



Briselē, 8.4.2019.
COM(2019) 168 final

**KOMISIJAS PAZIŅOJUMS EIROPAS PARLAMENTAM, PADOMEI, EIROPAS
EKONOMIKAS UN SOCIĀLO LIETU KOMITEJAI UN REĢIONU KOMITEJAI**

**VAIROJOT UZTICĒŠANOS ANTROPOCENTRISKAM MĀKSLĪGAJAM
INTELEKTAM**

1. IEVADS. EIROPAS MĀKSLĪGĀ INTELEKTA STRATĒGIJA

Mākslīgais intelekts (AI) mūsu pasauli varētu padarīt labāku: tas spēj uzlabot veselības aprūpi, mazināt enerģijas patēriņu, padarīt drošākus automobiļus un palīdzēt lauksaimniekiem lietderīgāk izmantot ūdens un dabas resursus. AI izmantojams vides un klimata pārmaiņu prognozēšanā, finansiālā riska pārvaldības uzlabošanā un par rīku, ar kuru rodas mazāk atkritumu un top produkti, kas labāk piemēroti mūsu vajadzībām. AI var arī palīdzēt atklāt krāpšanos un kiberdrošības apdraudējumu un ļauj tiesībsardzības iestādēm efektīvāk apkarot noziedzību.

AI kalpo vispārībai un tautsaimniecībai. Tā ir stratēģiska tehnoloģija, kas tiek strauji izstrādāta un izmantota visā pasaulē. Bet tik un tā AI nāk ar jauniem pārbaudījumiem turpmākajam darbam un liek uzdot tiesiskas un ētiskas dabas jautājumus.

Problēmu risināšanai un AI piedāvāto iespēju optimālai izmantošanai Komisija 2018. gada aprīlī publicēja Eiropas stratēģiju¹. AI attīstības centrā stratēģija liek cilvēku – AI ir antropocentrisks. Tā ir trejzaru pieeja, kas attīsta ES tehnoloģisko un rūpniecisko jaudu un AI ieviešanu visā tautsaimniecībā, gatavo sociālekonomiskām pārmaiņām un nodrošina atbilstīgas ētiskās un tiesiskās normas.

AI stratēģijas iedzīvināšanai **Komisija kopā ar dalībvalstīm ir izstrādājusi koordinētu AI plānu²**, ar ko tā iepazīstināja 2018. gada decembrī: veidot sinerģiju, sakopot datus – daudzu AI lietojumu izejmateriālu – un paplašināt kopīgu ieguldīšanu. Tā mērķis ir veicināt pārrobežu sadarbību un mobilizēt visus iesaistītos, lai nākamajos desmit gados³ publiskos un privātos ieguldījumus palielinātu **vismaz līdz 20 miljardiem EUR** gadā. Komisija ir divkārtšojusi savu ieguldījumu AI no “Apvāršņa 2020” un plāno katru gadu ieguldīt 1 miljardu EUR no programmām “Apvāršnis Eiropa” un “Digitālā Eiropa”, sevišķi atbalstot kopīgās datu telpas veselības aprūpē, transportā un ražošanā, kā arī lielas eksperimentālas iekārtas, piemēram, viedslimnīcas un automatizēto transportlīdzekļu infrastruktūras, un stratēģiskās pētniecības programmu.

Tādas kopīgas stratēģiskas pētniecības, inovācijas un ieviešanas programmas īstenošanai Komisija ir pastiprinājusi **dialogu ar visām attiecīgajām ieinteresētajām personām**: uzņēmumiem, pētniecības institūtiem un valsts iestādēm. Ar digitālās inovācijas centriem, pastiprinātu testēšanu un eksperimentālām iekārtām, datu telpām un apmācību programmām jaunā programma “Digitālā Eiropa” arī būtiski palīdzēs AI padarīt pieejamu maziem un vidējiem uzņēmumiem visās dalībvalstīs.

Vairojot Eiropas drošo un kvalitatīvo produktu labo slavu, tās ētiskā pieeja AI stiprina pilsoņu uzticēšanos digitālajām pārmaiņām, un tās mērķis ir nodrošināt visiem Eiropas AI uzņēmumiem priekšrocības konkurencē ar citiem reģioniem. Šā paziņojuma mērķis ir sākt vispusīgu izmēģinājuma posmu, visplašākā mērogā iesaistot ieinteresētās personas, lai pārbaudītu, kā praksē īstenojamas ētikas vadlīnijas par AI izstrādi un izmantošanu.

2. VAIROT UZTICĒŠANOS ANTROPOCENTRISKAM MĀKSLĪGAJAM INTELEKTAM

¹ COM(2018) 237.

² COM(2018) 795.

³ Lai būtu vieglāk sasniegt šo mērķi, Komisija likusi priekšā Savienībai nākamajā plānošanas periodā (2021.–2027. gadam) atvēlēt vismaz 1 miljardu EUR gadā ieguldījumu AI finansēšanai no “Apvāršņa Eiropas” un “Digitālās Eiropas”.

