

Andreas Sudmann (szerk.)

A mesterséges intelligencia demokratizálódása

A hálózati politika a tanuló algoritmusok korában

Forrás: SSOAR

Tartalom

A mesterséges intelligencia demokratizálódása

A hálózati politika a tanuló algoritmusok korában

Andreas Sudmann 9

Metaforák, amelyek szerint élünk

Három kommentár a mesterséges intelligenciáról és az emberi állapotról

Anne Dippel 33

AI, sztereotipizálás szteroidokon és Alan Turing biológiai fordulata

V. N. Alexander 43

Termékeny hangok

Az érintőhangos tárcsázás, a Call Center ipar
felemelkedése és a virtuális hangalapú asszisztensek
politikája

Axel Volmar 55

Algoritmikus kereskedés, mesterséges intelligencia és a megismerés politikája

Armin Beverungen 77

A használható adatok keresése

Gépi tanulási algoritmusok építése a közzsféra archívumaiból

Lisa Reutter/Hendrik Storstein Spilker 95

Plurális, helyhez kötött szubjektumok a mesterséges intelligencia kritikájában

Tobias Matzner 109

A mélytanulás kormányzati jellege

A másik fekete doboz

Jonathan Roberge/Kevin Morin/Marius Senneville 123

Csökkentés és részvétel

Stefan Rieger 143

A mesterséges intelligencia politikai vonatkozásai	
<i>Dan McQuillan</i>	163
Mesterséges intelligencia	
Láthatatlan ügynökségek a technológiai kultúrák ráncáiban	
<i>Yvonne Förster</i>	175
Verseny és számítógépes látás	
<i>Alexander Monea</i>	189
A mesterséges intelligencia demokratizálódásának feltérképezése a GitHubon	
Egy első megközelítés	
<i>Marcus Burkhardt</i>	209
A mesterséges intelligencia médiapolitikai dimenziójáról	
A mélytanulás mint fekete doboz és az OpenAI	
<i>Andreas Sudmann</i>	223
Hogyan védjük meg a mesterséges intelligenciát	
<i>Ina Schieferdecker/Jürgen Großmann/Martin A. Schneider</i>	245
Mesterséges intelligencia, demokrácia és jog	
<i>Christian Djeffal</i>	255
A tudás problémájának újragondolása a vállalati gigantizmus korában	
<i>Frank Pasquale</i>	285
A mesterséges intelligencia és a művészet demokratizálódása	
<i>Jens Schröter</i>	297
Köszönetnyilvánítás	333

A mesterséges intelligencia demokratizálódása

A hálózati politika a tanuló algoritmusok korában

Andreas Sudmann

Az idő diagnózisa természetesen nehéz vállalkozás. Mindazonáltal valószínűleg adekvát megállapítás, hogy jelenlegi történelmi helyzetünkben a demokrácia stabilitásáért és jövőjéért való aggodás különösen mélyreható (vö. Rapoza 2019). Ennek az aggodalomnak a tárgyai egyrészt olyan fejlemények, amelyek látszólag csak korlátozottan vagy közvetve kapcsolódnak a technológiai kérdésekhez, mint például a jobboldali populizmus és a tekintélyelvűség jelenlegi erősödése, különösen Európában és az Egyesült Államokban, vagy "a konfrontatív geopolitika újjáéledése" (Valladão 2018). Másfelől egyre inkább elterjedt diskurzusnak lehetünk tanúi, amely a mesterséges intelligencia (AI) legújabb fejleményeit a demokráciára és a demokratikus értékekre nézve potenciálisan komoly fenyegetésként tárgyalja, de amely - fontos kivételektől eltekintve - úgy tűnik, hogy nagyrészt elszakadt az egyes országok konkrét politikai körülményeitől és de- volúcióitól (vö. Webb 2019). Ebben a diskurzusban a mesterséges intelligencia demokratikus értékeket és elveket veszélyeztetőnek való prob- lematizálása különböző, de részben összefüggő jelenségekre utal. E viták központi hivatkozási pontjai például az AI-technológiák társadalmi-politikai következményei a jövőbeli munkaerőpiacra (jelszó: "a munka eltűnése"), az AI alkalmazása a vizuális információk manipulálására vagy az "álhírek" létrehozására, az autonóm fegyverrendszerek geopolitikai hatásai, vagy az AI-módszerek alkalmazása hatalmas megfigyelőhálózatokon keresztül az ítélkezési irányelvek és a bűnügyi igazságszolgáltatási rendszerekben a *viisszaesés kockázati* pro- fliktusainak előállítására, vagy a testek demográfiai és pszichográfiai célzására a reklám, a propaganda és az állami beavatkozás egyéb formái számára.¹

