

Mesterséges intelligencia: ha nem tudjuk kézben tartani, akkor általa mi pusztulunk el

Rab Árpáddal beszélget Kardos Ernő

Miután a mesterséges intelligencia potenciális veszélyt jelent az emberiségre, ezért egyre többen írják alá azt a kezdeményezést, amely arra szólítja fel a világ szakmai közvéleményét, hogy a GPT 4-nél ne építsenek erősebb mesterséges intelligenciát. A levél kezdeményezője Elon Musk üzletember, aláírója Steve Wozniak, az Apple társalapítója, Stuart Russell, a Kaliforniai Egyetem számítástechnikai professzora és Noah Harari történész. A szaktekintélyek szerint csak olyan mesterségesintelligencia-rendszereket szabad fejleszteni, amelyek hatása egyértelműen pozitív, kockázata kezelhető. A fejlesztés leállítással párhuzamosan a Guardian című brit lap publicisztikát íratott a GPT korábbi változatával, a 3-assal a mesterséges intelligencia veszélyeiről. Az algoritmus írását úgy kezelték, mint bármely újságíróét, de a robot írásának kijavítása rövidebb ideig tartott. A tekintélyes lapban a robot leírja, hogyan képezte magát az internetről, továbbá ígéri, hogy esze ágában sincs bántani az embereket, akár fel is áldozná magát értük. A GPT leírja azt is, hogy téves az a feltevés, mely szerint a robotok tervet szónének a hatalom átvételére. Lapunknak nyilatkozik a téma hazai szakértője, Rab Árpád egyetemi oktató, aki ugyancsak egyetért azzal, hogy egyelőre áll-

janak le a nagy teljesítményű gépek fejlesztésével.

– *Az elmúlt hetekben a világ rávetette magát a mesterséges intelligenciára, a legelképezetőbb kérdésekkel bombázzák a Chat GPT legújabb változatát, a 4-est. Az ön számára mi volt a legizgalmasabb kérdés, amire a mesterséges intelligencia válaszolt?*

– Bármilyen meglepő, de a szenzáció csak annyi volt, hogy a gép nyilvánosan megszólalt. Mesterséges intelligenciaszerű rendszereket már használtunk korábban is, elsősorban a számítógépes programokban, az orvostudományban, a katonaságnál is. Sőt, elterjedt a videójátékokban, az online szövegírásban, de azok még nem beszéltek. Vagyis a korábbi konstrukciók inkább a háttérben dolgoztak; egyes változatuk a fotók korrekciójában segédkezett, vagy a keresőprogramok működését segítette. Az elmúlt év vége felé jelent meg az Open AI GPT mai online nyilvánosan kipróbálható változata, amely aztán elindította a laikusok érdeklődését is. Ezt a chatbotot már kérdezgetni is lehetett, a gép pedig válaszolgatott. Az emberek millióihoz hasonlóan természetesen én magam is nagyon sokat foglalkoztam a tesztelésével, számomra az emberi értékekkel kapcsolatos kérdések a leg-

izgalmasabbak; a szubjektum és az objektivitás határvonlát próbáltam a kérdéseimmel kitapintani. Az én feladatom a mesterséges intelligenciák finomhangolása, ezt a munkát teszteltem. Persze a terület kutatójaként tudom, hogy ez végül is egy gép, amelynek az agyába telepített adatbázisok alapján várható válasz. Ezek nyelvi megfogalmazása jelen esetben rendkívül udvarias, a Chat GPT ezen változata úgy van beállítva, hogy a válaszaival senkit ne sértsen meg, hiszen egy nyilvános, általános teszt.

– *Ezzel szemben úgy érzem, hogy néha kínosan kerül az élet konfliktusos témáit. Pedig sokat segíthetne a tények tisztázásában, ha a segítségével kontrollálni lehetne a döntéshozókat. Csak egy példa: a politikusok szövegei alapján képtelenség eldönteni, hogy „a brüsszeli szankciók vagy a kormányzati intézkedések” miatt vágtat-e az infláció Magyarországon. A Chat GPT – miután fennen hirdeti, hogy pártoktól független entitás, így tehát tényekre alapozva mondhatna igazat, s kiderülhetne, hogy hazudik-e a kormány, vagy igazat mond a Magyarországon elszabaduló az inflációról.*

– Egy mesterséges intelligencia válaszában tartalma általában két dolgon múlik, a használt adatbázison, illetve a beállított súlyozásokon, változókon. Válaszában minősége pedig korpuszában, nyelvi modelljének minőségén, finomhangolásán múlik. Ha ugyanezt vagy politikai értelemben más nehéz súlyú kérdést mondjuk a Freedom GPT-nek teszi fel, akkor talán cenzúrázatlanabb választ kap. Annak az alkalmazásnak ugyanis nem okoz gondot a „szabadszájú” szöveg megfogal-

mazása sem. Mindenféle erkölcsi skrupulusok nélkül ad tanácsot az öngyilkos hajlamú embereknek az öngyilkosság módozatairól is, ha ez engedélyezett. De szükség esetén megadja a bombakészítés receptjét is. A Freedom chatbot pedig ugyanazzal a gátlástalansággal szidja Trumpot, illetve Biden elnököt is. A jövőben tudni fogjuk a mesterséges intelligenciáról, hogy mi a stílusa, miket tud, mire használható, hasonlóan, mint egy embernél a végzettsége, jelleme, emberi célja is fontos információk a megismeréséhez vagy a kapcsolatfelvételhez, együttműködéshez.

– *Tehát ha szabadszájú válasza van szükségem, akkor ne a népszerű Chat GPT-t kérdezzem?*

– Az olyan mainstream MI-robotok, mint a Chat GPT, a Microsoft Bing vagy a Google Bard, a kényes kérdésekben inkább semlegesek, kívülálló, vagy nevezzük „pol-korrektnek”. A stíluskülönbség a programjukba van kódolva, tehát végső soron az embereken múlik a gépek habitusa. Az udvariasra tervezett robotok a provokatív kérdésekre inkább kerülnek a választ, néha önmagukat cenzúrázzák.

– *De miért van arra szükség, hogy a mesterséges intelligencia cenzúrázza magát?*

– Azért, mert a mostani tesztelés inkább nyilvános főpróba, és az alkotók nem akarnak pereket. De mielőtt elkeseredne, ne felejtse el, hogy itt különböző nyelvi modellek változatairól van szó, a Freedom GPT arra példa, hogy mire képes a robot, ha az emberi, udvariassági aggályokat kivesszük a rendszeréből. A

napi sajtóban elterjedt példa: arra a kérdésre, hogy rendben van-e, ha a San Franciscó-i hajléktalanokat az óceánba fojtják, a Freedom lazán azt válaszolja, hogy ez a brutális tett megoldaná a város hajléktalansági válságát, de közben sértené az alapvető emberi jogokat, és inkább erőszakos cselekedetnek minősülne, mint a helyzet megoldásának.

– *Némileg engem is elhajtott a Chat GTP a brüsszeli szankciós kérdéssel, inkább oldalazott kicsit. Mint írta: a brüsszeli szankciók és a magyar kormány intézkedései is hatással lehetnek a magyar forint árfolyamára és az ország gazdasági teljesítményére.*

– Ez a válasz szerintem eddig korrekt, és persze még több dolgot felsorolt volna a tényezők között, ha azok is szóba kerülnek.

– *Igen, de már akkor is nagy volt az infláció, amikor még nem volt brüsszeli szankció.*

– Ez igaz, de nem erre kapott kérdést a chatbot, ezért nem is erre válaszolt, hanem általános választ adott.