Eiropas AI stratēģijā un koordinētajā plānā skaidri pateikts, ka **uzticēšanās ir priekšnoteikums antropocentriskai pieejai AI**: AI nav pašmērķis, bet rīks, kam jākalpo cilvēkam, lai uzlabotu cilvēka labklājību. Lai to panāktu, **jānodrošina AI uzticamība**. AI izstrādē ir pilnībā jāintegrē vērtības, uz kurām balstās mūsu sabiedrība.

Savienība ir dibināta uz **tādām vērtībām kā cilvēka cieņas respektēšana, brīvība, demokrātija, līdztiesība, tiesiskums un cilvēktiesību respektēšana**, ieskaitot mazākumiem piederīgo tiesības.⁴ Šīs vērtības ir kopīgas visu to dalībvalstu sabiedrībai, kurās dominē plurālisms, nediskriminēšana, iecietība, taisnīgums, solidaritāte un līdztiesība. **ES Pamattiesību harta** sakopo vienā tekstā visas personiskās, pilsoniskās, politiskās, saimnieciskās un sociālās tiesības, ko cilvēki bauda ES.

ES ir **spēcīgs tiesiskais regulējums**, kas tiecas noteikt pasaules standartu antropocentriskam AI. Vispārīgā datu aizsardzības regula nodrošina augstu aizsargātību persondatiem un prasa īstenot pasākumus, kas datu aizsardzību nodrošina integrēti un nevaicājo⁵. Nepersondatu brīvas plūsmas regula noņem šķēršļus brīvai nepersondatu aprītei un nodrošina visu kategoriju datu apstrādi Eiropas malā. Nesen pieņemtais Kiberdrošības akts palīdzēs stiprināt uzticēšanos tiešsaistes pasaulei, un tāds mērķis ir arī ierosinātajai E-privātuma regulai⁶.

Tomēr AI rada jaunas problēmas, jo tas ļauj mašīnām mācīties un pieņemt un īstenot lēmumus bez cilvēka iejaukšanās. Nebūs ilgi jāgaida, kad tāda funkcionalitāte kļūs par standartu daudzos preču un pakalpojumu veidos: no viedtālruniem līdz automatizētām automašīnām, robotiem un tiešsaistes lietojumiem. Tomēr algoritmu pieņemtie lēmumi var izrietēt no datiem, kuri ir nepilnīgi un tāpēc nav uzticami, kurus var sagrozīt kiberuzbrucēji vai kuri var būt neobjektīvi vai vienkārši kļūdaini. Nepārdomāti izmantojot tehnoloģiju, kas attīstījusies pati, rastos problemātisks iznākums, kā arī cilvēka nevēlēšanās to pieņemt un izmantot.

Jārīkojas otrādi – AI tehnoloģija jāattīsta tā, ka centrā nonāk cilvēks un tādējādi tā nopelna sabiedrības uzticēšanos. Ar to jāsaprot, ka AI lietojumiem ne tikai jāatbilst tiesību aktiem, bet arī jāievēro ētikas principi un jānodrošina, ka to īstenojums nevilšus nenodara kaitējumu. Ikvienā AI izstrādes posmā jānodrošina dažādība dzimuma, rases vai tautības, reliģijas vai ticības, invaliditātes un vecuma ziņā. AI lietojumiem jārada cilvēkiem iespējas un jārespektē viņu pamattiesības. Tiem jācenšas paplašināt cilvēku spējas, nevis viņi jāaizstāj, kā arī jānodrošina to pieejamība invalīdiem.

Tālab vajadzīgas **ētikas vadlīnijas**, kuras balstītos spēkā esošajā tiesiskajā regulējumā un kuras iekšējā tirgū piemērotu AI izstrādātāji, piegādātāji un lietotāji, tā radot noteikumus, kas ētikas ziņā vienādi visās dalībvalstīs. Šim nolūkam Komisija ir izveidojusi **AI augsta līmeņa ekspertu grupu**⁷, kas pārstāv plašu ieinteresēto personu loku, un uzdevusi tai izstrādāt AI ētikas vadlīnijas, kā arī sagatavot ieteikumu kopumu plašākai AI politikai. Tajā pašā reizē plašākam ieguldījumam AI augsta līmeņa ekspertu grupas darbā tika izveidota atvērta daudzinteresentu platforma **Eiropas AI alianse**⁸ ar vairāk nekā 2700 dalībniekiem.

⁴ Bez tam ES ir puse ANO Konvencijā par personu ar invaliditāti tiesībām.

⁵ Regula (ES) 2016/679. Vispārīgā datu aizsardzības regula (VDAR) garantē persondatu brīvu pārvietošanu Savienībā. Tajā ir noteikumi par lēmumu pieņemšanu pēc tikai automatizētas apstrādes, ieskaitot profilēšanu. Attiecīgajiem indivīdiem ir tiesības saņemt informāciju par automatizētu lēmumu pieņemšanu un saņemt jēgpilnu informāciju par automatizētajā lēmumu pieņemšanā izmantoto loģiku un par apstrādes nozīmi un paredzamajām sekām šiem indivīdiem. Tādos gadījumos viņiem ir arī tiesības prasīt cilvēka iejaukšanos, paust savu viedokli un lēmumu apstrīdēt.