Első látásra a demokrácia globális állapotával kapcsolatos aggodalmak mindkét formája nem

sok közös vonásuk van, de éppen emiatt kell mélyebb összefüggéseiket feltárni. Donald Trump amerikai elnök például nemrég indította el az úgynevezett "amerikai mesterséges intelligencia-kezdemenyezést", amelynek kifejezett célja, hogy az intelligens technológiák fejlesztését az amerikai érdekeket előtérbe helyező módon támogassa.

¹ Magától értetődik, hogy nem minden olyan szempont, amely valamilyen okból kritikára érdemesnek tűnik, jelent közvetlen veszélyt egy társadalom demokratikus rendjére. Ugyanakkor az is nyilvánvaló, hogy a kormányzatnak és a társadalomnak választ kell találnia a mesterséges intelligencia minden problémájára.

A Google/Alphabet nagyjából ugyanebben az időben jelentette be, hogy megnyitotta első mesterséges intelligencia laboratóriumát Ghánában. A Szilícium-völgyi központú technológiai óriás továbbra is folytatja stratégiáját, amelynek keretében világszerte AI-kutatóközpontokat hoz létre: New York, Tokió, Zürich és most Ghána fővárosa, Accra. A laboratórium vezetője, Moustapha Cisse szerint az egyik cél az lesz, hogy a fejlesztők számára biztosítsa a szükséges kutatásokat, amelyek segítségével olyan termékeket lehet létrehozni, amelyek képesek megoldani néhány olyan problémát, amellyel Afrika napjainkban szembesül. Az ilyen stratégiák sikeres megvalósítására példaként említik, hogy a Google nyílt forráskódú gépi tanulási library TensorFlow segítségével olyan alkalmazást lehetett fejleszteni okostelefonokra, amely lehetővé teszi a növényi betegségek észlelését Afrikában, akár offline is.

A Google/Alphabet és más technológiai vállalatok "humanista" mesterséges intelligencia programja első pillantásra éles ellentétben áll Donald Trump "Amerika az első" mesterséges intelligencia politikájával. Annak azonban, hogy a Szilícium-völgyi vállalatok egyre inkább igyekeznek előmozdítani az olyan demokratikus értékeket, mint a hozzáférhetőség, a részvétel, az átláthatóság és a sokszínűség, semmi köze ahhoz a motivációhoz, hogy elhatárolódjanak a jelenlegi amerikai kormány irányvonalától. Sokkal inkább növekszik azoknak a kritikusoknak a száma, akik a Google-t, a Facebookot és a többi technológiai óriást magukban komoly veszélyt látnak a demokráciára és/vagy a demokratikus értékekkel ellentétesen cselekszenek üzleti stratégiáik, adatkezelési gyakorlataik és hatalmas gazdasági és társadalmi-közéleti hatalmuk tekintetében.