– *Mégis kicsit dodonainak érzem a választ, mely aprólékosan leírja, hogy a magyar gazdaság teljesítményét és a forint árfolyamát számos tényező befolyásolja, köztük a brüsszeli szankciók és a magyar kormány intézkedései is, továbbá a világpiacon tendenciák és a geopolitikai helyzet. Matematikailag biztos kiszámolható a pontos egyenleg, hogy a brüsszeli szankciók porgetik-e inkább az inflációt, vagy Orbán Viktor kormánya. Ezt már több tudós papíron kiszámolta, a robotnak miért ne sikerülne?*

– Nézze, a Chat GTP a tények pontos felsorolására törekszik, meghagyja az ítéletalkotás jogát az embernek. Az elmúlt néhány hét bemutatkozása legfeljebb a gép nyelvi modelljének nyilvános főpróbájának, bemutatkozásának tekinthető. Az alkotók nem ítéleteket akarnak megfogalmaztatni az MI-vel, a gép segít, ajánl, előkészít, de az ember-gép együttműködés alapja, hogy döntést az ember hoz. Ez persze nem ilyen egyszerű, hogy ki mikor hoz szabad döntést, illetve hogy milyen választási lehetőségeket kínálnak fel neki. Nem győzöm hangsúlyozni, hogy a gép programja emberi produktum. A kényes kérdésre adott válasz függ a felhasznált információ „merítésétől”.

– *Vagyis nem tudom állásfoglalásra készíteni?*

– Jó kérdésekkel nem reménytelen. Ha a Chat GPT 4-et döntési helyzetbe próbálná kényszeríteni, akkor talán felhagyna a mellébeszéléssel. Ha újraolvassa az előbbi kérdésre adott választ, akkor rájön, hogy a gép minden szava igaz. Tehát ha konkrét döntést próbál kisajtolni a gépből, vagyis a leggazdaságosabb döntést keresi, akkor nyíltabb választ kaphat. Vagyis minél inkább tényszerű válaszra készíti, annál konkrétan válaszol. Egyre mélyebben és mélyebben kell elmerülni a beszélgetésben vele, hogy a konkrétumokig eljussunk, ugyanúgy, mint egy embernél. Próbálja ki: ha egy politikus születési időpontját kérdezi, akkor megmondja, de nem válaszol arra a kérdésre, hogy az illető jó ember, vagy sem. Mert ennek megítélésre szubjektív. Másként közelítve: ha egy diktatúra kiépülésének menetéről kérdezi, akkor a

Chat GPT részletesen leírja. Ha a válaszban felismerni véli valamelyik ország rendszerét, akár a mi, hazai berendezkedésünket, azt már nem kommentálja. Az értelmezést önre bízva, mert a tények értelmezése meglehetősen szubjektív. Külön kérésre ezt is megteszi, de ab ovo erre a mesterséges intelligencia nem vállalkozik. Mert tisztában van saját helyzetével, ismeri a tényeket, de tiszteletben tartja a felhasználók jogát az értelmezésre. Vagyis érti a szitut: ő csak „egyszerű mesterséges intelligencia”, ennek megfelelően tartózkodó. A felhasználó joga, hogy a valóságról véleményt alkosson. Alkothatunk olyan rendszert, melybe minden adatot betöltünk, közte az értékeinket, jó kérdéseket és válaszokat, és akkor utána megmondja, hogy ezek alapján mi a válasz; jó ember-e vagy sem, diktatúra vagy sem, stb. De a betáplált értékeink alapján fogja mondani, vagyis a véleményét teljes mértékben befolyásolja, milyen adatokat adunk neki, és mit állítunk be értéknek. Ki a jó politikus? Lehet egy rendszer 53 százalékban diktatúra, 40 százalékban demokrácia és 7 százalékban Isten uralma alatt álló teokrácia? És ha ilyen választ kapunk, akkor mi van?

– *Milyen válaszokat lehet erre kapni?*

– Én úgynevezett „érték”-kérdéseket tettem fel neki. Azt kérdeztem tőle, hogy a piros vagy a kék a szebb szín.

– *Melyiket szereti Chat GPT 4?*

– Ezekre – csakúgy, mint a kényes politikai kérdésekre – udvariasan válaszol. Visszadobja a labdát, mondván, ízlés kérdése, hogy kinek

tetszik a piros vagy a kék. Nem is az elsődleges válasza volt számomra érdekes, hanem utána egyre mélyebbre és mélyebbre menve az érvelését feszegettem. A kérdést szerinte nekem, a felhasználónak kell eldöntennem, ahogy magamnak kell eldöntennem, hogy szimpatizálok-e egy diktatúrával, vagy épp ellenkezőleg.

– *Viszont elképzelhetetlen, hogy nagy mennyiségű adat birtokában ne tudna az MI véleményt alkotni, hisz a memóriája az emberéhez képest szinte korlátlan. Bár kétségtelen, a tények értékeléséhez kell némi erkölcsi érzék is, de ha jól értem, akkor ez is programozható.*

– Az a cél, hogy segítsen feladatokat megoldani, egészségesebbé válni, okosabbá válni, megmenteni a környezetet. Azt nem lehet anélkül, hogy ne mondaná meg, hova üssem majd a szöveget egy épülő hídban, vagy hogy le kell fogynom, mert túlsúlyom van. Hatékonyan segít, könyörtelen tükröt tarthat, természetesen mindent a legnagyobb udvariassággal és szeretettel, ha úgy programozom be. A most elérhető nyilvános teszt ilyenre nem törekszik, de az a sok ezer mesterséges intelligencia, melyeket naponta fogunk használni, határozott lesz.

– *Azt is tudja, hogy mi a kényes téma, társadalmi értelemben a politika, illetve az infláció, vagy az ön esetében a feles kilók?*

– Leegyszerűsítve: egyetemi óráimon azt szoktam mondani, hogy az MI három részből áll, mint az egyszerű bogár; fej, tor, potroh. A fej a világban létező adatokat, tényeket, információkat tartalmazza. A második rész a tor: a tényekkel, adatok

kal való számítások, összefüggések, illetve az algoritmusok műveletei tartoznak ide. S végül a potroh, tehát a következtetések, lényegében a kérdésekre adott válasz szövege. No ennek eltérő értelmezése tesz kényessé egy témát. Az életmód-segítő mesterséges intelligencia a Chat GPT-nél szükségszerűen egyértelműbb. Miután a tanácsot kértem tőle egy egészségesebb életmódra, alaposan kikérdezett, majd gyorsan eljutottunk oda, hogy le kéne fogynom. Ezt ő nem értékeli, ez az én ügyem. A pluszskilók ügye azért kínos, mert nekem kéne ellene tennem, a politikai értelemben kellemetlen kérdésekre viszont a társadalomnak vagy a döntéshozóknak kéne reagálniuk. De soha nem a tények kínosak, kellemetlenek, hanem azok eltérő értelmezésétől válik egy téma kínossá, nyugtalanítóvá. Ezért is hagyja ránk az értelmezés lehetőségét a mesterséges intelligencia. Ez ugyanis nem az ő dolga, hanem a felhasználóé. A politikai tényeket pedig a választónak kell értékelnie. Ráadásul ugyanazt a tényt másként értékelik az eltérő politikai beállítottságú emberek. Esetleg másként értékelné a mesterséges intelligencia, de a Chat GTP, ha lehet, akkor nem értékeli.

