

JOHN BROCKMAN:
POSSIBLE MINDS – 25 WAYS OF LOOKING AT AI
A mesterséges intelligenciáról – haladóknak

„A társadalomra nem a gép jelenti a veszélyt,
hanem az, amivé az ember teszi a gépet.”

(Norbert Wiener)

Miben különbözik a gépi intelligencia és az emberi agy; utolérheti-e a mesterséges a természetes intelligenciát? Létezhet-e nem szénelapú, mesterséges tudat? A mesterséges intelligencia (MI) beilleszkedik a műszaki fejlődés történelmi vonalába, vagy egy új jelenséggel állunk szemben? Önállósodhat-e az MI? Képesek vagyunk/leszünk-e az MI emberbarát programozására? Veszélyt jelenthet-e az emberre? Elveszi-e a munkánkat, és feleslegessé teszi-e az embert? Az MI kapcsán ezek a leggyakrabban felmerülő, legvitatottabb kérdések, a *Valóság* 2019. decemberi számában én magam is foglalkoztam velük *A mesterséges intelligenciáról – haladóknak* című tanulmányomban.

Ismertetésünk tárgya a John Brockman New York-i kulturális-tudományos menedzser által szerkesztett, 2019-ben megjelent kötet: *Possible Minds – 25 Ways of Looking at AI* (Lehetséges elmék – A mesterséges intelligenciáról 25-féle megközelítésben). Brockman időszerű és fontos témákkal foglalkozó tudományos könyveket ad ki, a világ leghíresebb tudósait szólaltatva meg bennük. Nevéhez fűződik a „harmadik kultúra” fogalmának megalkotása is. E könyvében a világ huszonöt leghíresebb, mesterséges intelligenciával foglalkozó tudósát kérdezte meg a témáról, mely most ahhoz hasonló fogalom lett, amilyen sokáig Norbert Wiener kibernetikája volt. A megkérdezettek többsége híres amerikai egyetemek (főként az MIT és a Harvard) professzora, van köztük két kulturális menedzser is.

A szerzők a témát a legkülönbélebb szempontok szerint vizsgálják: lételméleti, antropológiai, szociológiai, informatikai, biológiai, evolúciós, technológiai, neurobiológiai, etikai, pszichológiai, neveléstudományi, kulturális és kognitív tudományok szempontjából. Megállapításaik – lévén a megszólaltattak a szakma legkiválóbbjai – részben további adalékul szolgálnak azokhoz az összefüggésekhez és dimenziókhöz, melyeket fentebb említettünk, részben túlmutatnak rajtuk. A recenzió műfajának megfelelően a továbbiakban csupán a szerzők legfontosabb megállapításait közöljük, azok kifejtése nélkül.

George M. Church úgy gondolkodik, hogy végső soron minden dolgot és szerkezetet legkisebb alkatrészeire, atomjaira lehet bontani, és ennek megfelelően, alulról fölfelé haladva, létrehozhatunk bármilyen kombinációjú szerkezetet az

élettelentől az élőig, az oktalantól az intelligensig terjedő skálán. Ezért azzal kell számolni, hogy a jövőben sokféle létező és sokféle elme vesz majd minket körül. A fő probléma nem az emberi és egy mesterséges intelligencia viszonya lesz, hanem az, hogy hogyan viszonyuljunk a legkülönbélebb intelligenciákhoz, és milyen jogokat adjunk a sokféle létezőnek. Mindezzel szemben áll Daniel C. Dennett leegyszerűsítő kijelentése: „Nem mesterséges tudattal rendelkező ágensekre, hanem intelligens eszközökre van szükségünk.” És hogy véleménye felől a „technofiliásoknak” ne maradjon kétségük, még ezzel is megtoldja: „Az MI jelenlegi formájában a humán intelligencia élősködője.”

Frank Wilczek szerint, ha a tudatot emergens jelenségként fogjuk fel, akkor – legalábbis elméletileg – mesterségesen is létre lehet hozni az atomok olyan kombinációját, mely azonos az aggyal, és amelyből majd – a komplexitás kellően magas fokán – emergens módon létrejön a tudat. Sőt, e logikai szálát tovább fonva, Wilczek arra a következtetésre jut, hogy az intelligencia általános formája valójában maga az MI, az emberi intelligencia pedig annak egy specifikus megnyilvánulása.