Ennek megfelelően ezekre a vállalatokra jelentős nyomás nehezedett, hogy reagáljanak erre a növekvő kritikára. A Facebook különösen két nagy botrányba keveredett, mindkettő Trump elnökválasztási kampányával kapcsolatban. Először is 2017-ben fokozatosan ismertté vált, hogy orosz szervezetek és személyek - többségük a szentpétervári székhelyű Internet Research Agency-hez (internetes trollfarm) köthető - hamis fiókokat hoztak létre olyan platformokon, mint a Facebook, a Twitter és az Instagram, és megpróbálták tőkét kovácsolni a 2016-os amerikai elnökválasztás körüli vitákból, részben álhírek létrehozásával. Egy másik botrány a Cambridge Analytica adatelemző és politikai tanácsadó céget érintette. Mint az 2018 márciusában nyilvánosságra került, a cég több mint 80 millió Facebook-felhasználó adataihoz férhetett hozzá, és feltehetően azok előzetes beleegyezése nélkül elemezte azokat, hogy Trump kampányát támogassa.

A botrányok következményeként nem csak Zuckerberg, hanem a Google vezérigazgatója, Sundar Pichai is nemrég tanúvallomást tett a washingtoni kongresszus előtt. A meghallgatások során különösen Zuckerberg ismerte el a múltban elkövetett számos mulasztást, és ígéretet tett arra, hogy fokozza az együttműködést a kormányzati intézményekkel és a nem kormányzati szervezetekkel, valamint hogy megvizsgálja az adatvédelem javítását célzó intézkedéseket, és végül ennek megfelelően hajtja végre azokat. Ami Európát illeti, az európai általános adatvédelmi rendelet ("GDPR") már tartalmazza az

adattvédelem javítására és betartására vonatkozó jogi követelményeket. A kongresszusi meghallgatásokon Zuckerberg kijelentette, hogy elvileg hajlandó támogatni a hasonló állami szabályozási intézkedéseket a

USA. Ugyanakkor félelmét fejezte ki, hogy a kínai versenytársak technikailag felülmúlhatják vállalatát, mivel az ország hagyományosan sokkal kisebb hangsúlyt fektet az adatvédelmi kérdésekre, mint Európa vagy az USA (vö. Webb 2019). A Facebook adatvédelmi politikában való együttműködési hajlandóságának azonban más okai is vannak: Legalábbis a WhatsApp és az Insta-gram felvásárlása óta a Facebook de facto monopolhelyzetbe került a közösségi média területén. Hasonló a helyzet az Amazon esetében az e-kereskedelemben és a Google esetében a keresőmotorok terén - és éppen ez a hatalmas hegemón pozíció az, ami egyre hevesebb viták tárgya. Nemrégiben még a Facebook társalapítója és korábbi szóvivője, Chris Hughes (2019) is kritizálta Zuckerberg cégét, mint az amerikai gazdaságra és demokráciára jelentett fenyegetést, és a cég feldarabolása mellett érvelt, hogy nagyobb verseny legyen a közösségi média szektorban. Különböző okok miatt meglehetősen kérdéses, hogy egy ilyen forgatókönyv bekövetkezhet-e a közeli vagy távoli jövőben. Mindazonáltal a globális "platformkapitalizmussal" (Srnicsek 2016) vagy "felügyeleti kapitalizmussal" (Zuboff 2018) szembeni kritika egyre erősödik, és ez a mesterséges intelligencia szerepét is érinti abban, amit az utóbbi időben néha új "adatgazdaságnak" neveznek (vö. például Bublies 2017).

Nem utolsósorban az eddig említett problémák és jelenségek tekintetében a A kötet célja, hogy feltárja a mesterséges intelligencia politikai dimenzióját, kritikusan összpontosítva a jelenlegi kezdeményezésekre, diskurzusokra és az úgynevezett "demokratizálás" koncepcióira. Ez utóbbi fogalom egyik sajátossága, hogy egyszerre homályos és kon-krétikus. Amint az a jelenlegi AI-diskurzusból kiderül, a fogalom számos különböző jelenségre utalhat, mégis többé-kevésbé összefüggő vagy koherens jelentésfelfogások együttesét idézi elő. Ennek megfelelően a demokratizálódás felfogható egy olyan etika megvalósulásaként, amely a politikai tájékoztatásra, a kritikára való hajlandóságra, a társadalmi felelősségvállalásra és aktivitásra, valamint egy olyan politikai kultúrára irányul, amely kritikusan a hatalommal szemben, részvételi és általános irányultságában befogadó. A demokratizálás tehát politikai, intervenciós gyakorlatként fogható fel, amely elvileg alkalmazható (és természetesen már alkalmazták is) a társadalomra általában, valamint annak számos alrendszerére vagy egyes területeire (például a technológiára).² A kötetben kritikusan vizsgálendő egyik központi kérdés, hogy a hálózati politika (és különösen a demokratizálási eszmékkel és tevékenységekkel kapcsolatosak) milyen mértékben kerültek új feltételek közé a mesterséges intelligencia technológiák széles körű meghonosítása és ipari megvalósítása érdekében. A hálózati politika fogalmát itt a következő fogalmak széles spektrumának heurisztikus gyűjtőfogalmaként értelmezzük