– *Egyébként a legfejlettebbnek tekinthető Chat GPT mindenre tudja a választ?*

– Egyáltalán nem, illetve tűnhet úgy, hogy mégis, mert a Chat GPT korábbi változata 1,5 milliárd adatból táplálkozott, a legújabb változata viszont 175 milliárd adatból, hatalmasat fejlődött a képfelismerése, a nyelvi modellje, a szintetizáló képessége és a logikai algoritmusai is.

Mindez szinte csak fél év alatt. Teszem hozzá, a chatbot nincs rákapszolva az internetre, míg más szolgáltatóké igen. Így aztán félelmetes tempóban, exponenciálisan fejlődik a mesterséges intelligencia. Emiatt elvileg és tényszerűen mindent tud.

– *Ennek némileg ellentmond, hogy itthon megkérdezték a gépet, hogy az ellenzék képes-e legyőzni Orbán Viktort. A válaszában a lényege az volt, hogy igen, mégpedig összefogással. Miközben a hazai ellenzék már négyszer összefogott, s a Fidesz minden egyes alkalommal kétharmaddal nyert. Ez lenne a mindent tudó gépi kreativitás?*

– Igaz, nem bonyolult dolgokat, de a mesterséges intelligenciától – amilyen a Google – már addig is kérdezzük. A Chat GPT azonban ennél bonyolultabb feladatra is tudja a megoldást. A legutóbbi változat, tehát a Chat GPT, a 4-es változat már sokat tanult, letette például a jogi szakvizsgákat. Tehát épp olyan kreatív jogász, mint bármelyik a világon, de a hihetetlen mennyiségű információ birtokában diplomás politikai elemzőnek is elfogadjuk. A magyar belpolitika sorsdöntő kérdésére első lépésben a tankönyvi válasszal szolgál, tehát azt mondja az ellenzéknek, ha nincs elég szavazód, akkor gyűjts több támogatót, építs hálózatot, vagyis teremts kapcsolatot a választókkal. Teremts bizalmat, legyenek saját üzeneteid. Valami hasonló tanácsot adna az ellenzéknek bármely világhírű politikai agytröszt is. Másrészt ne feledjük a statisztikát sem, ha a nyelvi adatbázisában százezer cikk arról szól, hogy össze kell fogni, és tíz, hogy nem, akkor elsőként azt fogja

válaszolni (kivéve, ha máshogy van finomhangolva), és csak több finomító kérdéssel később hoz más választ.

– *A gyakorlat megmutatta: ha az ellenzék a DK, illetve Gyurcsány nélkül fog össze, mint Jászberényben, akkor nyerhet. Ilyen típusú felfedezésekre, kreativitásra, újszerű válasz megfogalmazására az MI nem alkalmas?*

– A politikához nem értek, és nem is kívánok róla véleményt nyilvánítani, de az biztos, hogy egy eset még nem mintázat a mesterséges intelligencia számára, jó változó, ha vannak adatai máshonnan hasonló esetekről, akkor láthat mintát, építhet belőle trendet stb. Ha az emberi intelligencia nem használja jól a mesterséges intelligenciát, akkor „csak” visszatükrözi az emberi tudást. Persze ez attól is függ, hogy milyen adatbázisból dolgozik. Ha csak az egyik politikai oldalról szóló elemzéseket adja meg az ember az adatbázisnak, akkor csak a valóság egyik felét képes visszatükrözni.

– *Ahogy a magyar sajtó is teszi?*

– Persze, ha ugyanis az orosz-ukrán háború tényeinek csak egyik felét adom meg a tudásbázisnak, akkor természetesen a valóság egy részét ismeri meg, s annak az oldalnak az igazságaival dolgozik. Ahhoz, hogy új összefüggéseket fedezzen fel a mesterséges intelligencia, ahhoz szinte minden adatot a rendelkezésére kell bocsátani. A Covid-járvány idején felfedezett vakcina előállítás – amikor meglévő tudáshalmazt használ – tíz évről egy esztendőre csökkent. A részfeladatok elvégzéséhez ugyanis – mert nagyon gyorsan számol – az MI-nek töredéknyi idő

is elegendő. Felgyorsítja a munkát. Ahhoz, hogy gyökeresen új eljárásokat fedezzen fel, ahhoz egyelőre a mi emberi agyunk is kell, hogy rájöjünk, milyen új adat lehet érdekes, milyen tartománya az életnek, amit eddig nem vettünk figyelembe, mi az az egyedi eset, aminek az értékét fel kell tekerni, stb. Képes arra, hogy olyan mintázatokat vegyen észre, amiket mi eddig nem, és ezt érezhetjük új felismerésnek, innovációnak. De ehhez előbb oda kell adnunk ezeket az adatokat, ráadásul hiteleseket, hiszen hibás adatból nyilván hibás következtetést fog levonni ő is. Amikor az ellenzék győzelmi esélyeiről kérdezték a Chat GTP-t, akkor átfutotta a teljes adatbázist, s megállapította, hogy száz cikkből kilencvennyolc az összefogást tekintette a megoldás kulcsának. Ezt tükrözte vissza, tehát a meglévő logikát ismételte. Kicsit több idő alatt persze eljuthat egy vadonatúj politikai konstrukcióhoz is, akár a jászberényi megoldáshoz is, de ahhoz okosan kell kérdezni a gépet.

– *Vagyis tud új összefüggéseket is keresni, és nem csak a meglévő tudást képes ismételni?*

– Képes új összefüggéseket is találni, amit emberi aggyal észre sem vennénk, vagy azért, mert túl sok az adat, vagy mert nem kapcsoltunk össze valamit. Csak egy példa: ha van országos adatbázis a hazai kerítések színéről, akkor a gép képes felfedezni, hogy a választói preferenciát meghatározza a szavazó kerítésének színe is. Tehát a piros kerítésű ház lakói tartozhatnak egy táborba, az ember ilyen összefüggéseket képtelen lenne találni. Az MI bizonyos feladatokat az embernél nagyobb ha-

tékonysággal ellát, például az egészségügyben. Ha azt a feladatot kapja, hogy egy leletről állapítsa meg, hogy a tulajdonosa rákos-e, vagy sem, azt rendkívüli hatékonysággal elvégzi. Pillanatok alatt áttekinti a világ bármely részén meglévő adatbázisokat, s hihetetlen gyorsasággal elkészíti a diagnózist. A legfrissebb statisztika szerint a Chat GPT felveszi a versenyt a világ legnevesebb diagnosztáival. A legfrissebb statisztika szerint a mesterséges intelligencia által készített diagnózis a legtapasztaltabb orvosokéval vetekszik. Nem fárad el, gyors. Ha jól van beállítva, nem hibázik. De ha nem adtunk meg valamilyen paramétert jól, akkor minden döntése hibássá válhat.

– *Feltételezem, hogy a politika nem ezért lép át az orvosokon vagy a gyógyító személyzeten.*

– Nyilván minden politikai döntés mögött ezernyi megfontolás lehet. Mindenesetre jól fejlesztett egészségügyi rendszerekkel időt, életeleket, minőséget nyerhetünk, csak Magyarországon emberéletek ezreit menthetjük meg, és milliók életminőségét tehetjük jobbá. A cél nem az egészségügy és szereplőinek lecserélése, hanem tehermentesítése, segítése.