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>

Ētikas vadlīniju pirmo projektu AI augsta līmeņa ekspertu grupa publicēja 2018. gada decembrī. Pēc **apspriešanās ar ieinteresētajām personām**⁹ un **sanāksmēm ar dalībvalstu pārstāvjiem**¹⁰ AI ekspertu grupa 2019. gada martā iesniedza Komisijai jaunu dokumenta redakciju. Līdzšinējās atsauksmēs ieinteresētās personas visumā atzinīgi novērtējušas vadlīniju praktiskumu un lietišķos norādījumus AI izstrādātājiem, piegādātājiem un lietotājiem par to, kā nodrošināt uzticamību.

2.1. AI augsta līmeņa ekspertu grupas sagatavotās uzticama AI vadlīnijas

AI augsta līmeņa ekspertu grupas izstrādātās vadlīnijas¹¹, par kurām tiek runāts šajā paziņojumā, galvenokārt balstās uz Eiropas Dabaszinātņu un jauno tehnoloģiju ētikas grupas un Pamattiesību aģentūras veikumu.

Vadlīnijās teikts, ka “uzticama AI” panākšanai vajadzīgi trīs komponenti: 1) tam jāatbilst likumam, 2) tam jāievēro ētikas principi, 3) tam jābūt noturīgam.

Uz šo komponentu un 2. iedaļā minēto Eiropas vērtību pamata vadlīnijās noteiktas septiņas pamatprasības, kas AI lietojumiem jāizpilda, lai tos varētu uzskatīt par uzticamiem. Vadlīnijās iekļauts arī vērtēšanas saraksts, kas palīdzēs pārbaudīt, vai šīs prasības tiek pildītas.

Ir septiņas pamatprasības:

- Cilvēka subjektivitāte un virsvadība
- Tehniskā noturība un drošums
- Privātums un datu pārvaldīšana
- Pārredzamība
- Dažādība, nediskriminēšana un taisnīgums
- Sabiedrības un vides labklājība
- Atbildīgums

Lai gan šīs prasības ir paredzētas piemērot visādām AI sistēmām dažādos apstākļos un nozarēs, to konkrētā un samērīgā īstenošanā ir jāņem vērā īpatnējais konteksts, kādā tās tiek īstenotas, un jāievēro uz ietekmi balstīta pieeja. Tā, AI lietojums, kas iesaka lasīt nepiemērotu grāmatu, ir daudz mazāk bīstams nekā vēža nepareiza diagnosticēšana, un tāpēc varētu tikt pakļauts mazāk stingrai uzraudzībai.

AI augsta līmeņa ekspertu grupas izstrādātās vadlīnijas nav saistošas un tādējādi nerada jaunus juridiskus pienākumus. Protams, daudzas pastāvošās Savienības tiesību normas (kas mēdz būt specifiskas lietošanas veidam vai jomai) jau atspoguļo vienu vai vairākas no šīm pamatprasībām, piemēram, drošuma, persondatu aizsardzības, privātuma vai vides aizsardzības normas.

⁹ Apspriešanās tika saņemtas piezīmes no 511 organizācijām, apvienībām, uzņēmumiem, pētniecības iestādēm, privātpersonām u. c. Atsauksmju kopsavilkums pieejams šajā vietnē: https://ec.europa.eu/futurium/en/system/files/ged/consultation_feedback_on_draft_ai_ethics_guidelines_4.pdf

¹⁰ Padomes 2019. gada 18. februāra secinājumos dalībvalstis atzinīgi novērtēja ekspertu grupas darbu, cita starpā ņemot vērā gaidāmo ētikas vadlīniju publicēšanu un atbalstot Komisijas pūliņus ES ētikas pieeju iznest pasaules arēnā: <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>

Komisija atzinīgi vērtē AI augsta līmeņa ekspertu grupas darbu un uzskata, ka tas ir vērtīgs ieguldījums Komisijas politikas veidošanā.

2.2. Pamatprasības uzticamam AI

Komisija atbalsta tālāk aprakstītās pamatprasības uzticamam AI, kuras balstās uz Eiropas vērtībām. Tā mudina visus ieinteresētos piemērot šīs prasības un izmēģināt vērtēšanas sarakstu, kas šīs prasības padara darbspējīgas, lai radītu vidi, kas uzticas AI sekmīgai attīstībai un izmantošanai. Komisija atzinīgi vērtē ieinteresēto personu atsauksmes, kas palīdz izvērtēt, vai vadlīnijās dotajā vērtēšanas sarakstā ir vajadzīgas vēl kādas korekcijas.

