná skupina na vysokej úrovni pre umelú inteligenciu na základe hodnotenia spätnej väzby, ktorá sa získa počas pilotnej fázy, usmernenia prehodnotí a aktualizuje. Na základe preskúmania a získaných skúseností **Komisia vyhodnotí výsledok a navrhne prípadné ďalšie kroky.**

Etická umelá inteligencia je návrh prospešný pre všetkých. Zaručenie dodržiavania základných hodnôt a práv je nevyhnutné nielen samo osebe, ale uľahčuje aj prijatie zo strany verejnosti, pričom zvyšuje konkurenčnú výhodu európskych spoločností v oblasti umelej inteligencie tým, že zavádza dôveryhodnú umelú inteligenciu sústredenú na človeka ako značku pre etické a bezpečné produkty. Všeobecnejšie to vychádza z dobrej povesti európskych spoločností v oblasti poskytovania kvalitných bezpečných a zabezpečených produktov. Pilotná fáza pomôže zabezpečiť, aby produkty umelej inteligencie tento sľub splňali.

2.4. Vypracovanie medzinárodných etických usmernení v oblasti umelej inteligencie

Medzinárodné diskusie o etike umelej inteligencie sa zintenzívnili po tom, ako japonské predsedníctvo G7 zaradilo túto tému na popredné miesto programu v roku 2016. Vzhľadom

¹⁴ Pozri oznámenie Komisie o prepojenej a automatizovanej mobilite, COM(2018) 283.

¹⁵ V rámci Európskeho obranného fondu Komisia takisto vypracuje osobitné etické usmernenia pre hodnotenie návrhov projektov v oblasti umelej inteligencie na účely obrany.

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

na medzinárodné prepojenia v oblasti rozvoja umelej inteligencie, pokiaľ ide o obeh údajov, vývoj algoritmov a investície do výskumu, **Komisia bude pokračovať vo svojom úsilí presadiť prístup Únie na celosvetovej úrovni a budovať konsenzus v oblasti umelej inteligencie sústredenej na ľudí**¹⁷.

Práca expertnej skupiny na vysokej úrovni pre umelú inteligenciu, presnejšie zoznam požiadaviek a proces zapojenia zainteresovaných strán, poskytuje Komisii ďalšie cenné príspevky, ktorými môže prispieť do medzinárodných diskusií. Európska únia môže zohrávať vedúcu úlohu pri vypracúvaní medzinárodných usmernení v oblasti umelej inteligencie a podľa možnosti aj súvisiaceho mechanizmu posudzovania.

Komisia preto bude:

posilňovať spoluprácu s podobne zmýšľajúcimi partnermi:

- skúmať, do akej miery možno dosiahnuť konvergenciu s návrhom etických usmernení tretích krajín (napr. Japonsko, Kanada, Singapur), a prostredníctvom spolupráce s touto skupinou podobne zmýšľajúcich krajín pripraviť širšiu diskusiu podporenú opatreniami, ktorými sa vykonáva nástroj partnerstva pre spoluprácu s tretími krajinami¹⁸, a
- skúmať, ako môžu spoločnosti z krajín mimo EÚ a medzinárodné organizácie prispieť k „pilotnej fáze“ usmernení prostredníctvom testovania a overovania,

nadálej zohrávať aktívnu úlohu v medzinárodných diskusiách a iniciatívach:

- prispievať k mnohostranným fóram, ako sú G7 a G20,
- zapojiť sa do dialógu s krajinami mimo EÚ a organizovať dvojstranné a mnohostranné stretnutia s cieľom dosiahnuť konsenzus v oblasti umelej inteligencie sústredenej na ľudí,
- prispievať k príslušným normalizačným činnostiam v medzinárodných organizáciách pre rozvoj noriem na podporu tejto vízie a
- posilňovať získavanie a šírenie poznatkov o verejných politikách v spolupráci s príslušnými medzinárodnými organizáciami.

3. ZÁVERY

EÚ je založená na súbore základných hodnôt, pričom na týchto základoch vybudovala silný a vyvážený regulačný rámec. V nadväznosti na tento existujúci regulačný rámec treba vypracovať etické usmernenia pre rozvoj a používanie umelej inteligencie vzhľadom na jej novosť a špecifické výzvy, ktoré tieto technológie prinášajú. Umelá inteligencia sa môže

¹⁷ Vysoká predstaviteľka Únie pre zahraničné veci a bezpečnostnú politiku bude s podporou Komisie vychádzať z konzultácií v Organizácii Spojených národov, Globálnej technickej pracovnej skupine a na iných mnohostranných fórach a bude najmä koordinovať návrhy na riešenie príslušných zložitých bezpečnostných výziev.