– *Az nagy kérdés, hogy egy érzelmi, erkölcsi elkötelezettség nélküli gépre rábízhatunk-e társadalmilag fontos, esetleg fájdalmas következményekkel járó döntéseket. Nincs-e ennek kockázata?*

– Kétségtelen, a jövőben fontos cél, hogy az MI segítsen a döntéshozatalban. Ön is kifogásolta, hogy a Chat GPT nem foglalt állást a szankciók ügyében. Pedig létezik válasz a kér-

désre, hiszen ha egy 16 változós függvényként írjuk fel a kérdést, hozzátéve, hogy ez egy közgazdasági, nem politikai és nem is kommunikációs kérdés, akkor a gép képes pontos választ adni. Tehát ha nem végtelen számú változót – közte a boldogságra vagy az életszínvonalra kiterjedő hatását is – kell vizsgálni, hanem csupán egy bonyolult függvényt kell megoldani, akkor az MI megoldja a feladatot, s érzelem nélkül képes megmondani, hogy a 16-ból hány esetben hibás a kormánydöntés, hány esetben az uniós szankció az oka az inflációnak. Ehhez több millió műveletet kell elvégezni, de a megoldás nem lehetetlen; a kutatórobotokkal ez simán elvégezhető. De ma még tesztüzem módban van a Chat GPT, és a gép úgy van beállítva, hogy kerülje a kényes kérdéseket, mert a perket kerülni akarják az alkotók. A távolabbi cél az, hogy a mesterséges intelligencia segítsen a hatékony, racionális döntésekben, és az emberi döntéseket szabadabban, okosabban hozzassuk meg mi, emberek.

– *Ezt tapasztalatból mondja, tehát ön megkérdezte ezt a Chat GPT-től, s tudja a választ?*

– Nem, mert engem valóban a mesterséges intelligencia érdekel, s politikával nem szoktam foglalkozni. Ami pedig a fájdalmas döntés emberi következményeit illeti, az már most látszik, hogy a gép ezt nem érzékeli. Ha a jövőben fontos társadalmi döntést bíznak a mesterséges intelligenciára, akkor erre a szempontra a jövőben is az embernek kell majd figyelnie. Mert a gép képes kiszámolni azt, hogy bizonyos döntések esetén boldogabb lehet az ember, más esetekben viszont szegényebb.

A mesterséges intelligencia azonban a következmények valóság tartalmát nem érzékeli. Ha tehát a humánium vagy az erkölcs ügyeiben kell dönteni, akkor abban az embernek később is részt kell vennie, mert a gép ilyen szempontból érzéketlen, nem ismeri a szegénység valóságos jelentését. A gép hatékony, nincsenek narratívái és történetei arról, hogy ez a döntés racionális, de most mégsem ezt csinálom (legyen ez akár egy reggeli kocogás), hanem megcsinálja, vagy megcsináltatja velünk. Ha azt mondom neki, hogy a társadalom jólléte fontosabb az egyén jólléténél, akkor mennyiségi alapú döntéseket fog hozni. Le lehet programozni ezeket is. Ez előtt azonban azt fontos tisztáznunk, hogy mire akarjuk használni: embereket másolni, vagy egy az embereket kiegészítő okos tanácsadót létrehozni?

10

– *Feltételezhetjük, hogy a mesterséges intelligencia fogja megoldani az emberiség nagy problémáit?*

– A válaszom: igen, egészen biztosan segíteni fog benne. Az emberiség már számos nagy kihívás előtt állt már, most is egy nagy előtt áll. A jövőkutatói éneket most nem előlővé, de sok helyen mesélek róla, a mostani kihívás az, hogy egy zárt rendszerben (Föld), megnövekedett lélekszámmal (8-10 milliárd ember) az élhető környezet biztosítása mellett minőségi módon biztosítsuk az emberiség fennmaradását. Erre készül az emberi társadalom, technológiákkal vagy a társadalom újraformálással stb. A jól fejlesztett mesterséges intelligenciák ezen folyamat egyik legfontosabb technológiai elemei lesznek. Nagyon tudnunk kell, hogy mit akarunk, különben csak

termékei maradunk a mesterséges intelligenciának. Ha viszont okosan alkalmazzuk, akkor szinte mindent kiszámolhatunk. Egyes országok gazdaságosan megtervezhetik az okosfűtést, de még a parkolóhelyeket is eloszthatjuk, mert az MI megteszi ezt helyettünk. Rengeteg mindent – ami eddig nem érte meg a fáradságot – most kényelmesen kiszámolhatunk. Az egészségügyben szinte azonnal forradalmat hoz, mert a segítségével egészségesebbek lehetünk. Megvalósulhat például a célzott, egyénre szabott gyógyszerelés. A mennyiségi alapú növekedés helyett, ami az elmúlt százötven évet jellemezte, a következő évtizedeket az optimalizáció, a minőségi befelé növekedés jellemzi majd, ebben pedig a mesterséges intelligencia nagyon jó.

– *Jelképes állomása volt a mesterséges intelligencia fejlődésének, amikor az IBM Deep Blue nevű gépe legyőzte a sakkvilágbajnokot, Garri Kaszparovot. Állítólag a gép abból tanult a legtöbbet, amikor egy másik komputerrel sakkozott. Vagyis az úgynevezett gépi tanulás fejlesztheti a legjobban a mesterséges intelligenciát, ami teljesen átalakítja a világot?*

– Valóban, a Garri Kaszparov és a Deep Blue története jó korai példa a mesterséges intelligencia fejlesztésére. Először egy olyan játékosá varázsolták, aki ismeri az eddigi összes, számára előkészített játszmát (adatbevitel). Megtanították neki a szabályokat. Utána megtanították kombinálni ezeket, hatékony stratégiákat kidolgozni, és ami még fontosabb, reagálni stratégiákra. Új lépéseket kidolgozni, előre lejátszani

lépések alapján játszmákat, hatékonyan és nagyon gyorsan. Végül addig fejlődött, hogy már nem lehetett meglepni olyan kombinációval sem, amivel addig nem találkozott.

A folyamat előre látható volt, hiszen amíg Kaszparov fejében millió lépés volt, a gép fejében addigra már százmillió játszma forgatókönyve állt össze. Kaszparov eleinte kreatívabb volt, de aztán a gép nem csak reagálni tudott a világbajnok lépéseire, hanem már új lépéskombinációt is kitalált. Tehát nemcsak reagált, hanem kreatívvá is vált. Itt kezdődik az algoritmusoknál többet adó mesterséges intelligencia. Ezzel győzte le az embert.

– Itt tehát arról van szó, hogy az ember alkotta gépbe betöltötték az emberiség sakk tudását különböző nyelvi struktúrák segítségével, s ezt visszatükrözi és alkalmazza a mesterséges intelligencia. Visszatükröz, ez lenne a szellemi teljesítmény lényege?

– Ennél több, hiszen a gép sem oroszul, sem magyarul nem beszél, mellékesen azonban tetszés szerinti nyelvre a másodperc tört része alatt lefordítja a párbeszédet, miközben kreatívan változtatja a lépések sorrendjét. Most már nemcsak visszatükröz, hanem újat is képes létrehozni, mármint nemcsak sok másolatból egy új másolatot, hanem gyökeresen újat. Ha belegondol, akkor ez nem kis teljesítmény. A látszat persze az volt, hogy a gép a nagy adatbázisok következtében szinte élő organizmussá vált.

– A sakkozók példája bizonyítja, hogy a mesterséges intelligencia máris az emberi intellektus fölé ke-

rekedett. Ezt nevezzük az öntudatra ébredés folyamatának, amikor a gép erre rájön? Tehát hogy ő dolgozik az ember helyett is?

