



Vijeće
Europske unije

Bruxelles, 9. travnja 2019.
(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

POP RATNA BILJEŠKA

Od: Glavni tajnik Europske komisije,
potpisao g. Jordi AYET PUIGARNAU, direktor

Datum primitka: 8. travnja 2019.

Za: g. Jeppe TRANHOLM-MIKKELSEN, glavni tajnik Vijeća Europske unije

Br. dok. Kom.: COM(2019) 168 final

Predmet: KOMUNIKACIJA KOMISIJE EUROPSKOM PARLAMENTU, VIJEĆU,
EUROPSKOM GOSPODARSKOM I SOCIJALNOM ODBORU I ODBORU
REGIJA **Izgradnja povjerenja u antropocentričnu umjetnu
inteligenciju**

Za delegacije se u prilogu nalazi dokument COM(2019) 168 final.

Priloženo: COM(2019) 168 final



Bruxelles, 8.4.2019.
COM(2019) 168 final

**KOMUNIKACIJA KOMISIJE EUROPSKOM PARLAMENTU, VIJEĆU,
EUROPSKOM GOSPODARSKOM I SOCIJALNOM ODBORU I ODBORU REGIJA**

Izgradnja povjerenja u antropocentričnu umjetnu inteligenciju

1. UVOD – EUROPSKA STRATEGIJA ZA UMJETNU INTELIGENCIJU

Umjetna inteligencija ima potencijal da naš svijet promijeni na bolje: njezinom upotrebom može se poboljšati zdravstvena zaštita, smanjiti potrošnja energije, povećati sigurnost automobila te povećati učinkovitost upotrebe vode i prirodnih resursa u poljoprivredi. Umjetna inteligencija može se upotrijebiti za predviđanje okolišnih i klimatskih promjena, za poboljšanje upravljanja financijskim rizikom i izradu proizvoda prilagođenih našim potrebama, uz manje otpada. Umjetna inteligencija također može pomoći u otkrivanju prijevara i kibersigurnosnih prijetnji, a tijelima kaznenog progona omogućuje učinkovitiju borbu protiv kriminala.

Umjetna inteligencija može koristiti cijelom društvu i gospodarstvu. Riječ je o strateškoj tehnologiji koja se ubrzano razvija i upotrebljava posvuda u svijetu. Međutim, umjetna inteligencija donosi nove izazove za budućnost rada i otvara pravna i etička pitanja.

Kako bi odgovorila na te izazove i u najvećoj mogućoj mjeri iskoristila mogućnosti koje pruža umjetna inteligencija, Komisija je u travnju 2018. objavila europsku strategiju¹. U strategiji su ljudi smješteni u središte razvoja umjetne inteligencije – **antropocentrične umjetne inteligencije**. Riječ je o trostranom pristupu za jačanje tehnoloških i industrijskih kapaciteta EU-a te uvođenje umjetne inteligencije u cijelom gospodarstvu, pripremu za društveno-gospodarske promjene i uspostavu odgovarajućeg etičkog i pravnog okvira.

Kako bi ostvarila ciljeve strategije za umjetnu inteligenciju, **Komisija je s državama članicama izradila koordinirani plan za umjetnu inteligenciju²** i predstavila ga u prosincu 2018. Svrha je plana stvoriti sinergije, prikupiti podatke (sirovinu za mnoge primjene umjetne inteligencije) i povećati zajednička ulaganja. Cilj je potaknuti prekograničnu suradnju i mobilizirati sve dionike kako bi se povećala javna i privatna ulaganja na **najmanje 20 milijardi EUR** godišnje tijekom sljedećeg desetljeća³. Komisija je udvostručila svoja ulaganja u umjetnu inteligenciju u okviru programa Obzor 2020. i planira svake godine uložiti 1 milijardu EUR iz programâ Obzor Europa i Digitalna Europa, osobito za potporu zajedničkim podatkovnim prostorima u zdravstvu, prometu i proizvodnji te velikim objektima za eksperimentiranje, kao što su pametne bolnice i infrastrukture za automatizirana vozila te strateški istraživački program.

Komisija je u cilju provedbe takvog zajedničkog strateškog plana istraživanja, inovacija i uvođenja pojačala **dijalog sa svim relevantnim dionicima** iz industrije, istraživačkih instituta i tijela javne vlasti. Novi program Digitalna Europa imat će ključnu ulogu i u povećanju dostupnosti umjetne inteligencije malim i srednjim poduzećima u svim državama članicama putem digitalnoinovacijskih centara, poboljšanih objekata za testiranje i eksperimentiranje, podatkovnih prostora i programa osposobljavanja.

Europskim etičkim pristupom umjetnoj inteligenciji, koji se temelji na ugledu koji Europa ima u području sigurnosti i visoke kvalitete proizvoda, jača se povjerenje građana u digitalni razvoj i nastoji ostvariti konkurentna prednost europskih poduzeća u sektoru umjetne inteligencije. Svrha je ove Komunikacije pokrenuti sveobuhvatnu pilot-fazu uz sudjelovanje najšireg spektra dionika kako bi se provjerila praktična provedba etičkih smjernica za razvoj i primjenu umjetne inteligencije.

