

ÚJ FEJEDELEM

CSEPELI GYÖRGY

Eötvös Loránd Tudományegyetem Társadalomtudományi Kar, Budapest

Beérkezett: 2023. szeptember 13., elfogadva: 2024. január 9.

Nick Bostrom 2014-ben az „uralkodó” típusú mesterséges intelligenciáról írva profetikusan előre látta a datafikáció, a digitalizáció és a hálózatosodás talaján kibontakozó kollektív szuperintelligencia lehetőségét, melynek megvalósulását a generatív nyelvi modellek megjelenése 2022-ben elérhető közelségbe hozott. A természetes nyelveken értelmes szövegtestekben létező és a jövőben létező mondatokon tanuló mesterséges intelligencia fejlődése végső soron felveti a kérdést, hogy az ember képes lesz-e maga által létrehozott eszköz feletti uralomra vagy az elfeledtetni vele a lét igazságát.

Kulcsszavak: digitalizáció, datafikáció, hálózatosodás, kollektív szuperintelligencia, generatív nyelvi modellek

THE NEW PRINCE

Nick Bostrom, writing in 2014 about the “sovereign” type of artificial intelligence, prophetically foresaw the unfolding of cognitive and robotic applications of artificial intelligence that has brought the emergence of collective superintelligence within reach through datafication, digitalization, and networking. The article discusses the implications of the rise to dominance of the ‘new prince’ in economy, health, education, and politics. The development of artificial intelligence that exists in natural languages in meaningful bodies of text and learns from sentences that will exist in the future ultimately raises the question of whether humans will be able to dominate the tool they have created or whether it will subjugate the one that created it.

Keywords: digitalization, datafication, connectedness, collective superintelligence, generative language models

Nick Bostrom a szuperintelligenciáról 2014-ben írt könyvében a mesterséges intelligencia válfajairól, az ő kifejezésével élve „kasztjairól” beszélve megkülönbözteti az orákulumot, a dzsinnt és az uralkodót. Az orákulum természetes nyelven kérdezhető kérdésmegválaszoló rendszer, mely a könyv megírását követően pár évvel később korábbi szűkös keretei közül kilépve, generatív nyelvi modellre alapozva meg is jelent, és az interneten különböző nyelveken elérhető mondatrengetegből szemantikai logika alapján válogatva kérdésekre tud válaszolni.

A dzsinn egy parancsvégrehajtó rendszer, mely végrehajt egy adott utasítást, majd megáll, várva a következő parancsot. Az uralkodónak a dzsinnel szemben nyílt végű felhatalmazása van széles körű és akár nagyon hosszú távú célok végrehajtására (Bostrom 2015: 213).

A mesterséges intelligencia elképesztő iramú fejlődése a három kaszt közötti határok elmosódását eredményezte, előre vetítve a Bostrom által szuperintelligenciának nevezett rendszer megjelenését, mely ugyanazon intelligenciafunkcióra képes, mint az emberi elme, és magában rejtj annak lehetőségét, hogy meghaladja a természetes intelligencia biológiai korlátok miatt lehatárolt képességeit.

Az orákulum, a dzsinn és az uralkodó képességeit egyesítő új fejedelem planetáris hatókörű szuperintelligenciájának lehetősége nem jöhetett volna létre a Földön zajló emberi és természeti történések által keletkeztetett adatok megállíthatatlan ütemű bővülése és interneten történő centralizációja nélkül. Az új fejedelem nem léphetne színre a korábban bekövetkezett datafikáció, digitalizáció és hálózatosodás nélkül. E kollektív szuperintelligencia közvetlen előzménye a különböző területekről származó digitális adatokat az interneten egymástól elkülönülve tároló adatközpontok létrehozása volt. A fejlődés következő szakaszában az adatközpontok az adattárolási kapacitást jelentékenyen megnövelő felhőbe kerültek, ami megkönnyítette a működés révén tanuló mesterséges intelligencia alkalmazásainak megjelenését, aminek végső eredménye a planetáris kollektív szuperintelligencia, azaz az új fejedelem színrelépése lesz. Az új fejedelem első, még messze nem tökéletes működésének már most a tanúi vagyunk.

A generatív, tanuló mesterségesintelligencia-rendszerekről bebizonyosodott, hogy szövegírásban, zeneszerzésben, mozgó és álló képalkotásban, tudományos és technikai problémák megoldásában felveszik a versenyt az átlagos emberi intelligencia produktumaival. A generatív mesterséges intelligencia megjelenése lépés az új fejedelem uralmához vezető úton, melynek végén az emberiség gazdasági, politikai és kulturális életének radikális átalakulása, s attól elválaszthatatlanul magának az embernek a változása lesz várható.