I. Cilvēka subjektivitāte un virsvadība

AI sistēmām jāpalīdz indivīdiem izdarīt labāku, zinīgāku izvēli, kas atbilst viņu mērķiem. Tām jādarbojas kā zeļošas un vienlīdzīgas sabiedrības veicinātājām, atbalstot cilvēka darbību un **pamattiesības**, nevis vājinot, iegrožojot vai novīrējot no ceļa cilvēka patstāvību. Sistēmas funkcionalitātes uzmanības centrā jābūt lietotāja vispārējai **labklājībai**.

Cilvēka virsvadība palīdz nodrošināt, ka AI sistēma neapdraud cilvēka patstāvību un neraisa citādas nelabvēlīgas sekas. Atkarā no katras uz AI balstītas sistēmas un tās pielietošanas jomas attiecīgā mērogā jānodrošina **kontroles pasākumi**, kas aptvertu arī pielāgojamību, precizitāti un uz AI balstītu sistēmu izskaidrojamību¹². **Virsvadībai** var izmantot pārvaldīšanas mehānismus, kas nodrošina pieeju *human-in-the-loop*, *human-on-the-loop* vai *human-in-command*¹³. Ir jānodrošina, ka valsts iestādes spēj īstenot savas virsvadības pilnvaras saskaņā ar saviem uzdevumiem. Pārējiem faktoriem nemainoties – jo mazāka AI sistēmas pārraudzība iespējama, jo plašāka testēšana un stingrāka pārvaldība vajadzīga.

II. Tehniskā noturība un drošums

Uzticams AI no algoritmiem prasa pietiekamu drošību, bezatteiksmi un noturību, kas novērstu kļūdas vai nekonsekvenci visos AI sistēmas cikla posmos un pienācīgi tiktu galā ar kļūmīgiem rezultātiem. AI sistēmām jābūt **neklūdīgām** un pietiekami aizsargātām, lai tās būtu **noturīgas** pret atklātu uzbrukumu un smalkākiem mēģinājumiem manipulēt ar datiem vai pašiem algoritmiem, un jānodrošina **atkāpšanas plāns** problēmu gadījumiem. To lēmumiem jābūt **precīziem** vai vismaz pareizi jādarbojas tiem noteiktajā precizitātes pakāpē, un to rezultātiem jābūt **reproducējamiem**.

Bez tam AI sistēmās jābūt integrētiem iestrādātās drošības un automātiskās drošības mehānismiem, lai panāktu, ka tās katrā posmā ir **pārbaudāmi drošas**, ņemot vērā visu

¹² Vispārīgā datu aizsardzības regula dod indivīdam tiesības, ka par viņu nevar pieņemt lēmumu, kura pamatā ir tikai automatizēta apstrāde, ja tas rada tiesiskas sekas lietotājiem vai tos līdzīgā veidā būtiski skar (VDAR 22. pants).

¹³ *Human-in-the-loop* (HITL) ir cilvēka iesaistīšanās katrā sistēmas lēmuma pieņemšanas ciklā, kas daudzos gadījumos nav ne iespējama, ne vēlama. *Human-on-the-loop* (HOTL) ir iespēja cilvēkam iejaukties sistēmas projektēšanas cikla laikā un pārraudzīt sistēmas darbību. *Human-in-command* (HIC) ir iespēja pārraudzīt AI sistēmas vispārējo darbību (arī plašāku ekonomisko, sociālo, juridisko un ētisko ietekmi) un spēja izlemt, kad un kā izmantot sistēmu katrā atsevišķā situācijā. Te var būt lēmumi neizmantot AI sistēmu atsevišķā situācijā, noteikt cilvēka rīcības brīvību sistēmas izmantošanā vai nodrošināt iespēju pārlabot sistēmas lēmumu.

iesaistīto personu fizisko un garīgo drošību. Tas ietver sistēmas darbībā negribētu seku vai kļūdu minimalizēšanu un, kur iespējams, izlabojamību. Jāievieš procesi, kas noskaidro un novērtē iespējamos riskus, kas saistīti ar AI sistēmu izmantošanu dažādās lietojumu jomās.

III. Privātums un datu pārvaldīšana

Privātums un **datu aizsardzība** garantējama **visos** AI sistēmas dzīves cikla **posmos**. Digitālas liecības par cilvēku uzvedību var ļaut AI sistēmām atminēt ne tikai indivīda paradumus, vecumu un dzimumu, bet arī viņa dzimumorientāciju un reliģiskos vai politiskos uzskatus. Lai indivīds varētu uzticēties datu apstrādei, ir jānodrošina, ka viņa pārziņā ir pilnīgi visi viņa dati un ka datus par viņu nevar izmantot, lai viņam kaitētu vai viņu diskriminētu.