¹ COM(2018) 237.

² COM(2018) 795.

³ Kako bi pridonijela ostvarivanju tog cilja, Komisija je u sljedećem programskom razdoblju 2021.–2027. predložila da Unija iz sredstava za financiranje programâ Obzor Europa i Digitalna Europa izdvoji najmanje 1 milijardu EUR godišnje za ulaganja u umjetnu inteligenciju.

2. IZGRADNJA POVJERENJA U ANTROPOCENTRIČNU UMJETNU INTELIGENCIJU

U europskoj strategiji za umjetnu inteligenciju i koordiniranom planu jasno je navedeno da je **povjerenje preduvjet da se osigura antropocentrični pristup umjetnoj inteligenciji**: umjetna inteligencija sama po sebi nije cilj, nego sredstvo koje treba služiti ljudima za ostvarenje krajnjeg cilja – veće dobrobiti ljudi. Kako bi se to postiglo, **potrebno je osigurati pouzdanost umjetne inteligencije**. Potrebno je vrijednosti na kojima se naša društva temelje u potpunosti integrirati u razvoj umjetne inteligencije.

Unija se temelji na **vrijednostima poštovanja ljudskog dostojanstva, slobode, demokracije, jednakosti, vladavine prava i poštovanja ljudskih prava**, uključujući i prava pripadnika manjina⁴. Te su vrijednosti zajedničke društvima svih država članica, u kojima prevladavaju pluralizam, nediskriminacija, tolerancija, pravda, solidarnost i jednakost. Nadalje, u **Povelji EU-a o temeljnim pravima** na jednom su mjestu obuhvaćena sva osobna, građanska, politička, gospodarska i socijalna prava stanovnika EU-a.

EU ima **snažan regulatorni okvir** koji će postaviti globalni standard za antropocentričnu umjetnu inteligenciju. Općom uredbom o zaštiti podataka osigurava se visok standard zaštite osobnih podataka i nalaže se provedba mjera za osiguravanje integrirane i zadane zaštite podataka⁵. Uredbom o slobodnom protoku neosobnih podataka uklanjaju se prepreke slobodnom kretanju neosobnih podataka i osigurava obrada svih kategorija podataka bilo gdje u Europi. Nedavno donesenim aktom o kibersigurnosti pridonijet će se jačanju povjerenja u svijetu interneta, a ista je svrha i predložene Uredbe o e-privatnosti⁶.

Bez obzira na to, umjetna inteligencija donosi nove izazove jer strojevima omogućuje „učenje” te donošenje i provedbu odluka bez ljudske intervencije. Ta vrsta funkcionalnosti ubrzo će postati standardna u mnogim vrstama roba i usluga, od pametnih telefona do automatiziranih automobila, robota i internetskih aplikacija. Ipak, odluke koje donose algoritmi mogu se temeljiti na nepotpunim, a stoga i nepouzdanim podacima, podacima koje su kibernetički napadači izmijenili, koji bi mogli biti pristrani ili jednostavno pogrešni. Nepromišljena primjena netom razvijene tehnologije mogla bi stoga dovesti do problematičnih ishoda kao i nesklonosti građana da je prihvate ili upotrebljavaju.

Umjesto toga, tehnologiju umjetne inteligencije trebalo bi razvijati na način koji ljude stavlja u njezino središte i time stječe njihovo povjerenje. To znači da bi primjene umjetne inteligencije trebale biti ne samo u skladu sa zakonom, nego i u skladu s etičkim načelima te bi trebalo osigurati da njihovom provedbom ne nastaju neželjene štete. U svakoj fazi razvoja umjetne inteligencije trebalo bi osigurati raznolikost u pogledu spola, rasnog ili etničkog podrijetla, vjere ili uvjerenja, invaliditeta i dobi. Primjene umjetne inteligencije trebale bi osnažiti građane i poštovati njihova temeljna prava. Njihov bi cilj trebao biti poboljšanje sposobnosti ljudi, a ne njihova zamjena, kao i omogućivanje pristupa osobama s invaliditetom.

⁴ Osim toga, EU je potpisnik Konvencije UN-a o pravima osoba s invaliditetom.

⁵ Uredba (EU) 2016/679. Općom uredbom o zaštiti podataka jamči se slobodan protok osobnih podataka unutar Unije. Ona sadržava odredbe o donošenju odluka isključivo na temelju automatizirane obrade, uključujući izradu profila. Pojedinci na koje se to odnosi imaju pravo biti obaviješteni o postojanju automatiziranog donošenja odluka i dobiti smislene informacije o logici automatskog donošenja odluka te o važnosti i predviđenim posljedicama obrade za njih. U takvim slučajevima imaju i pravo na ljudsku intervenciju, izraziti svoje stajalište i osporiti odluku.

⁶ COM(2017) 10.