Nincs messze az idő, amikor az új fejedelem fejlesztői a planetáris, kollektív szuperintelligencia-alkalmazásokat önfejlesztésre teszik képessé, miáltal azok a rendszer fejlesztőitől elszakadva önálló életre támadnak. A mesterségesintelligencia-rendszerek egyre kisebb erőfeszítésekkel, egyre nagyobb teljesítményekre képesek. Három évvel azután, hogy az Open Ai útjára bocsátotta a GPT-3 modellt, a fejlesztők olyan modelleket alkottak, amelyek egyhatodnyi ráfordítással ugyanannyira képesek, mint az eredeti, ráadásul ingyen, mindenki számára hozzáférhetőek az interneten. Soha még korábban nem volt olyan technológia, mely ilyen gyorsan, ilyen széles körben elterjedt volna a felhasználók körében. Ez a tömeges használat pedig hozzájárul a modellek további fejlődéséhez.

A világ a jövőben nem lesz biztonságosabb. Az új fejedelemnek a változékonyság, a bizonytalanság, az előreláthatatlanság és a kiismerhetetlenség korában kell majd uralkodnia, ami azt jelenti, hogy szembe kell néznie az általa a munkából kiszorított emberek tömegeivel, a kiberbűnözéssel, a vallási, illetve politikai terrorizmussal, és a klímaváltozás okán bekövetkező globális változásokkal, melyek következményeként milliók kényszerülnek majd lakóhelyük elhagyására.

A mesterséges intelligencia sokféle alkalmazása egyre több embert segíthet a jobb, teljesebb, boldogabb élethez, de látnunk kell, hogy ugyanazok az alkalmazások visszajukra is fordíthatók. Az önvezető autókat a városokban és a közutakon ugyanaz a rendszer működteti, mint az önvezető tankokat a csatatereken. A polgári és a katonai célú mesterséges intelligencia alkalmazásai között a határ elmosódik. Tíz év múlva nem lesz a világon hely, ahova az új fejedelem nem teszi be a lábát. A mesterséges intelligencia alkalmazásai ott lesznek a politikában, a nyilvánosságban, a közigazgatásban, a bíróságokon, a rendőrségen, az iskolákban, a kórházakban, az egyetemeken, a közlekedésben, a hadseregben.

A kollektív szuperintelligencia letéteményeseként fellépő új fejedelem ma még üres trónjára a világban többen pályáznak. A mesterséges intelligencia fejlődése a világot megosztó gazdasági, kulturális és politikai szakadékok elmélyülésével fenyeget. A mesterséges intelligenciát fejlesztő technológiai óriásvállalatok a nemzetállamok határain túllépve globális geopolitikai szereplőkké válnak, ellenőrzésük, működésük kordában tartása megoldhatatlan feladat lesz a nemzetállamok kormányai számára.

Az egyes nemzetállamokban megszülető új fejedelmeknek az alattvalóiktól folyamatosan egyre nagyobb tömegben érkező adatok birtokában minden eszközük meglesz ahhoz, hogy leépítsék, kiüresítsék a politikai akaratképzés demokratikus intézményeit és totálisan ellenőrizzék a hatalomnak magukat önként alávető alattvalókat. Minden új fejedelem elemi érdeke a mesterségesintelligencia-technológiáinak fejlesztése, ugyanakkor látnivaló, hogy az egyes nemzetállamok erői végesek, a verseny kiélezett, s csak néhány államnak van reális lehetősége arra, hogy a kollektív szuperintelligencia adta kompetenciákat megszerezve esélye legyen a világ sorsának befolyásolására. Az országok túlnyomó többsége rá lesz szorulva a mesterséges intelligencia fejlesztésében élen járó technológiai óriásvállalatok és az azokkal összefonódott nagy országok segítségére. A forrásokban szegény és forrásokban gazdag országok közötti kapcsolat nem a nulla összegű játszmák logikája szerint alakul. Ellenkezőleg, mint azt az internet, az okos telefonok, a digitális oktatás és egészségügy példáin látjuk, a szegény országok az új technológiák hasznélvezői lehetnek akkor is, ha a fejlesztés fölött nincs is uralmuk.

Másfelől a világ feletti digitális uralmat gyakorló új fejedelem trónjáért folytatott küzdelem nyers és kegyetlen lesz. A mesterséges intelligencia – mint azt a jelenleg Ukrajna és Oroszország közötti háború mutatja – páratlan értékű harci eszköz, melynek alkalmazásával szemben szükségképpen alul marad a pusztá emberfölnyben bízó ellenfél.