Bez privātuma un persondatu aizsargāšanas ir jāizpilda arī prasības, kas AI sistēmām nodrošina augstu kvalitāti. AI sistēmu sekmīgā darbībā pati lielākā nozīme ir izmantoto datu kopu kvalitātei. Datu vākumos var atspoguļoties sociāli konstruēta neobjektivitāte vai neprecizitātes, kļūdas un maldi. Tas jāatrisina, pirms AI sistēmai liek apgūt kādu datu kopumu. Jānodrošina arī datu **veselums**. Izmantotie procesi un datu kopas ir jāpārbauda un jādokumentē katrā posmā: plānošanā, apmācībā, testēšanā, ieviešanā u. c. Tas attiecas arī uz AI sistēmām, kuras nav pašu izstrādātas, bet iegūtas citur. Visbeidzot, ir pienācīgi jāpārvalda un jākontrolē **piekļuve** datiem.

IV. Pārredzamība

Jānodrošina AI sistēmu **trasējamība**. Ir svarīgi reģistrēt un dokumentēt gan sistēmu lēmumus, gan visu procesu (ieskaitot aprakstu par datu vākšanu un marķēšanu, kā arī izmantotā algoritma aprakstu), kura rezultātā pieņemti lēmumi. Šai sakarā iespēju robežās jānodrošina algoritmiskā lēmumu pieņemšanas procesa **izskaidrojamība**, to pielāgojot iesaistītajām personām. Jāturpina pētījumi, kas izstrādā izskaidrojamības mehānismus. Jābūt pieejamiem arī paskaidrojumiem par to, kādā mērā AI sistēma ietekmē un veido organizatorisko lēmumu pieņemšanas procesu, sistēmas uzbūves izvēles iespējām, kā arī tās ieviešanas pamatojumu (tā nodrošinot ne tikai datu un sistēmas, bet arī uzņēmējdarbības modeļa pārredzamību).

Visbeidzot, ir svarīgi dažādām ieinteresētajām personām adekvāti **darīt zināmas** AI sistēmas spējas un ierobežojumus tādā veidā, kas piemērots attiecīgajam lietošanas gadījumam. Turklāt AI sistēmām jābūt identificējamām kā AI, nodrošinot, ka lietotāji zina, ka mijdarbojas ar AI sistēmu un kuras personas par to atbild.

V. Dažādība, nediskriminēšana un taisnīgums

Datu kopas, ko AI sistēmas izmanto (gan apmācībai, gan darbībai), var ciest no tajās iekļautas nejaušas vēsturiskas neobjektivitātes, nepilnīguma un sliktas pārvaldīšanas modeļiem. Neobjektivitātes turpināšanās var izraisīt (ne)tiešu diskrimināciju. Kaitējumu var radīt arī (patērētāja) neobjektivitātes tīša izmantošana vai iesaistīšanās negodīgā konkurencē. Turklāt neobjektīva var būt arī pati AI sistēmas izstrādāšana (piemēram, algoritma programmkoda sastādīšana). Piesardzība jāievēro no paša sistēmas izstrādes sākuma.

Palīdzēt risināt šīs problēmas var **dažādu projektēšanas grupu** veidošana un mehānismi,

kas nodrošina **dalību**, sevišķi pilsoņu dalību AI izstrādē. Ieteicams apspriesties ar ieinteresētajām personām, kuras sistēma var tieši vai netieši ietekmēt visā savā dzīves ciklā. AI sistēmām jāņem vērā viss cilvēku spēju, prasmju un prasību spektrs un pieejamība jānodrošina ar universālā dizaina pieeju, cenšoties panākt vienlīdzīgu pieejamību arī invalīdiem.

VI. Sabiedrības un vides labklājība

Lai AI būtu uzticams, ir jāņem vērā tā ietekme uz **vidi un jūtīgām būtnēm**. Būtu ideāli, ja visiem cilvēkiem, arī nākamajām paaudzēm, būtu bioloģiski daudzveidīga un apdzīvojama vide. Tālab jāveicina AI sistēmu ilgtspēja un **ekoloģiskā atbildība**. Tas pats attiecas uz AI globāli risināmās jomās, piemēram, ANO ilgtspējīgas attīstības mērķu sasniegšanā.

AI sistēmu ietekme jāskata ne tikai no indivīda, bet arī no **vispārības** viedokļa. AI sistēmu izmantošana rūpīgi jāapsver, it īpaši situācijās, kas saistītas ar demokrātijas procesu, ieskaitot viedokļu veidošanu, politisko lēmumu pieņemšanu vai vēlēšanu kontekstu. Bez tam jāņem vērā AI **sociālā ietekme**. Lai gan AI sistēmas ir izmantojamas sociālo iemaņu uzlabošanai, tās var arī veicināt to pasliktināšanos.

VII. Atbildīgums

Ir jāievieš mehānismi, kas nodrošinātu atbildību un norēķinu par AI sistēmām un to darbības rezultātiem – gan pirms, gan pēc to īstenošanas. Šajā ziņā liela nozīme ir AI sistēmu **pārbaudāmībai**, jo iekšējo un ārējo revidentu veiktā AI sistēmu novērtēšana un novērtēšanas ziņojumu pieejamība ievērojami veicina tehnoloģijas uzticamību. Ārējā pārbaudāmība īpaši jānodrošina lietojumos, kas skar pamattiesības, aptverot arī drošuma jomā kritiskus lietojumus.