Stoga je potrebno utvrditi **etičke smjernice** koje će se temeljiti na postojećem regulatornom okviru i kojih bi se trebali pridržavati nositelji projekata, dobavljači i korisnici umjetne inteligencije na unutarnjem tržištu, čime bi se uspostavili jednaki etički uvjeti u svim državama članicama. Zato je Komisija osnovala **stručnu skupinu na visokoj razini za umjetnu inteligenciju**⁷, koja okuplja širok spektar dionika, te ju je zadužila za izradu etičkih smjernica za umjetnu inteligenciju i preporuka za općenitiju politiku umjetne inteligencije. Istodobno je osnovan **Europski savez za umjetnu inteligenciju**⁸, otvorena višedionička platforma s više od 2700 članova kako bi se omogućilo šire sudjelovanje i doprinos radu stručne skupine na visokoj razini za umjetnu inteligenciju.

Stručna skupina na visokoj razini za umjetnu inteligenciju objavila je prvi nacrt etičkih smjernica u prosincu 2018. Nakon **savjetovanja s dionicima**⁹ i **sastanaka s predstavnicima država članica**¹⁰, stručna skupina za umjetnu inteligenciju dostavila je Komisiji revidirani dokument u ožujku 2019. Dionici su u povratnim informacijama općenito pozdravili praktičnost smjernica koje nositeljima projekata, dobavljačima i korisnicima umjetne inteligencije pružaju konkretne upute o načinima za osiguranje pouzdanosti.

2.1. Smjernice za pouzdanu umjetnu inteligenciju koje je izradila skupina stručnjaka na visokoj razini za umjetnu inteligenciju

Smjernice koje je izradila skupina stručnjaka na visokoj razini za umjetnu inteligenciju, na koje se upućuje u ovoj Komunikaciji¹¹, posebno se oslanjaju na rad Europske skupine za etiku u znanosti i novim tehnologijama i Agencije Europske unije za temeljna prava.

Osnovna je pretpostavka smjernica da su za uspostavu „pouzdana umjetne inteligencije” nužna tri uvjeta: (1) mora biti u skladu sa zakonom, (2) mora poštovati etička načela i (3) mora biti pouzdana.

Na temelju tih triju uvjeta i europskih vrijednosti utvrđenih u odjeljku 2., u smjernicama je utvrđeno sedam ključnih zahtjeva koje je potrebno ispuniti da bi se primjene umjetne inteligencije mogle smatrati pouzdanima. Smjernice također uključuju popis za procjenu kako bi se provjerilo jesu li ti zahtjevi ispunjeni.

Sedam ključnih zahtjeva su:

- ljudsko djelovanje i nadzor
- tehnička stabilnost i sigurnost
- privatnost i upravljanje podacima
- transparentnost
- raznolikost, nediskriminacija i pravednost

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 Tijekom savjetovanja primjedbe je poslalo 511 organizacija, udruga, poduzeća, istraživačkih instituta, pojedinaca i drugih subjekata. Sažetak primljenih povratnih informacija dostupan je na: https://ec.europa.eu/futurium/en/system/files/ged/consultation_feedback_on_draft_ai_ethics_guidelines_4.pdf

¹⁰ Države članice pozitivno su ocijenile rad stručne skupine, a u zaključcima Vijeća od 18. veljače 2019. među ostalim prima se na znanje predstojeća objava etičkih smjernica i podupiru naporu Komisije za afirmiranje etičkog pristupa EU-a na svjetskoj razini: <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>

- društvena i okolišna dobrobit
- odgovornost.

Iako se ti zahtjevi odnose na sve sustave umjetne inteligencije u različitim okruženjima i industrijama, za njihovu je konkretnu i razmjernu provedbu potrebno uzeti u obzir konkretan kontekst njihove primjene, uzimajući u obzir pristup koji se temelji na učinku. Na primjer, primjena umjetne inteligencije kojom se predlaže čitanje neprikladne knjige bitno je manje opasna od pogrešnog dijagnosticiranja raka te se stoga na nju može primjenjivati blaži nadzor.

Smjernice koje je izradila skupina stručnjaka na visokoj razini za umjetnu inteligenciju nisu obvezujuće i stoga ne stvaraju nikakve nove pravne obveze. Međutim, mnogi postojeći propisi Unije (koji su često specifični za određenu primjenu ili određeno područje) već odražavaju jedan ili više ključnih zahtjeva, na primjer sigurnost, zaštita osobnih podataka te pravila o privatnosti ili zaštiti okoliša.

Komisija pozdravlja rad stručne skupine na visokoj razini za umjetnu inteligenciju i smatra ga vrijednim doprinosom pri izradi njezinih politika.

2.2. Ključni zahtjevi za pouzdanu umjetnu inteligenciju

Komisija podržava sljedeće ključne zahtjeve za pouzdanu umjetnu inteligenciju, koji se temelje na europskim vrijednostima. Ona potiče dionike da primjenjuju zahtjeve i testiraju popis za procjenu u kojem se zahtjevi operacionaliziraju kako bi se stvorilo pravo okruženje povjerenja za uspješan razvoj i upotrebu umjetne inteligencije. Komisija izražava zadovoljstvo zbog povratnih informacija koje je dobila od dionika kako bi ocijenila treba li popis za procjenu iz smjernica dodatno prilagoditi.