Az új fejedelem a világban nem lehet más csak az Egyesült Államok vagy Kína. A harc kettőjük között fog eldőlni, s a párharcot nulla összegű játszmaként felfogva egyikük sem tűnik hajlandónak az együttműködésre. Mindketten mindent megtesznek annak érdekében, hogy a saját országaikban zajló mesterségesintelligencia-fejlesztés során a másik ne tehessen szert behozhatatlan katonai és gazdasági előnyökre.

A kollektív mesterséges szuperintelligencia kialakulásának most is zajló folyamatában a fordulópontot a természetes nyelvi programozás jelentette, aminek révén a gépi és emberi intelligencia korábban elképzelhetetlen közelségbe került. Az első lépést a gépi és az emberi agy közelítése felé Turing tette meg, leírva az univerzális intelligens gépet, mely minden feladatot megold, amit az ember is meg tud oldani. Mint Turing életrajzírója, Andrew Hodges írja, Turing szerint „létezhetne egy olyan gép, amely a »szalagjáról« olvasva más gépek leírását az ember mentális aktivitásával azonos munkát tudna végezni. Egy gép, amely fel tudná váltani a számítást végző embert. Elektromos agy?” (Hodges 2015: 184.) A gépi és emberi intelligencia közötti határt Turing a gép és az ember közé beiktatott természetes nyelvi teszttel kívánta megállapítani, melynek révén a géppel írásban kommunikáló ember eldönthette, hogy a párbeszéd másik szereplője ember-e vagy gép.

A generatív nyelvi modellek robbanásszerű fejlődése nyomán a Turing-tesztek alkalmazhatósága rohamosan csökken, mivel a mesterséges intelligencia által természetes nyelveken létrehozott szövegekkel szembesülő emberi intelligencia egyre kevésbé képes megállapítani, hogy az adott szöveg forrása gép vagy ember.

Az emberi és a mesterséges intelligencia között a nyelv döntő kapocs, mivel a nyelv – mint Heidegger mondja – „a lét háza”, ami bővebben kifejtve azt jelenti, hogy „a nyelv lényegét a létnek való megfeleléséből gondoljuk el, méghozzá mint ezt a megfelelést, azaz mint az emberi lény hajlékát” (Heidegger 2003: 293). A létbe hatoló gondolkodás alapvetően a nyelv közegében lehetséges. A lét végtelen, miként a nyelv is az. A nyelvészetet forradalmian megújító Chomsky szerint a nyelv

a grammatika szabályai által formált mondatok végtelen halmaza, s ezáltal képes a lét sokdimenziós, komplex igazságának feltárására (Chomsky 1985). Az emberi tudatban a kinti és a benti végtelenség találkozik, amit az emberrel veleszületett nyelvi kompetencia tesz lehetővé. A kompetencia azonban csak lehetőség, melyet a nyelvi performancia aktualizál. A mesterséges intelligencia generatív nyelvi modelljei nem a hazátlanságában a nyelvben hazára talált ember hajlékának őrzői, a költők és a gondolkodók nyelvi teljesítményére támaszkodnak. A generatív nyelvi modellek forrásai az akárik által leírt vagy kimondott mondatok.

Heidegger arra figyelmeztet, hogy a mondatok szavai csak akkor visznek a lét igazságához, ha az ember, mielőtt megszólalna, némán hagyja, hogy a lét megszólítsa. A gondolkodás a lét meghallott szavát juttatja kifejezésre. A már kimondott szavak, mondatok iszonyatosan nagy és folyamatosan bővülő tömegéből válogató generatív nyelvi modellek nem gondolkodnak, csak utánozzák a gondolkodást. A chat GPT és társai létrehozott szövegekbe olyan mondatok kerülnek, amelyek alapján már eleve eldől, hogy mi érthető, és mit kell mint értelmetlent elvetni. A mindent mindenki számára közel hozó nyelv a nyilvánosság diktatúrája alá kerülve az akarki nyelveként elveszti az igazsághoz vezető utat. A gép és az ember elsőkélyesedett közös nyelve az ember lényegét veszélyezteti. A mesterséges intelligencia az emberi gondolkodást leutánozva a gondolkodás technikai értelmezése előtt nyit teret, szem elől veszve a létet, mint a gondolkodás első számú célpontját.

IRODALOM

- BOSTROM, N. (2015) *Szuperintelligencia*. Budapest, Ad Astra.
- CHOMSKY, N. (1985) *Generatív grammatika*. Budapest, Magvető.
- HEIDEGGER, M. (2003) Levél a humanizmusról. In: PONGRÁCZ T. (ed.) *Útjelzők*. Budapest, Osiris. pp. 293–334.
- HODGES, A. (2015) *Kódjátzsma*. Budapest, Gabo.