Apzināma, novērtējama, dokumentējama un minimizējama ir AI sistēmu **potenciālā negatīvā ietekme**. Šo procesu atvieglo ietekmes novērtējumu izmantošana. Novērtēšanai jābūt samērīgai ar AI sistēmu radītajiem riskiem. **Kompromisi** starp prasībām reizēm nav izbēgami, bet tie jārisina racionāli un metodiski un par tiem jāsniedz pārskats. Visbeidzot, netaisnīgas negatīvas ietekmes gadījumos jābūt pieejamiem mehānismiem, kas nodrošina **pienācīgu tiesisko aizsardzību**.

2.3. Nākamais solis – izmēģinājuma posms, kurā visplašākā mērogā tiek iesaistītas ieinteresētās personas

Panākt vienprātību par AI sistēmām izvirzāmajām pamatprasībām ir pirmais svarīgais solis ētikas vadlīniju izstrādē. Nākamais – Komisija nodrošinās, ka norādījumus var pārbaudīt un iedzīvināt.

Šim nolūkam Komisija tagad sāks mērķtiecīgu izmēģinājuma posmu ar mērķi iegūt strukturētas atsauksmes no ieinteresētajām personām. Īpaša vērība pasākumā tiks veltīta vērtēšanas sarakstam, ko AI augsta līmeņa ekspertu grupa izstrādājusi visām galvenajām prasībām.

Šim darbam būs divas daļas: i) vadlīniju izmēģinājuma posms, kurā tiks iesaistītas ieinteresētās personas, kas izstrādā vai izmanto AI, ieskaitot valsts pārvaldes iestādes, ii) nemitīga konsultēšanās ar visiem ieinteresētajiem un izpratnes veidošanas process dalībvalstīs un dažādās ieinteresēto personu grupās, aptverot arī rūpniecību un pakalpojumus.

- (i) No 2019. gada jūnija visi ieinteresētie un interesenti tiks aicināti testēt vērtēšanas sarakstu un sniegt atsauksmes par to, kā to uzlabot. Turklāt AI augsta līmeņa ekspertu grupa kopā ar ieinteresētajām personām no privātā un publiskā sektora sarīkos pamatīgu pārskati, lai iegūtu detalizētas atsauksmes par to, kā vadlīnijas var īstenot visdažādākajās lietojumu jomās. Visas atsauksmes par vadlīniju izmantojamību darbā un īstenojamību izvērtēs līdz 2019. gada beigām.
- (ii) Līdztekus Komisija organizēs turpmākus informatīvus pasākumus, dodot AI augsta līmeņa ekspertu grupas pārstāvjiem iespēju iepazīstināt ar vadlīnijām attiecīgās ieinteresētās personas dalībvalstīs, ieskaitot rūpniecības un pakalpojumu nozares, un dot tām papildu iespēju izteikties un ieguldīt darbu AI vadlīniju izstrādē.

Komisija ņems vērā satīklotu un automatizētu transportlīdzekļu braukšanas ētikas ekspertu grupas darbu¹⁴ un pamatprasību īstenošanā sadarbosies ar ES finansētiem pētniecības projektiem AI jomā un attiecīgām publiskā un privātā sektora partnerībām¹⁵. Piemēram, Komisija sadarbībā ar dalībvalstīm atbalstīs kopīgas veselības attēlu datubāzes izstrādi, kura sākotnēji specializēsies visizplatītākajās vēža formās, lai varētu apmācīt algoritmus ar ļoti augstu precizitāti diagnosticēt simptomus. Līdzīgi Komisijas un dalībvalstu sadarbība ļauj palielināt pārrobežu koridoru skaitu satīklotu un automatizētu transportlīdzekļu izmēģināšanai. Šajos projektos vadlīnijas būtu jāpiemēro un jāizmēģina, un rezultāti tiks izmantoti novērtēšanas procesā.

Izmēģinājuma posms un apspriešanās ar ieinteresētajām personām izmantos Eiropas Mākslīgā intelekta aliansi un AI4EU – AI pieprasījumuplatformu. 2019. gada janvārī sāktais AI4EU projekts¹⁶ apvieno algoritmus, rīkus, datu kopas un pakalpojumus, lai palīdzētu organizācijām, it īpaši maziem un vidējiem uzņēmumiem, īstenot AI risinājumus. Eiropas Mākslīgā intelekta alianse kopā ar AI4EU turpinās mobilizēt AI ekosistēmu visā Eiropā, arī ņemot vērā AI ētikas vadlīnijas un vairojot cieņu pret antropocentrisku AI.

2020. gada sākumā, pamatojoties uz novērtējumu par atsauksmēm, kas būs saņemtas izmēģinājuma posmā, **AI augsta līmeņa ekspertu grupa vadlīnijas pārskatīs un atjauninās**. Balstoties uz šo pārskati un gūto pieredzi, **Komisija izvērtēs rezultātus un ierosinās nākamos pasākumus**.