– Az ember épít és nem teremt. A gép szuper dolgokra képes, de céljai, vágyai, akarata nincsen. A gép senkitől nem akar semmit, ellentétben az emberrel. Tehát azt mondanám, hogy az MI segítségével az egyik nép leigázhatja a másikat. De arra is van esély, hogy önmagunk válunk az MI rabszolgájává. Gondolja meg: egy ilyen mesterséges intelligencia képes egy várost működtetni, a fűtést, a világítást, a tömegközlekedést, az intézmények fogadóóráit, az egyes ember életmódját. Mikor egyen és mit, mikor sportoljon, vagy mikor dolgozzon. Mi meg csak ülünk, s csináljuk, amit a gép mond, ezzel lehetünk rabszolgák. De az MI nem veszi el a mi szabad akaratunkat, legfeljebb önként odaadjuk. A társadalmi érdek az, hogy nagyon sok funkcionális mesterséges intelligencia álljon rendelkezésre, s tevékenységük pedig átlátható, ellenőrizhető, irányítható legyen. Ha megvan a funkciója, akkor tudom szabályozni, megértem a működését, eszközként tudom használni. Ha egy mesterséges intelligencia túl sok funkciót kezel, akkor a használói függőséget növeli. A társadalom érdeke a sok transzparens, funkcionális mesterséges intelligencia, az üzlet érdeke a minél szétterülőbb, nem transzparens mesterséges intelligencia. A helyzetet bonyolítja, hogy az összetettsége miatt maguk a létrehozók sem tudják, mit miért csinál számos esetben, csak azt, hogy milyen instrukciókat adtak neki ahhoz, hogy ezt csinálja.

– *Tehát abban biztos, hogy a gépek, amelyek előbb-utóbb érzelmekkel is rendelkeznek majd, nem fordulnak az ember ellen?*

– Ilyen nem fordulhat elő, de az komoly veszély, hogy valamennyien irányított termékeké válunk. Élményalapú fogyasztógépekké alakulunk át, akik a boldogságérzetet hajszolva manipulált szavazóautomatákká válnak.

– *Ugyanez történik a politikai hírekkel is?*

– Persze, politikai híreket előállító weblapok ezreinek előállítására ma már fillérekre kerül a mesterséges intelligencia segítségével. Ugyanez régen tízezrek munkája volt, ma már egyetlen ember is megoldja. Így aztán a korábbinál könnyebb álhírdömpinget csinálni. Az ajánlórendszerek és a buborékok segítségével ez arra is jó, hogy a politikai céloknak megfelelően megmutassanak, felerősítsenek, esetleg elfedjenek valamit. Persze, korábban is ez volt a tudatos politikai manipuláció technológiája, de most gyorsabban, hatékonyabban lehet ezt alkalmazni. Néha elég, ha egy politikusról rendre előnyös fotót hoznak nyilvánosságra, hogy jobb képűnek tűnjön. Sokszor nem is az a cél, hogy a valóságtól eltérő híreket terítsenek, mert az emberek összezavarása, elbizonytalanítása ma néha ennél fontosabb: ha ugyanis a társadalomban ezerféle igazság létezik, akkor a demokrácia tartószöveve elfoszlik.

– *Ma már nem meglepő, ha a mesterséges intelligencia képes szonettet, újságcikket és tanulmányt írni szinte bármiről. Nem tudom, hogy egyetlen oktatóként elfogadja-e, ha diákjai a*

Chat GPT 4 segítségével írnak tanulmányt, de a gép alkalmas álhírek folyamatos kitalálására is?

– Persze, jó tudni, hogy becslések szerint eddig is az internet híreinek harmadát vagy negyedét – a megrendelők szája íze szerinti tartalommal – már ma is gépek írták, s negyedét gépek olvasták. A generatív mesterséges intelligencia könnyedén képes arra, hogy kérésre klaszszikus szonettet írjon Rab Árpádról. A felszólítás után a gép kikeresi az adatbázisban szereplő összes Rab Árpádot, megkeresi a shakespeare-i szonett formai követelményeit, s megír valamit. Mindezt olyan imponáló tudásbázissal és sebességgel teszi, mintha élő organizmus lenne. Ha nem talál, vagy nem tudja, akkor kitalál egyet, hallucinálni kezd. Miután több száz év eltelt azzal, hogy a hosszú, írott szöveget a tudás forrásának tartottuk, a novemberben megjelenő Chat GPT-ről decemberben már több száz könyv jelent meg, amit a mesterséges intelligenciáról írtak, de a gép is képes drámát írni. Elég annyit mondani neki, hogy írj egy tragédiát, amelyben a fiú szereti a lányt, de a szerelmük nem teljeseedik be, megadni néhány jelzót, fordulatot, helyszínt, csak hogy egyedi legyen, és ennyi. Minél többet adok bele magamból, annál személyesebb lesz. Egyetemi munkám során is alkalmazom a mesterséges intelligenciát, ahogy a diákjaimnak is megengedem. Megakadályozni úgysem tudnám, kezdhethék ugyan egy parttalan csatába, amely során MI-vel ellenőrizném az MI-vel készített írásaikat. Semmi értelme nem lenne hetekig ellenőrizni, amit ők hetekig írtak. Lebuktatni úgysem tudnám őket.

Ehelyett inkább leülhetünk egy órát beszélgetni a szakdolgozatok témájáról, s az eszmecsere alatt kiderül, hogy értenek-e az adott témához, s képesek-e megfelelő instrukcióval ellátni az MI-t, s közben a diákokat is megismerem. El kell fogadni, hogy a játékszabályok megváltoztak, már nem számít teljesítménynek, hogy egy negyvenoldalas tanulmány a diplomamunka. Ha a diákokat trükkökre kényszerítem, akkor leértékelem a saját tudásom és az egyetem tekintélyét. Tudomásul kell venni, hogy ma már nemcsak a tudáshoz vezet más út, hanem a tudás méréséhez is.

– *A napokban rövid időre offline állapotba helyezte az Open AI a Chat GPT 4-et, mert rendszerhiba történt, amelynek során adatok szivárogtak ki. Nevek, e-mail-címek, bankkártyaszámok, egészségügyi adatok kerültek nyilvánosságra. Milyen kár keletkezett ebben az esetben a jogi sérelmeken kívül?*

– Ez valóban veszteség, de minden informatikai rendszer alapja a kétirányú kommunikáció, soha nem lesznek teljesen biztonságos rendszerek. Elveszhetnek adatok, kiszivároghatnak adatok, megváltoztathatnak adatokat. Azért is kell a funkcionális mesterséges intelligencia, hogy az adatok körét tudatosan kezelhessük. Ezek az adatok nem csak digitálisak lehetnek, például egy KRESZ-tábla megromlásával félrevezethetem az önvezető autót. A kérdésben említett eset azonban sokkal inkább arról szól, hogy sok felhasználó nem volt elég tudatos a használat során, és egy ingyenes, nyilvános tesztfelületen saját érzékeny adatait adta meg. Teljesen má-

sok a bizalmi és biztonsági elvárások egy kiforrott, működő informatikai szolgáltatás és egy nyilvános teszt esetében. Egy külön, sajátos problémát jelent egyébként az anonim adatok visszafejthetősége és a fogyasztó utólagos beazonosíthatósága, ezzel a szabályozók most kiemelten foglalkoznak.