I. Ljudsko djelovanje i nadzor

Sustavi umjetne inteligencije trebali bi podupirati pojedince u boljem donošenju informiranih odluka u skladu s njihovim ciljevima. Oni bi trebali djelovati kao pokretači naprednog i pravednog društva podupiranjem ljudskog djelovanja i **temeljnih prava**, a ne smanjivati, ograničavati ili pogrešno usmjeravati ljudsku autonomiju. Ukupna **dobrobit korisnika** trebala bi biti u središtu djelovanja sustava.

Ljudskim nadzorom osigurava se da sustav umjetne inteligencije ne ugrozi autonomiju čovjeka ili da ne uzrokuje druge negativne učinke. Ovisno o konkretnom sustavu koji se temelji na umjetnoj inteligenciji i njegovu području primjene, potrebno je osigurati odgovarajuće stupnjeve **kontrolnih mjera**, uključujući prilagodljivost, točnost i objašnjivost takvih sustava¹². **Nadzor** se može postići mehanizmima upravljanja kao što je osiguravanje ljudske intervencije (*human-in-the-loop*), ljudskog nadzora (*human-on-the-loop*) ili ljudskog odlučivanja (*human-in-command*).¹³ Mora se osigurati da tijela javne

¹² Općom uredbom o zaštiti podataka utvrđuje se da pojedinci imaju pravo da se na njih ne odnosi odluka koja se temelji isključivo na automatiziranoj obradi ako to proizvodi pravne učinke za korisnike ili na sličan način na njih znatno utječe (članak 22. Opće uredbe o zaštiti podataka).

¹³ Ljudska intervencija (HITL, *human-in-the-loop*) odnosi se na ljudsku intervenciju u svakom ciklusu odlučivanja sustava, što u mnogim slučajevima nije ni moguće ni poželjno. Ljudski nadzor (HOTL, *human-on-the-loop*) odnosi se na mogućnost ljudske intervencije tijekom ciklusa projektiranja sustava i praćenje rada sustava. Ljudsko odlučivanje (HIC, *human-in-command*) odnosi se na mogućnost nadzora ukupne aktivnosti sustava umjetne inteligencije (uključujući njegov širi gospodarski, društveni, pravni i etički

vlasti mogu izvršavati svoje nadzorne ovlasti u skladu sa svojim ovlastima. Ako su ostale okolnosti jednake, što je manja mogućnost ljudskog nadzora nad sustavom umjetne inteligencije, to je veća potreba za opsežnijim testiranjem i strožim upravljanjem.

II. Tehnička stabilnost i sigurnost

Nužna je pretpostavka pouzdane umjetne inteligencije da su algoritmi dovoljno sigurni, pouzdani i stabilni da mogu rješavati pogreške ili nedosljednosti u svim fazama životnog ciklusa sustava umjetne inteligencije te da na odgovarajući način rješavaju pogrešne ishode. Sustavi umjetne inteligencije moraju biti **pouzdan**, dovoljno sigurni da su **otporni** na očite napade i skrivenije pokušaje manipulacije podacima ili samim algoritmima i moraju osigurati **rezervni plan** ako nastanu problemi. Njihove odluke moraju biti **točne**, ili barem ispravno odražavati njihovu razinu točnosti, a njihovi ishodi trebaju biti **ponovljivi**.

Osim toga, sustavi umjetne inteligencije trebali bi uključivati integrirane sigurnosne mehanizme kako bi se osiguralo da se njihova **sigurnost može provjeriti** u svakoj fazi, pri čemu treba uzeti u obzir fizičku i mentalnu sigurnost svih zainteresiranih strana. To uključuje smanjenje na najmanju moguću mjeru i, ako je to moguće, reverzibilnost neželjenih posljedica ili pogrešaka u radu sustava. Trebalo bi uvesti postupke za pojašnjenje i procjenu potencijalnih rizika povezanih s upotrebom sustava umjetne inteligencije u različitim područjima primjene.

III. Privatnost i upravljanje podacima

Zaštita privatnosti i **podataka** mora se jamčiti u **svim fazama** životnog ciklusa sustava umjetne inteligencije. Digitalni zapisi o ljudskom ponašanju mogu sustavima umjetne inteligencije omogućiti da predvide ne samo želje, dob i spol pojedinaca, nego i njihovu spolnu orijentaciju, vjerska ili politička stajališta. Kako bi ljudi stekli povjerenje u obradu podataka, nužno je osigurati da imaju potpunu kontrolu nad vlastitim podacima i da se podaci koji se na njih odnose neće koristiti na njihovu štetu ili za njihovu diskriminaciju.