Ētisks AI ir ieguvums bez zaudējumiem. Kopīgo vērtību un pamattiesību garantēšana ir būtiska ne tikai pati par sevi, bet tā arī atvieglo akceptēšanu sabiedrībā un palielina Eiropas AI uzņēmumu priekšrocības konkurencē, veidojot antropocentrisku, uzticama AI zīmolu, kurš ir pazīstams kā ētisks un drošs produkts. Tas kopumā balstās uz Eiropas uzņēmumu labo slavu un piedāvā drošus un aizsargātus kvalitatīvus ražojumus. Izmēģinājuma posms palīdzēs nodrošināt, ka AI produkti atbilst cerētajam.

2.4. Ceļā uz starptautiskām AI ētikas vadlīnijām

Pasaules reģionu diskusijas par AI ētiku ir pastiprinājušās pēc tam, kad Japānas G7 prezidentūra šo jautājumu izvirzīja priekšplānā 2016. gadā darba kārtībā. Tā kā AI attīstība pasaules mērogā izpaužas datu apritē, algoritmu attīstībā un ieguldījumos pētniecībā,

¹⁴ Sk. Komisijas paziņojumu par savienoto un automatizēto mobilitāti (COM(2018) 283).

¹⁵ Eiropas Aizsardzības fonda ietvaros Komisija izstrādās arī īpašus ētikas norādījumus par projektu priekšlikumu novērtēšanu AI jomā aizsardzībai.

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

Komisija turpinās Savienības pieeju izplatīt pasaules arēnā un veidot vienprātību jautājumā par antropocentrisku AI¹⁷.

AI augsta līmeņa ekspertu grupas darbs, proti, prasību saraksts un saistīšanās ar ieinteresētajām personām sniedz Komisijai vērtīgu ieguldījumu starptautisko diskusiju veicināšanai. Eiropas Savienībai var būt līderes loma starptautisku AI vadlīniju un, ja iespējams, ar to saistīta novērtēšanas mehānisma izstrādē.

Tādēļ Komisija:

- stiprinās sadarbību ar līdzīgi noskaņotiem partneriem,

- pētot, kādu konvergenci ar trešo valstu (piemēram, Japānas, Kanādas, Singapūras) ētikas vadlīniju projektiem var panākt, un, izmantojot šo līdzīgi noskaņoto valstu grupu, gatavosies plašākai diskusijai, ko atbalstītu darbības, ar kurām tiek īstenots Partnerības instruments sadarbībai ar trešām valstīm¹⁸, un
- pētot, kā uzņēmumi no ārpussavienības valstīm un starptautiskās organizācijas varētu palīdzēt vadlīniju izmēģinājuma posmā, tās testējot un apstiprinot to derīgumu.

- turpinās aktīvi piedalīties starptautiskās apspriedēs un iniciatīvās,

- veicinot tādus daudzpusējos forumus kā G7 un G20,
- iesaistoties dialogā ar ārpussavienības valstīm un rīkojot divpusējas un daudzpusējas sanāksmes, lai panāktu vienprātību par antropocentrisku AI,
- veicinot attiecīgās standartizācijas darbības starptautiskajās standartu izstrādes organizācijās, lai sekmētu ieceres īstenošanos, un
- sadarbībā ar attiecīgajām starptautiskajām organizācijām stiprinot informācijas vākšanu un izplatīšanu publiskās politikas jautājumos.

3. NOSLĒGUMS

ES balstās uz pamatvērtību kopumu, un uz tā ir izveidots stingrs un līdzsvarots tiesiskais regulējums. Balstoties uz esošo tiesisko regulējumu, ir vajadzīgas ētikas vadlīnijas AI izstrādei un izmantošanai, jo šī tehnoloģija ir jauna un tai ir īpatnējas problēmas. AI var uzskatīt par uzticamu tikai tad, ja tas tiek izstrādāts un izmantots, ievērojot plaši akceptētas ētiskās vērtības.

Komisija atzinīgi vērtē AI augsta līmeņa ekspertu grupas sagatavoto informāciju, kas kalpos šā mērķa sasniegšanai. Pamatojoties uz pamatprasībām, kas jāievēro, lai AI varētu uzskatīt par uzticamu, Komisija tagad sāks mērķpilnu izmēģinājuma posmu, lai nodrošinātu, ka ētikas vadlīnijas par AI izstrādi un izmantošanu ir īstenojamas praksē. Komisija arī strādās, lai

¹⁷ Savienības augstā pārstāve ārlietās un drošības politikas jautājumos ar Komisijas atbalstu izvērsīs apspriešanu ANO, Vispasaules tehnoloģiju ekspertu grupā un citos daudzpusējos forumos un, pats svarīgākais, koordinēs priekšlikumus par sarežģīto drošības problēmu risināšanu.