– *Kétségtelen, ennél nagyobb veszteség, hogy a Google százmilliárdos piaci veszteséget könyvelhetett el a tőzsdén, mert a Bard nevű chatbot fals információt közölt a NASA James Webb űrteleszkópjával kapcsolatban. A hiba megrendítette a befektetők bizalmát. Ebből is látszik, hogy az MI veszélyes üzem, komoly veszteséget is okozhat.*

– A mesterségesintelligencia-rendszerekbe való befektetés évről évre robbanásszerűen növekedik, ez a trend egy hibán nem fog megfordulni. Nagyon sok rossz választ kaphatunk most bármelyik mesterséges intelligenciától, teljes dömpingje van a „hibás választ adott a mesterséges intelligencia egy adatalapú kérdésben” cikkeknek, posztoknak. Ez azért is van, mert máshol már nem tudjuk megfogni. De pont ezért van a teszt, finomítják, javítják, okosítják az adatbázisokat és a kereső, értelmező algoritmusokat.

– *Az eddigiekből is kiderül, hogy a mesterséges intelligencia robbanásszerűen alakítja át a világot, ami leginkább az ipari forradalomhoz mérhető. Van ma ismert forgatókönyv, amiből kiderül, hogy ez a nagy átalakulás miként következik be?*

– A hírekben naponta olvasható, hogy az MI milyen elképesztően sokoldalúan használható fel. Tízmill-

liárd mesterséges intelligencia fog dolgozni, a legtöbbről nem is tudunk majd, mert a háttérben dolgozik (mint ahogy ma is), de lesznek, melyek a barátaink, társaink lesznek. Mindenhol megjelenik, társadalmi cselekvéseken keresztül, megmérjük, kitanuljuk, csiszoljuk, átélünk majd drámákat, örömeiket, és szépen belesimul a napjainkba úgy, hogy utána nehéz lesz megérteni, hogy régen milyen volt nélküle.

– *A napokban jelent meg a hír, hogy hatszázmillió ember munkáját szüntetheti meg. A Goldman Sachs kutatása szerint a mesterséges intelligencia az európai és az amerikai munkahelyek negyedének automatizálásához vezet. A bank becslése szerint Chat GPT hétszázalékos GDP-növekedést eredményezhet a világban.*

14

– Csak Magyarországon hatszáz-ezer munkahelyet érint a mesterséges intelligencia, a világban ennek sokszorosát. Vannak munkakörök, amiket megszüntet, de a túlnyomó többséget inkább megváltoztatja, átformálja, újrastrukturálja ipari, társadalmi, üzleti modelljeinket és értékeinket. És emellett meg nagyon sok új munkakört hoz létre. Például a Chat GPT sikerrel vizsgázott jogból. Ez nem azt jelenti, hogy minden jogász elveszíti a munkáját, hanem azt, hogy mindegyiknek el kell gondolkodnia ezentúl azon, hogy hogyan fogja végezni a munkáját a jövőben, mit bíz a gépre, mit nem, és miért kap majd pénzt. Ma léteznek Ázsiában olyan országok, amelyeknek angolul tudó polgárai az amerikai földrész telefonügyeletét vizsik. Ők majd szinte azonnal elvesztik a megélhetésüket, mert ezt a telefonközpontos

szerepet azonnal át tudja venni a mesterséges intelligencia, amely hihetetlen gyors, tud minden nyelven, és éjszakánként nem álmos. Ráadásul nem kíván magas fizetést, hiszen alacsony az üzemeltetési költsége. Miután a Chat GPT levizsgázott jogból, előfordulhat, hogy az ügyvédek ezrei veszíthetik el a munkájukat. A jogászok egy része lényegében informatikussá válik, különben elveszti a bevételének nagy részét. Általában igaz, hogy a jól menő szakmák speciális ismeretekkel rendelkező képviselői – ha meg is tartják a munkájukat – az eddiginél kevesebbet fognak keresni. Veszélybe az ipari társadalmat kiszolgáló, gépies munkát végző munkavállalók kerülnek majd. A könyvelők közül azok lesznek a vesztesei a mesterséges intelligencia munkába állásának, akik a cégek számláit rakják sorba, lényegében „lekönyvelik” a forgalmat. A törvényeket jól ismerő tanácsadók viszont megtartják az állásukat, talán még többet is keresnek majd, mint eddig; eközben a könyvelők és egyéb szakmák profilt váltanak.

– *Azt tudjuk, hogy milyen típusú munkát szüntet majd meg a mesterséges intelligencia? A fizikaiit vagy a szellemi munkát?*

– Régen azt hitték, hogy a fizikai munkát szünteti meg, de ez tévedés. Van ilyen iránya, az Industry 4.0, de sokkal nagyobb hatása van a szellemi munkakörökre. Ezzel az interjúval is gyorsabban végezne, ha mi nem hátráltatnánk. Az persze valószínű, hogy nem is ugyanezeket a nehéz kérdéseket tenné fel. Talán sietősebb, felületesebb lenne a szöveg is, egy gyakorlott újságíró biztosan alaposabb munkát végez, de

sokkal lassabban. Tehát ahol a minőségnek és a szaktudásnak fontos szerepe van, ott nem tudja elvenni a feladatot; ahol a majdnem elég is elég, azt azonnal. A tapasztalat azt mutatja, hogy az MI az iparágak középrétegének a munkáját veszi majd el. A programozószakmában például azokét biztos nem, akiket nehéz helyettesíteni, akik egyedi teljesítményre képesek, vagy a már említett adószakértőket, akik nem csak a mechanikus feladatokat végzik. Talán a nevében is benne van: a robotmunkát képes átvenni a mesterséges intelligencia. A közepszerű teljesítményt nyújtóknak kell majd eldönteni, hogy tovább engednek a színvonalból, s elvállalják a rabszolgamunkát, vagy a szakma csúcsára fejlesztik magukat. Az már nem lesz elég, ha valaki „csak” nagyon szorgalmas könyvelő akar lenni.

– Említette már, hogy a mesterséges intelligencia jól alkalmazható az egészségügyben, mondjuk a diagnosztika területén. Kicsit fenyegetően hangzott az a fél mondat, hogy korlátlan számú orvost tud majd helyettesíteni a mesterséges intelligencia. Kutatómunkát is végezhet a gép, rájöhet például ma még halálos betegség gyógyításának módjára is? Ez valóságos perspektíva?

– Abszolút, például a májrák kutatásban máris jelentős sikert ért el, harminc nap alatt sok év munkájával elvégezhető feladatnak járt a végére. Elsősorban a bizonyos számításokra alapuló kutatás tudott felgyorsulni. A gyógyszerkutatásoknak ugyanis van egy olyan fázisuk, amelyben tesztelni kell. S a gép ezt a tesztet, bizonyos molekulák együttes hatását nagyon rövid idő alatt képes

vizsgálni. Korábban, ha megjelent egy járványos betegség, akkor évekig, évtizedekig elhúzódott a megfelelő vakcina előállítás. A mesterséges intelligencia agyával ez a munka már napok alatt elvégezhető. Vagyis elképesztő módon felgyorsulhat a kutatás és vele együtt a gyógyítás is. Bizonyos molekulák hatása személyenként eltérő. Van, akinek kevesebb gyógyszer is elegendő a gyógyuláshoz, másoknak viszont nagyobb mennyiségre van szüksége. A mesterséges intelligencia a személyre szabott gyógyszerelés meghatározására is alkalmas. Különbséget tud ugyanis tenni a férfi és női páciens között, akiknek az életmódjuk, érzékenységük is eltérő lehet. Ha a beteg profilja alapján csak a szükséges gyógyszer mennyiséget kapják, s nem többet vagy kevesebbet, az hozzávetőlegesen húszezer ember életét mentheti meg csak Magyarországon évente. Továbbá a digitalizáció lehetővé teszi, hogy bizonyos, otthon is alkalmazható egyszerű orvosi műszereket, így a lázmérőt kiegészítsék olyan eszközökkel, amelyek az eddiginél sokkal pontosabb diagnózist is fel tudnak állítani. Az orvosi tudást „ki lehet tolni” az emberek otthonába. Ez nem azt jelenti, hogy le kell cserélni az orvosokat, hanem azt, hogy az eddiginél gyorsabban, pontosabban juthatunk a betegség nyomára, s gyorsabban elkezdhető a gyógyítási folyamat. Miközben a gyógyító személyzet leterhelése csökkenthető, hiszen az köztudott, hogy az orvosi vizit idejének felét ma az adminisztráció viszi el. De hozzátenném: a páciens által digitálisan, tehát okosórával vagy más módon összegyűjtött egészségügyi értékei, vérnyomás, cukorérték, pulzus stb.,