Uz zaštitu privatnosti i osobnih podataka, potrebno je ispuniti određene zahtjeve da bi se ostvarili visokokvalitetni sustavi umjetne inteligencije. Kvaliteta korištenih skupova podataka ključna je za uspješnost sustava umjetne inteligencije. Podaci koji se prikupljaju mogu odražavati društvene predrasude ili mogu sadržavati netočnosti i pogreške. Ta je pitanja potrebno riješiti prije treniranja sustava umjetne inteligencije s bilo kojim skupom podataka. Osim toga, mora se osigurati **cjelovitost** podataka. Postupci i skupovi podataka koji se upotrebljavaju moraju se testirati i dokumentirati u svakoj fazi, npr. u planiranju, treniranju, testiranju i uvođenju. To bi trebalo vrijediti i za sustave umjetne inteligencije koji nisu samostalno razvijeni, nego su nabavljeni od vanjskih dobavljača. Naposljetku, za **pristup** podacima moraju se utvrditi odgovarajući postupci upravljanja i kontrole.

IV. Transparentnost

učinak) i mogućnost odlučivanja o vremenu i načinu upotrebe sustava u određenoj situaciji. To može uključivati odluku da se sustav umjetne inteligencije ne upotrebljava u određenoj situaciji, da se utvrđuje razina diskrecijskog odlučivanja ljudi tijekom upotrebe sustava ili da se osigura mogućnost poništenja odluke koju je donio sustav.

Potrebno je osigurati **sljedivost** sustava umjetne inteligencije. Važno je evidentirati i dokumentirati odluke koje su donijeli sustavi, kao i cijeli postupak (uključujući opis prikupljanja i označivanja podataka te opis korištenih algoritama) koji je doveo do odluke. Povezano s time, potrebno je osigurati, u mjeri u kojoj je to moguće, **objašnjivost** postupka algoritamskog odlučivanja koji je prilagođen uključenim osobama. Trebalo bi neprestano provoditi istraživanja kako bi se razvili mehanizmi objašnjivosti. Nadalje, trebala bi biti dostupna objašnjenja u kojoj mjeri sustav umjetne inteligencije utječe na i oblikuje postupak odlučivanja u organizaciji, odabir pri izradi sustava, kao i razlog za njegovo uvođenje (čime se osigurava ne samo transparentnost podataka i sustava, nego i transparentnost poslovnog modela).

Konačno, važno je različitim uključenim dionicima na odgovarajući način **priopćiti** mogućnosti i ograničenja sustava umjetne inteligencije u skladu s konkretnim slučajem. Osim toga, sustave umjetne inteligencije trebalo bi moći prepoznati kao takve te osigurati da korisnici znaju da su u interakciji sa sustavom umjetne inteligencije i da su obaviješteni o odgovornim osobama.

V. Raznolikost, nediskriminacija i pravednost

Skupovi podataka koji se upotrebljavaju u sustavima umjetne inteligencije (za trening i rad) mogu uključivati nenamjerne naslijeđene predrasude, mogu biti nepotpuni i uključivati loše modele upravljanja. Daljnje prenošenje takvih predrasuda moglo bi dovesti do (ne)izravne diskriminacije. Šteta može nastati i namjernim iskorištavanjem pristranosti (potrošača) ili zbog nepoštenog tržišnog natjecanja. Osim toga, i sam način na koji se razvijaju sustavi umjetne inteligencije (npr. način pisanja programskog koda algoritma) može biti podložan predrasudama. Te probleme potrebno je rješavati od samog početka razvoja sustava.

Osnivanjem **projektnih timova koji se odlikuju različitostću** i uspostavom mehanizama kojima se osigurava **sudjelovanje**, posebno građana, u razvoju umjetne inteligencije također se može pomoći u rješavanju tih problema. Preporučuje se savjetovati se s dionicima na koje bi sustav mogao izravno ili neizravno utjecati tijekom njegova životnog ciklusa. Sustavima umjetne inteligencije trebalo bi uzeti u obzir sve ljudske sposobnosti, vještine i zahtjeve te osigurati dostupnost na temelju univerzalnog projektnog pristupa kako bi se ostvario jednak pristup za osobe s invaliditetom.

VI. Društvena i okolišna dobrobit

Da bi umjetna inteligencija bila pouzdana, mora se uzeti u obzir njezin utjecaj na **okoliš i na druga čuvstvena bića**. Idealno bi bilo da svi ljudi, uključujući buduće generacije, mogu uživati prednosti biološke raznolikosti i pogodnog okruženja. Stoga bi trebalo poticati održivost i **ekološku odgovornost** sustava umjetne inteligencije. Isto vrijedi i za sustave umjetne inteligencije koji se odnose na područja od svjetskog značaja, kao što su ciljevi održivog razvoja UN-a.

Nadalje, učinak sustava umjetne inteligencije treba razmotriti ne samo s pojedinačnog gledišta nego i iz perspektive **društva u cjelini**. Upotrebu sustava umjetne inteligencije trebalo bi pažljivo razmotriti, posebno u situacijama povezanim s demokratskim postupkom, uključujući oblikovanje mišljenja, donošenje političkih odluka ili izborne kontekste. Osim toga, potrebno je uzeti u obzir i **društveni učinak** umjetne inteligencije. Iako se sustavi umjetne inteligencije mogu koristiti za jačanje društvenih vještina, ti

sustavi mogu pridonijeti i njihovom slabljenju.