¹⁸ Eiropas Parlamenta un Padomes 2014. gada 11. marta Regula (ES) Nr. 234/2014, ar ko izveido Partnerības instrumentu sadarbībai ar trešām valstīm (OV L 77, 15.3.2014., 77. lpp.). Piemēram, plānotais projekts „Starptautiskā aliance antropocentriskai pieejai mākslīgajam intelektam” veicinās kopīgas iniciatīvas ar līdzīgi noskaņotiem partneriem, lai popularizētu ētikas vadlīnijas un pieņemtu kopīgus principus un secinājumus tālākam darbam. Tā ļaus ES un līdzīgi noskaņotām valstīm apspriest secinājumus darbam, kuri izriet no ekspertu grupas ierosinātajām ētikas vadlīnijām par AI, lai panāktu kopīgu pieeju. Tā arī nodrošinās AI tehnoloģijas ieviešanas pārraudzību visā pasaulē. Visbeidzot, projekts plāno organizēt tautas diplomātijas pasākumus līdztekus starptautiskiem pasākumiem, ko rīko G7, G20 un Ekonomiskās sadarbības un attīstības organizācija u. c.

panāktu plašu sabiedrības vienprātību par antropocentrisku AI, arī ar visām iesaistītajām ieinteresētajām personām un partneriem citur pasaulē.

AI ētiskais aspekts nav greznumlieta vai fakultatīva lasāmviela – tam jāklūst par AI attīstības sastāvdaļu. Cenšoties panākt antropocentrisku AI, kas balstītos uz uzticēšanos, mēs nodrošinām, ka tiek ievērotas mūsu sabiedrības pamatvērtības, kā arī darinām izcilu preču zīmi Eiropai un tās rūpniecībai kā līderēm tāda AI izstrādē, kam var uzticēties visa pasaule.

Lai nodrošinātu AI attīstību Eiropā plašākā kontekstā, Komisija īsteno vispusīgu pieeju, kas aptver šādas darbības jomas, kuras jāīsteno līdz 2019. gada trešajam ceturksnim:

- sākt veidot **AI pētniecības izcilības tīkla** centrus, izmantojot “Apvārsni 2020”, izvēlēties ne vairāk kā četrus tīklus, koncentrējoties uz tādām svarīgām zinātniskām vai tehnoloģiskām problēmām kā izskaidrojamība un cilvēka un mašīnas mijdarbība, kuras ir uzticama AI galvenie komponenti,
- sākt veidot **digitālās inovācijas centru tīklus**¹⁹, koncentrējoties uz AI izmantošanu ražošanā un uz lielajiem datiem.
- Kopā ar dalībvalstīm un ieinteresētajām personām Komisija sāks sagatavošanās diskusijas, lai izstrādātu un ieviestu **datu apmaiņas modeli un pēc iespējas labāk izmantotu kopīgās datu telpas**, īpaši pievēršoties transportam, veselības aprūpei un rūpnieciskajai ražošanai²⁰.

Komisija arī izstrādā ziņojumu par AI izraisītajām problēmām drošuma un atbildības jomā un vadlīniju dokumentu par Produktatbildības direktīvas īstenošanu²¹. Tajā pašā laikā Eiropas Augstas veikspējas datošanas kopuzņēmums ("EuroHPC")²² izstrādās nākamās paaudzes superdatorus, jo datošanas jauda ir būtiska datu apstrādei un AI apmācīšanai, un Eiropai ir jāapgūst visa digitālā vērtības veidošanas ķēde. Pašreizējās partnerības ar dalībvalstīm un nozari, kuras attiecas uz mikroelektronikas komponentiem un sistēmām (ECSEL)²³, kā arī Eiropas procesoru iniciatīva²⁴ veicinās energotaupīgu procesoru tehnoloģijas izstrādi uzticamai un drošai augstas veikspējas datošanai un perifērdatošanai.

Tāpat kā AI ētikas vadlīniju izstrāde, visas šīs iniciatīvas balstās uz **visu attiecīgo ieinteresēto personu**, dalībvalstu, rūpnieku, sabiedrisko procesu dalībnieku un pilsoņu **ciešu sadarbību**. Kopumā Eiropas pieeja AI rāda, ka ekonomikas konkurētspējai un sabiedrības uzticībai jāizaug no vienām pamatvērtībām un vienai otra savstarpēji jāpapildina.

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

²⁰ Nepieciešamos resursus sagādās no “Apvārsņa 2020” (kur aptuveni 1,5 miljardi EUR paredzēti AI laikposmam no 2018. līdz 2020. gadam) un tā plānotās aizstājējas programmas “Apvārsnis Eiropa”, no Eiropas infrastruktūras savienošanas instrumenta digitālās daļas un sevišķi no topošās programmas “Digitālā Eiropa”. Projektos izmantos arī privātā sektora un dalībvalstu programmu līdzekļus.

²¹ Sk. Komisijas paziņojumu "Mākslīgais intelekts Eiropai" (COM(2018) 237).

²² <https://eurohpc-ju.europa.eu>

²³ www.ecsel.eu

²⁴ www.european-processor-initiative.eu