A fordulópontot, az igazi veszélyt az ÁMI (általános mesterséges intelligencia) jelenti, amely minden területen eléri és meghaladja majd az emberi intelligenciát. Ezt a határt különösen Max Tegmark tekinti létfontosságúnak. Nyomatékosan hirdeti, hogy még az ÁMI kialakulása/megalkotása előtt ki kell dolgozni azokat az elveket és szabályokat, amelyek az MI használatát biztonságossá teszik. Az ÁMI-val kapcsolatos legsúlyosabb probléma az lesz, hogy milyen célokat és értékeket tápláljunk a gépbe. Stuart Russel szerint, mivel elménk egyes alrendszereinek preferenciái inkompatibilisek egymással, a vágyaink, óhajaink kielégítésére törekvő ÁMI megoldhatatlan feladattal találja szemben magát. Az emberi viselkedés és gondolkodás jelentős mértékben irracionális, a feladat ezért is tűnik megoldhatatlannak: ki határozza majd meg ezeket az értékeket? Steven Pinker azonban az ÁMI-nak tulajdonított veszélyeket a tudományos alapok nélküli, ótestamentumi átkozódásokhoz és jeremiádákhoz hasonlítja. Szerinte a rettegek téves platformon állnak: azt hiszik, hogy az intelligenciából következik a dominancia vágya. Azonban inkább ennek az ellenkezője valószínű: a szuperintelligencia jóhiszemű lesz. David Deutsch is alaptalannak tartja félelmünket, hogy az ÁMI-k okosabbak lesznek nálunk. Létrehozásuk során a mi gondolkodásunk is fejlődik. Az emberi agy bármit el fog tudni gondolni, amit az ÁMI, a különbség csak a gyorsaságban lesz, az a hátrány viszont majd külső eszközök segítségével behozható lesz.

Az embert utánzó mesterséges intelligencia kifejlesztésén dolgozunk, miközben nem ismerjük jól magunkat, a mintát. A robotok tanításához, ösztönzéséhez is azt kell megértenünk, hogy hogyan tanulnak az emberek, mi ösztönzi őket. Ez az „inverz tanulási ösztönzés” (inverse reinforcement learning) témája. Ezeket a kérdéseket tárgyalja Tom Griffiths. Judea Pearl a *deep learning*et a végtelenül

lassú darwini szelekcióhoz hasonlítja, amikor például a kígyó vagy a sas látása évmilliók során tökéletesedett. Ezzel szemben, az embernek csak évszázadokra, legfeljebb évezredekre volt szüksége ahhoz, hogy látását nagyítóval, szemüveggel majd idővel mikroszkóppal tökéletesítse. Ez az emberi tulajdonság ugyanis nem a darwini evolúció eredménye; hanem annak a következménye, hogy a Homo sapiens képes „a valóság mentális reprezentációjára”. Alison Gopnik az MI fejlesztéséhez a gyermeki tanulás csodálatos elemeit ajánlja felhasználásra: kíváncsiságot kell bevinni a gépbe, és lehetővé kell tenni, hogy a gép aktívan interakcióba lépjen a külvilággal, mint ahogy az a gyermek tanulásának is fontos feltétele. Képesek vagyunk gondolatok és intelligencia automatikus előállítására – ez az MI. De tudunk-e anyagot is automatikusan (azaz digitális technikával) előállítani? Ez a kérdés foglalkoztatja Neil Gershenfeldet.

Seth Lloyd kétségbe vonja, hogy a műszaki fejlődés tartósan exponenciális jelleget ölthet, és ezáltal a szingularitást és a transzhumán lények kialakulását. W. Daniel Hillis hibrid intelligenciákról beszél, és ebbe a kategóriába sorolja az államot és a nagyvállalatokat is. Végül egyes szerzők olyan következtetésekhez is eljutnak, hogy az emberi agy mégiscsak magasabb rendű az akármilyen elképzelt szuperintelligenciánál (George Church és Fred Wilchek), George Dyson szerint pedig az analóg számítás nem halt ki, sőt, már visszatérőben van. Alex Pentland a Big Datát korunk prométheuszi tüzének nevezi, mely ugyanúgy fordítható jóra, mint rosszra. Ugyanis fennáll a veszélye az algoritmusok zsarnokságának is, amikor senki által sem választott adatkezelők irányítják a világot.

John Brockman az MI-hez fűződő reményeket és félelmeket még a szokásosnál is bombasztikusabban tálalja: „Az MI korunk sztorija; ez áll minden mai történetés mögött, ez Krisztus második eljövetele, és ez az Apokalipszis.” Könyvében a téma súlyához megfelelő neves tudósokat szólaltatott meg, akik jelentősen bővítik a mesterséges intelligenciára vonatkozó ismereteinket.

(John Brockman: Possible Minds – 25 Ways of Looking at AI. Penguin Press, 2019)

Kiss Károly

közgazdász, társadalomkutató, kandidátus,
a Budapesti Corvinus Egyetem nyugalmazott egyetemi docense