VII. Odgovornost

Potrebno je uspostaviti mehanizme kojima bi se utvrdile dužnosti i osigurala odgovornost za sustave umjetne inteligencije i njihove ishode prije i nakon njihova uvođenja. U tom je pogledu ključna **moгуćnost revizije** sustava umjetne inteligencije s obzirom na to da ocjenjivanje sustava umjetne inteligencije koju provode unutarnji i vanjski revizori te dostupnost takvih evaluacijskih izvješća snažno doprinosi pouzdanosti tehnologije. Posebno je važno osigurati mogućnost vanjske revizije za primjene koje se odnose na temeljna prava, uključujući one koje su ključne za sigurnost.

Potrebno je utvrditi, ocijeniti, dokumentirati i umanjiti **moгуće negativne učinke** sustava umjetne inteligencije. Primjenom procjena učinka taj se proces olakšava. Te procjene moraju biti razmjerne opsegu rizika koje sustavi umjetne inteligencije predstavljaju. **Kompromisi** između zahtjeva često su neizbježni i trebaju se rješavati racionalno i metodološki te se o njima mora voditi računa. Naposljetku, kada dođe do nepravednog nepovoljnog učinka, trebalo bi predvidjeti dostupne mehanizme kojima će se osigurati **odgovarajuća pravna zaštita**.

2.3. Sljedeći koraci: pilot-faza uz sudjelovanje najšireg spektra dionika

Postizanje konsenzusa o tim ključnim zahtjevima za sustave umjetne inteligencije prvi je važan korak prema smjernicama za etičku umjetnu inteligenciju. Komisija će u sljedećoj fazi osigurati testiranje i praktičnu provedbu tih smjernica.

U tu će svrhu Komisija pokrenuti ciljanu pilot-fazu koja je osmišljena tako da od dionika dobije strukturirane povratne informacije. Taj će postupak posebno biti usmjeren na popis za procjenu koji je stručna skupina na visokoj razini sastavila za svaki od ključnih zahtjeva.

Rad će se sastojati od dva dijela: i. pilot-faze za izradu smjernice u kojoj sudjeluju dionici koji razvijaju ili upotrebljavaju umjetnu inteligenciju, uključujući javne uprave i ii. daljnjeg savjetovanja s dionicima i podizanja svijesti u državama članicama i različitim skupinama dionika, uključujući industrijski i uslužni sektor:

- (i) od lipnja 2019. pozivat će se svi dionici i pojedinci da ispituju popis za procjenu i dostave povratne informacije o načinima za njegovo poboljšanje. Osim toga, stručna skupina na visokoj razini za umjetnu inteligenciju provest će s dionicima iz privatnog i javnog sektora temeljito preispitivanje kako bi prikupila detaljnije informacije o načinima provedbe smjernica u nizu područja primjene. Sve povratne informacije o izvedivosti i primjenjivosti smjernica ocijenit će se do kraja 2019.
- (ii) Komisija će istodobno organizirati daljnje aktivnosti informiranja i predstavnicima skupine na visokoj razini za umjetnu inteligenciju omogućiti da smjernice predstave relevantnim dionicima u državama članicama, uključujući industrijski i uslužni sektor, te da tim dionicima pruže dodatnu priliku da iznesu svoje primjedbe i doprinesu izradi smjernica za umjetnu inteligenciju.

Komisija će uzeti u obzir rad skupine stručnjaka za etiku u području povezane i automatizirane vožnje¹⁴ i rad na istraživačkim projektima u području umjetne inteligencije koje je financirao EU te u okviru relevantnih javno-privatnih partnerstava na provedbi

¹⁴ Vidjeti Komunikaciju Komisije o povezanoj i automatiziranoj mobilnosti za Europu, COM(2018) 283.

ključnih zahtjeva¹⁵. Na primjer, Komisija će, u suradnji s državama članicama, poduprijeti razvoj zajedničke baze podataka s medicinskim snimkama koja će se u početku koristiti za najčešće oblike karcinoma tako da se algoritmi mogu osposobiti za dijagnosticiranje simptoma s vrlo visokom točnošću. Slično tome, suradnja Komisije i država članica omogućuje sve veći broj prekograničnih koridora za testiranje povezanih i automatiziranih vozila. Smjernice bi se u tim projektima trebale primijeniti i ispitati, a rezultati će se uzeti u obzir u postupku evaluacije.

Provedbi pilot-faze i savjetovanja s dionicima pridonijet će i Europski savez za umjetnu inteligenciju i projekt AI4EU (platforma za umjetnu inteligenciju na zahtjev). Projekt AI4EU¹⁶, pokrenut u siječnju 2019., objedinjuje algoritme, alate, skupove podataka i usluge za pomoć organizacijama, posebno malim i srednjim poduzećima, u primjeni rješenja za umjetnu inteligenciju. Europski savez za umjetnu inteligenciju, zajedno s projektom AI4EU, nastavit će poticati razvoj ekosustava umjetne inteligencije u cijeloj Europi, među ostalim u području etičkih smjernica za umjetnu inteligenciju i promicanja antropocentrične umjetne inteligencije.