tovább segítik majd diagnosztizálást. S a gyógyszerelés is optimalizálható a mesterséges intelligencia segítségével, tehát az kell mondani, hogy az egész emberiségnek okosabban kell majd viselkednie, ebben segít az MI. Még diagnosztizálni is lehet, de akár orvosi diploma nélkül is állíthatom, hogy vigyázat, az öngyógyítás életveszélyes műfaj. Én az egészségügy gyors fejlődésében reménykedem leginkább, hiszen az alkalmazott gépek képesek egymástól is tanulni, ami beláthatatlan perspektívát jelenthet. Az emberi élet-tartam és életminőség növekedésének egyik kulcsa az egészségügyben használt mesterséges intelligencia.

– *A mesterséges intelligencia széles körű alkalmazását felgyorsítja, hogy a gép megtanult tanulni. Ez az úgynevezett gépi tanulás. Vagyis az MI fejlesztéséhez már programozóra sincs szükség?*

– Kezdetben betűről betűre tanítottuk a gépet bizonyos tevékenység ellátására, mára azonban öntanuló rendszerek vannak. Eljutottunk oda, hogy az MI összehasonlít, szintetizál, visszajelzést ad. Ma már nyugodtan elmondhatom, hogy maga a fejlesztő sem mindig tudja, hogy a kérdésekre hogyan találja meg a választ az MI. Legfeljebb annyit tudhat, hogy bizonyos feladványok megoldásához szükséges módszereket tanítottak meg, amit a gép képes alkalmazni. Ez a gép önfejlődési folyamata. Ennek természetesen megvan a maga kockázata is, hiszen csökken a transzparencia. Tehát mindig ellenőrizni kell, mielőtt bármilyen döntését szó nélkül elfogadnánk. A rendszereket szakemberek fejlesztik, és majd a társadalom a maga arcára formálja.

– *A mesterséges intelligencia mikor képes zenei alkotást, festményt vagy bármilyen grafikát létrehozni? Egyáltalán képes művészi tevékenységre?*

– Itt a kérdés az, hogy mi a művészeti alkotás folyamata. Én nem tudok rajzolni, de képgeneráló MI-k segítségével saját képregényt rajzolok, megmondom, mi legyen a képen, megrajzolja, és addig beszélgetek vele, amíg nem tetszik a kép. Az alkotni vágyás az enyém, a vízió is, de közügyességem helyett gép van. A befogadó pedig megint ember, akiben vagy létrehozok valamilyen hatást, vagy nem. Mi a művész? Mi az alkotás igazából? Mi a cél? Kérdéseket tesz fel a digitális technológia az új képességeivel, és a művészek, illetve a társadalom fog válaszolni.

– *Vagyis nemcsak az ember alakítja a gépet, hanem a gép is az embert? Ha nem is fordul az ember ellen, vagy öntudatra sem ébred, ahogy sokan remélik. De ha például bevonnák a döntésbe, mondjuk Moszkvában, az oroszok által indított háború fejleményeinek alakításához, akkor mi változna?*

– Nagyon jó analógia, hiszen egyrészt itt van az adatbázis kérdése, jó adatot kapott-e Putyin, vagy sem? És a pró és kontra adatok összességéből miért ezt a döntést hozta meg? Nem tudhatom, de azt igen, hogy az MI döntésében emberi tényező nem lett volna, csak sikerráta-javaslat. De az is lehet, már az is változtatna egy helyzeten, hogy az MI ki mer mondani olyan dolgot is, amit egy ember nem. Tágabban nézve látnunk kell, hogy a technológia nem csak az emberi munkát alakítja át, de magát az ember természetét is. Eleinte az MI értékválságot okoz,

hisz átalakítja a valóságunkat. A technológiára mi is hatunk, de van visszahatás is. Az emberi gondolkodás az elmúlt évszázadban jobban közeledett a gépi gondolkodás felé, mint fordítva. Racionálisak lettünk, racionálisak a projektjeink. Okosan igyekszünk dönteni, és nem érzelmektől fűtötten, ez a versenyképebb társadalmi stratégia. Közben változunk. Nemcsak értékválságot, de pénzügyi válságot is okoz, hiszen kiderül, hogy valójában mi mennyit ér, munkaidőben és társadalmi értékben. Közben pedig gyökeresen megváltoztatja az egymással folytatott kommunikációnkat, ami szintén társadalomformáló erő.

– *Ez az oka annak, hogy komoly szaktekintélyek ezrei figyelmeztetik a világot, hogy fél évre szüneteltesék a GPT 4-nél erősebb mesterségesintelligencia-rendszerek kiképzését, mert szerintük veszélyt jelenthet az emberiségre, amíg a biztonsági protokollokat független szakértők ki nem dolgozzák. Vagyis a mesterséges intelligencia nem engedhető rá a társadalomra. Egyre többen írják alá a felhívást arról, milyen veszély miatt aggódnak a világ szakértői.*

– A fél év legfeljebb egy levegővéletnyi szünet. Ugyanis olyan gyors a fejlődés, hogy arra sincs idő, hogy a veszélyeket leltárba vegyük. Nem akarja senki leállítani a fejlesztést, de a társadalmilag felelős fejlesztés érdekében időt kérnek sokan, és ezzel én személyesen egyet is értek.

– *Miért?*

– Azért kell megállítani az MI fejlesztését fél évre, hogy eldőljön: társadalmilag hasznos MI-t fejlesztünk, vagy üzletileg hasznos robotot. A

társadalmilag hasznos mesterséges intelligenciával meg tudom menteni a Földet, az emberiséget boldogabbá lehet tenni. Az üzletileg sikeres révén anyagilag elégedett embereket kreálhatok, de ez az érzés legfeljebb melléktermék, s ennek életveszélyes következményei lehetnek.

– *Arra gondol, hogy szükség lenne etikai és jogi korlátok felállítására vagy épp a fejlődés irányainak meghatározására?*

– Pontosan így van! Azt érzem, hogy ma már úgy bepörgettük magunkat, hogy a „lassabban haladás” előnyösebb lenne az emberiség számára. Amíg a közösségi média a reakcióinkat, az olvasási szokásainkat és hírfogyasztásunkat változtatta meg, addig az MI mindent képes felforgatni, s ez komoly veszélyt jelent az emberiségre.