¹⁸ Nariadenie Európskeho parlamentu a Rady (EÚ) č. 234/2014 z 11. marca 2014, ktorým sa ustanovuje nástroj partnerstva pre spoluprácu s tretími krajinami (Ú. v. EÚ L 77, 15.3.2014, s. 77). Napríklad plánovaný projekt „Medzinárodná aliancia pre koncepciu umelej inteligencie sústredenej na ľudí“ umožní spoločné iniciatívy s podobne zmýšľajúcimi partnermi s cieľom podporiť etické usmernenia a prijať spoločné zásady a operatívne závery. Umožní EÚ a podobne zmýšľajúcim krajinám diskutovať o operatívnych záveroch vyplývajúcich z etických usmernení o umelej inteligencii, ktoré navrhla expertná skupina, s cieľom dosiahnuť spoločný prístup. Okrem toho bude zabezpečovať globálne monitorovanie využívania technológií umelej inteligencie. Napokon sa v rámci projektu plánuje organizácia aktivít verejnej diplomacie popri medzinárodných podujatiach, napr. G7, G20 a podujatiach Organizácie pre hospodársku spoluprácu a rozvoj.

považovať za dôveryhodnú, len ak sa vyvíja a používa spôsobom, ktorý je v súlade so všeobecne zdieľanými etickými hodnotami.

Vzhľadom na tento cieľ Komisia víta príspevok, ktorý vypracovala expertná skupina na vysokej úrovni pre umelú inteligenciu. Komisia teraz na základe kľúčových požiadaviek na to, aby sa umelá inteligencia považovala za dôveryhodnú, začne cieľenú pilotnú fázu s cieľom zabezpečiť, aby sa výsledné etické usmernenia pre rozvoj a používanie umelej inteligencie mohli zaviesť do praxe. Komisia bude okrem toho pracovať na dosiahnutí širokého spoločenského konsenzu v oblasti umelej inteligencie sústredenej na človeka vrátane všetkých zainteresovaných strán a našich medzinárodných partnerov.

Etický rozmer umelej inteligencie nie je nadštandardným ani doplnkovým prvkom: musí byť neoddeliteľnou súčasťou rozvoja umelej inteligencie. V úsilí o dosiahnutie umelej inteligencie sústredenej na človeka, ktorá je založená na dôvere, zabezpečujeme rešpektovanie našich základných spoločenských hodnôt, pričom formujeme osobitnú ochrannú známku pre Európu a jej priemysel ako celosvetového lídra v oblasti špičkovej dôveryhodnej umelej inteligencie.

Aby sa zabezpečil etický rozvoj umelej inteligencie v Európe v jej širšom kontexte, Komisia uplatňuje komplexný prístup, ktorý zahŕňa najmä tieto smery činností, ktoré sa majú realizovať do tretieho štvrtého roka 2019:

- začne zavádzať súbor **sietí centier excelentnosti pre výskum umelej inteligencie** prostredníctvom programu Horizont 2020. Vyberie až štyri siete, ktoré sa zamerajú na hlavné vedecké alebo technologické výzvy, ako je vysvetliteľnosť a pokročilá interakcia medzi človekom a strojom, ktoré sú kľúčovými zložkami dôveryhodnej umelej inteligencie,
- začne vytvárať **siete centier digitálnych inovácií**¹⁹ so zameraním na umelú inteligenciu vo výrobe a na veľké dáta,
- Komisia spolu s členskými štátmi a zainteresovanými stranami začne prípravné diskusie s cieľom vypracovať a zaviesť **model spoločného využívania údajov a čo najlepšieho využívania spoločných dátových priestorov** so zameraním najmä na dopravu, zdravotníctvo a priemyselnú výrobu²⁰.

Komisia okrem toho vypracúva správu o výzvach v oblasti umelej inteligencie, pokiaľ ide o rámce bezpečnosti a zodpovednosti, a usmerňovací dokument o vykonávaní smernice o zodpovednosti za výrobky²¹. Európsky spoločný podnik vysokovýkonnej výpočtovej techniky (EuroHPC)²² zároveň vyvinie ďalšiu generáciu superpočítačov, keďže výpočtová kapacita je nevyhnutná na spracovanie údajov a výcvik umelej inteligencie a Európa potrebuje pokryť celý digitálny hodnotový reťazec. Prebiehajúce partnerstvo s členskými štátmi a priemyslom v oblasti mikroelektronických komponentov a systémov (ECSEL)²³, ako aj európska iniciatíva v oblasti procesorov²⁴ prispievajú k rozvoju technológie nízkoenergetických procesorov pre dôveryhodnú a bezpečnú vysokovýkonnú špičkovú výpočtovú techniku.

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

²⁰ Potrebne zdroje sa budú čerpať z programu Horizont 2020 (v rámci ktorého je na umelú inteligenciu na roky 2018 – 2020 vyčlenených približne 1,5 miliardy EUR) a z jeho plánovaného nástupcu Horizont Európa, z digitálnej zložky Nástroja na prepájanie Európy, a najmä z budúceho programu Digitálna Európa. Projekty budú čerpať aj zo zdrojov súkromného sektora a z programov členských štátov.

²¹ Pozri oznámenie Komisie „Umelá inteligencia pre Európu“, COM(2018) 237.

²² <https://eurohpc-ju.europa.eu>.

²³ www.ecsel.eu.

²⁴ www.european-processor-initiative.eu.

Rovnako ako práca na etických usmerneniach pre umelú inteligenciu, všetky tieto iniciatívy vychádzajú z **úzkej spolupráce všetkých príslušných zainteresovaných strán**, členských štátov, priemyslu, spoločenských aktérov a občanov. Prístup Európy k umelej inteligencii vo všeobecnosti ukazuje, ako hospodárska konkurencieschopnosť a spoločenská dôvera musia vychádzať z rovnakých základných hodnôt a navzájom sa posilňovať.