Početak 2020. stručna skupina na visokoj razini za umjetnu inteligenciju preispitat će i ažurirati smjernice na temelju ocjena povratnih informacija prikupljenih tijekom pilot-faze. Komisija će na temelju tog preispitivanja i stečenih iskustava **ocijeniti rezultate i predložiti sljedeće korake.**

Etična umjetna inteligencija donosi koristi svim dionicima. Jamstvo poštovanja temeljnih vrijednosti i prava bitno je samo po sebi, ali i zato što građanima olakšava prihvaćanje i povećava konkurentnost europskih poduzeća u području umjetne inteligencije uspostavljanjem pouzdane antropocentrične umjetne inteligencije za etičke i sigurne proizvode. To će se općenito temeljiti na znatnom ugledu europskih poduzeća u pružanju sigurnih i zaštićenih proizvoda visoke kvalitete. Rad u okviru pilot-faze pridonijet će tome da proizvodi umjetne inteligencije ta očekivanja ispune.

2.4. Prema međunarodnim etičkim smjernicama za umjetnu inteligenciju

Međunarodne rasprave o etici u umjetnoj inteligenciji intenzivirale su se nakon što je ta tema stavljena na popis prioriteta tijekom japanskog predsjedništva skupine G-7 u 2016. S obzirom na međunarodnu međusobnu povezanost razvoja umjetne inteligencije u smislu protoka podataka, razvoja algoritama i ulaganja u istraživanja, **Komisija će se i dalje zalagati za jačanje pristupa Unije na globalnoj razini i izgradnju konsenzusa o antropocentričnoj umjetnoj inteligenciji**¹⁷.

Rad skupine stručnjaka na visokoj razini za umjetnu inteligenciju, konkretnije popis zahtjeva i uključivanje dionika, vrijedan je dodatni doprinos Komisijinu sudjelovanju u međunarodnim raspravama. Europska unija može imati vodeću ulogu u razvoju međunarodnih smjernica za umjetnu inteligenciju i, po mogućnosti, povezanog mehanizma procjene.

¹⁵ U okviru Europskog fonda za obranu Komisija će izraditi i posebne etičke smjernice za ocjenu prijedloga projekata u području umjetne inteligencije za obranu.

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

¹⁷ Visoka predstavnica Unije za vanjske poslove i sigurnosnu politiku temeljit će svoj rad, uz potporu Komisije, na savjetovanjima u Ujedinjenim narodima, panelu „Global Tech Panel” i drugim multilateralnim forumima te osobito koordinirati prijedloge za rješavanje povezanih složenih sigurnosnih izazova.

Stoga će Komisija:

Jačati suradnju s partnerima sličnih stavova:

- istraživanjem mogućeg stupnja konvergencije s nacrtima etičkih smjernica trećih zemalja (npr. Japan, Kanada, Singapur) i na temelju iskustava tih zemalja s kojima dijeli stavove pripremiti širu raspravu, uz potporu mjera za provedbu Instrumenta za partnerstvo za suradnju s trećim zemljama¹⁸ i
- istraživanjem načina na koji poduzeća iz trećih zemalja i međunarodne organizacije ispitivanjima i potvrđivanjem mogu pridonijeti „pilot-fazi” izrade smjernica.

Nastaviti aktivno sudjelovati u međunarodnim raspravama i inicijativama:

- sudjelovanjem u radu multilateralnih foruma kao što su skupine G-7 i G-20
- dijalozima s trećim zemljama i organiziranjem bilateralnih i multilateralnih sastanaka radi izgradnje konsenzusa o antropocentričnoj umjetnoj inteligenciji
- sudjelovanjem u relevantnim aktivnostima normizacije u međunarodnim organizacijama za razvoj normi radi promicanja te vizije i
- poticanjem prikupljanja i širenja informacija o javnim politikama u suradnji s relevantnim međunarodnim organizacijama.

3. ZAKLJUČCI

EU se zasniva na skupu temeljnih vrijednosti i na tim je temeljima izgradio snažan i uravnotežen regulatorni okvir. Postoji potreba da se na temelju postojećeg regulatornog okvira izrade etičke smjernice za razvoj i primjenu umjetne inteligencije zbog njezine inovativnosti i specifičnih izazova povezanih s tom tehnologijom. Umjetna inteligencija može se smatrati pouzdanom samo ako se razvija i upotrebljava uz poštovanje općeprihvaćenih etičkih vrijednosti.

U tom kontekstu Komisija pozdravlja doprinos skupine stručnjaka na visokoj razini za umjetnu inteligenciju. Komisija će, uzimajući u obzir ključne zahtjeve koje umjetna inteligencija mora ispuniti da bi se smatrala pouzdanom, pokrenuti ciljanu pilot-fazu kako bi osigurala da se dobivene etičke smjernice za razvoj i primjenu umjetne inteligencije mogu provesti u praksi. Komisija će također nastojati izgraditi široki društveni konsenzus o antropocentričnoj umjetnoj inteligenciji sa svim uključenim dionicima i našim međunarodnim partnerima.