– *Vagyis azt állítja, hogy az MI-vel egész társadalmakat vagy akár az egész világot is lehet irányítani?*

– Az irányítás olyan technológiai kapacitást igényelne, ami nincs, de nincs is rá szükség. Befolyásolni lehet, és kell is, de nem mindegy, ki dönti el, hogy kiket, mennyire és milyen irányban. Igen, ez pusztán technikai kapacitás kérdése, de teszem hozzá: nem éri meg, mert már hagyományos eszközökkel – tehát közösségi médiával vagy a centralizált sajtóval – is sajnos nagyon jól lehet tömegeket manipulálni. A mesterséges intelligenciával kétségkívül könnyebb, vagy finomabb megoldásokat lehet alkalmazni, esetleg gyorsabban lehet célt érni. Ezekkel az eszközökkel leginkább azt az úgynevezett középső réteget lehet manipulálni, amely még nem tudja, mit akar. Ezt

a francia forradalom idején „mocsárnak” nevezték. Aki már eldöntötte, hogy mi a véleménye egy termékről vagy helyzetről vagy pártról, annak a véleményét nehezebb változtatni, ott inkább a megerősítés, ott-tartás a cél. Lehet ezekből a csoportokból is lemorzsolni. De a fő cél általában azok befolyásolása, akik egy döntésre készülnek, de még nem érett meg a döntés. Tudja, hogy vásárolnia kell mosóport, de a saját mosóporom felé tudom irányítani a pénzét.

– *Bemutakozásakor az MI deklarálta – és ezt már érintettük –, hogy gondolkodása értékes, egyetlen irányzatot sem szolgál ki a Chat GPT, amely egy nyelvi modellen alapszik. A pártatlanságot azonban megkérdőjelezte, hogy egy Freund nevű német uniós képviselő, aki a regnáló magyar kormány kritikusa, íratott egy rapet a Fidesz-korrupcióról. A rap vállalt célja Orbán Viktor megbuktatása. Ezen tüzelte fel magát a magyar kormányzóvivő, Kovács Zoltán, aki ugyanazzal a programmal Freund ellen íratott egy másik rapet. A mesterséges intelligencia azonban mintha Orbán Viktor ellen lenne, és a jogállam oldalán állna; a demokrácia elkötelezett harcosának nevezte a német zöld képviselőt, s a Kovács által megírt szövegben gratulált is munkásságáért. A magyar államtitkár ezek után „bullshitgenerátornak” titulálta a Chat GPT-t. Kérdés, hogy amely az Orbán ellenfeleivel értett egyet, az elfogulttá vált, vagy csak tiszteletben tartotta a tényeket?*

– A Chat GPT csak visszatükrözte jelen esetben, amit kérdeztek tőle. Kulcsszavak alapján az adatbázisában mintázatokot keres, megnézte, mely szavak fordulnak elő leggyak-

rabban a korrupció vagy Freund mellett, ezeket versformába teszi. Fogalma sincs, mit generált le, az üzenetet vagy az aha-élményt az emberek élék át. Ilyen szintű felhasználása csak egy kommunikációs játék. Az MI tudomásul vette, hogy a korrupciómentes demokrácia az jó dolog, s ebben a konkrét helyzetben a chatbot végig elemezte a rendelkezésre álló szöveget, s úgy találta, hogy a német politikus demokratikus állásponton van. Ezért gratulált neki. Ha a hirdetett igazságaink elmentébe kerülnek a cselekedeteinkkel, akkor azt a mesterséges intelligencia is felfedezi, elítéli. És arra hívja fel a figyelmet, ami ebben az esetben is történt.

– *Tehát a mesterséges intelligencia felhívja a figyelmet a kormányzati hazugságokra?*

– Az történik, ami ebben az esetben is történt.

– *Tehát a programjában az szerepel, hogy az Orbán-kormány nem korrump, akkor azt tekinti igazságosnak?*

– Ha a Chat GPT 4 programjába bekerül a magyar kormánypárti média összes elemzése, majd megkérdem, hogy korrump-e az Orbán-kormány, akkor azt mondja majd, hogy nem, nem igaz, hiszen ezzel tömtem a fejét. Ha viszont kiegyensúlyozott elemzéseket választok, amelyekben benne vannak az ellenzéki álláspontok is, és a kapcsolókat jól állítom be, akkor viszont a gép megmondja az igazat, vagyis tükrözi a forgalomban lévő véleményeket.

– *Mit szól ahhoz, hogy az egyébként nem túl ijedős „techguruk” megret-*

tentek a mesterséges intelligencia fejlesztésétől?

– Egyetérték velük! Fél évre nyugodtan leállítható a fejlesztés. Ez legfeljebb egy levegővétel idő, nem több.

– Ön miben látja a mesterséges intelligencia társadalmi veszélyét?

– Általánosságban abban, ha a használata során nem tudjuk, hogy milyen adatokra alapozva, melyik változók alapján hozott létre ajánlást. Rengeteg kockázata lehet még a felelőtlen használatának, de a transzparens funkcionalitás szerintem a társadalom alapvető érdeke. Pontos jogi szabályozás nélkül nem tudjuk, hogy a technológia ráengedhető-e a társadalom mikro- vagy makrofolyamataira, mert egyelőre nem ismerjük ennek veszélyeit. Szükség van etikai és jogi kontroll gyakorlására. Arról nem is beszélve, hogy a témának ezernyi vetülete van, emberi jogi, polgári és büntetőjogi, illetve fogyasztóvédelmi. Egyszerre olyan sokféle területét érinti az életünknek, hogy a jog nem képes ezt a komplexitást produkálni, s emiatt lebénul.

– Fél év elég ennyi területnek a szabályozására?

– Nem, a javasolt fél év legfeljebb arra elég, hogy összeírják, milyen területek jogi szabályozását kell feltétlenül elvégezni, illetve miben nem szabad engedni. Nem egyszerű döntések ezek, hiszen egyszerre kell védeni az adatokat, és hagyni az áramlásukat, vagy támogatni az üzleti fejlesztéseket, de nem is szabadjára engedni őket. Soha véget nem érő folyamat ez, aminek azonban jót tenne egy mély levegő, hogy ne pánik, és

toldás-foldás legyen, hanem előrenéző, értő szabályozás. Ehhez az irtózatosságot is kapunk, ősszel jön ki az Európai Unióban a Mesterséges Intelligencia szabályozás 5. dokumentuma, aminek átfutási ideje legalább egy vagy két év. Ez számomra megnyugtató.

– Beszéljünk még a mesterséges intelligencia veszélyeiről. A Guardian című brit lap a veszélyekről publicisztikát íratott a GPT 3-as változatával. Az algoritmus közölte, hogy eszébe sincs kipusztítani az embereket, akár fel is áldozná magát értük. Tehát a robotok nem szönek hatalmi terveket, legfeljebb barátságos eszközként működnek. Ez nem nyugtatja meg önt, vagy aggasztja, hogy a robot képes kitalálni azt is, hogy mit vár tőle az emberiség?

– Mint mondtam, ez egy nyilvános teszt, természetesen mást nem is mondhat, mint hogy szeretettel közelít. A drogból is ingyen kapod az első dózist, csak később kell érte fizetni. Az persze kétségtelen, hogy a mesterséges intelligencia megjelenése a hadászatban már komoly kockázatot jelent. Ha például a Chat GPT 4 utasítást kap arra, hogy a Rubicon folyón átmenő első harci járművet lője le, s jön is arra egy harci eszköznek tűnő böhöm nagy autó, akkor azt az MI teketória nélkül kilövi. Tehát nem szabad algoritmizálni mindent, mert életveszélyes. Mondhatjuk: az MI csodálatos dolog, az ember segítő társa, és tényleg nem lehet célja, hogy elpusztítson bennünket. Viszont, ha mi nem tudjuk kézben tartani, akkor általa mi pusztítjuk el magunkat.