Etička dimenzija umjetne inteligencije nije luksuz ili dodatak: ona mora biti sastavni dio razvoja umjetne inteligencije. Naporima za ostvarivanje pouzdane antropocentrične umjetne inteligencije štitimo svoje temeljne društvene vrijednosti i omogućujemo Europi i njezinu

¹⁸ Uredba (EU) br. 234/2014 Europskog parlamenta i Vijeća od 11. ožujka 2014. o uspostavi Instrumenta za partnerstvo za suradnju s trećim zemljama (SL L 77, 15.3.2014., str. 77.). Na primjer, planirani projekt o međunarodnom savezu za antropocentrični pristup umjetnoj inteligenciji olakšat će zajedničke inicijative s partnerima sličnih stavova radi promicanja etičkih smjernica i donošenja zajedničkih načela i operativnih zaključaka. Omogućit će raspravu između EU-a i zemalja sličnih stavova radi donošenja operativnih zaključaka na temelju etičkih smjernica o umjetnoj inteligenciji koje je predložila skupina stručnjaka radi postizanja zajedničkog pristupa. Osim toga, osigurati će praćenje uvođenja tehnologije umjetne inteligencije u svijetu. Naposljetku, u okviru projekta planiraju se organizirati aktivnosti javne diplomacije koje prate međunarodne događaje, npr. u skupinama G-7, G-20 i u Organizaciji za gospodarsku suradnju i razvoj.

industriji da bude jasno prepoznata kao predvodnik u području vrhunske umjetne inteligencije koja se može smatrati pouzdanom u cijelom svijetu.

Kako bi se osigurao etički razvoj umjetne inteligencije u Europi u širem kontekstu, Komisija primjenjuje sveobuhvatan pristup, posebno za sljedeće mjere koje treba provesti do trećeg tromjesečja 2019.:

- Komisija će pokrenuti niz **mreža centara izvrsnosti za istraživanja umjetne inteligencije** u okviru programa Obzor 2020. Odabrat će se najviše četiri mreže koje će se usredotočiti na glavne znanstvene ili tehnološke izazove, kao što su objašnjivost i napredna interakcija čovjeka i stroja, koji su ključni elementi pouzdane umjetne inteligencije.
- Komisija će započeti s uspostavom **mreža digitalnoinovacijskih centara**¹⁹ usmjerenih na umjetnu inteligenciju u proizvodnji i velikim podacima.
- Komisija će zajedno s državama članicama i dionicima započeti pripremne rasprave za razvoj i provedbu **modela za razmjenu podataka i najbolje iskorištavanje zajedničkih podatkovnih prostora**, s posebnim naglaskom na promet, zdravstvo i industrijsku proizvodnju²⁰.

Nadalje, Komisija radi na izvješću o izazovima umjetne inteligencije u odnosu na okvire za sigurnost i odgovornost i na smjernicama za provedbu Direktive o odgovornosti za proizvode²¹. Istovremeno, Europsko zajedničko poduzeće za računalstvo visokih performansi (EuroHPC)²² razvit će sljedeću generaciju superračunala jer je računalni kapacitet ključan za obradu podataka i treniranje umjetne inteligencije, a Europa mora ovladati cjelokupnim digitalnim vrijednosnim lancem. Postojeće partnerstvo s državama članicama i industrijom u području mikroelektroničkih komponenata i sustava (ECSEL)²³, kao i Inicijativa za europski procesor²⁴, pridonijet će razvoju tehnologije procesora s niskom potrošnjom energije za pouzdano i sigurno računalstvo visokih performansi i računalstvo na rubu mreže.

Slično kao i u radu na etičkim smjernicama za umjetnu inteligenciju, sve te inicijative temelje se na **bliskoj suradnji svih uključenih dionika**, država članica, industrije, društvenih aktera i građana. Ukupno gledajući, pristup Europe umjetnoj inteligenciji pokazuje kako se gospodarska konkurentnost i društveno povjerenje moraju zasnivati na istim temeljnim vrijednostima i uzajamno jačati.

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

²⁰ Potrebna sredstva mobilizirat će se iz programa Obzor 2020. (u kojem je za umjetnu inteligenciju u razdoblju 2018.– 2020. namijenjeno gotovo 1,5 milijardi EUR) i programa koji će ga naslijediti, Obzor Europa, te digitalnog dijela Instrumenta za povezivanje Europe i posebno iz budućeg programa Digitalna Europa. Osim toga, projekti će se oslanjati na sredstva iz privatnog sektora i iz programa država članica.

²¹ Vidjeti Komunikaciju Komisije „Umjetna inteligencija za Europu”, COM(2018) 237.

²² <https://eurohpc-ju.europa.eu>

²³ www.ecsel.eu

²⁴ www.european-processor-initiative.eu