2 Természetesen a politikai elméletben a kifejezés egy demokratikusabb rendszerbe való átmenetet is jelent, vagy a demokráciák fejlődésének történelmi folyamatait írja le. A demokrácia és a demokratizálódás fogalmának tárgyalásához lásd Birch (1993), a demokrácia és a technológia kapcsolatának megvitatásához lásd például Mensch/Schmidt (2003), Diamond/ Plattner (2012) vagy Rockhill (2017) írásait.

kritikai kutatás, hogy megvilágítsa a hálózatok és a politika összefonódásának és kapcsolatának különböző formáit, mind társadalmi-technikai diskurzusokként, mind pedig gyakorlatokként. Mint ilyen, a politika hálózati dimenziójával, valamint a különböző típusú társadalmi, kulturális vagy technológiai hálózatok - beleértve, de nem kizárólagosan az internetet vagy az úgynevezett közösségi médiát - politikai feltételeivel, következményeivel és hatásaival foglalkozik.³ Ennek megfelelően a kötet nem csupán a szűkebb értelemben vett mesterséges intelligencia és az internetes technológiák kapcsolatának politikai aspektusait kívánja feltárni (pl. le- gal keretek, politikai tartalmak a közösségi médiában stb.). A kritikai fókusz inkább a jelenlegi és történelmi AI-technológiák előállításában és kialakításában részt vevő *valamennyi* entitás hálózati és közvetített dimenziójának vizsgálatát foglalja magában.

Először is, egy ilyen feladathoz tisztázni kell a mesterséges intelligencia fogalmát, mivel a kifejezés különböző megközelítéseket foglal magában, amelyek nem mindig különülnek el pontosan, különösen a közbeszédben. Amikor manapság az emberek a mesterséges intelligenciáról beszélnek, akkor többnyire az úgynevezett gépi tanulási technikákra és különösen a mesterséges neurális hálózatokra (ANN) összpontosítanak. Sőt, azt is mondhatjuk, hogy ezek a beszámolók állnak a jelenlegi AI reneszánsz középpontjában. Néha a két kifejezést szinonimaként használják, de ez egyszerűen téves. A gépi tanulás a mesterséges intelligenciában alkalmazott algoritmusok különböző formáinak gyűjtőfogalma, amelyek lehetővé teszik a számítógépes rendszerek számára, hogy komplex adatstruktúrákban statisztikai mintákat elemezzenek és tanuljanak, hogy egy bizonyos *x* bemenetre megjósolják a megfelelő *y* eredményt, anélkül, hogy kifejezetten erre a feladatra lennének felkészülve (vö. Samuel Mitchell 1959, 1997). Az ANN viszont a gépi tanulás egy sajátos, de nagyon hatékony megközelítése, amelyet lazán a biológiai neurális hálózatok ihlettek, és alapvetően a következő jellemzők jellemzik (vö. Good- fellow/Bengio/Courville 2016):

1. az információ feldolgozásának/szimulálásának masszív párhuzamossága a mesterséges neuronok hálózatán keresztül.
2. az információfeldolgozás hierarchikus felosztása, amely az egyszerű minták megtanulásából az egyre összetettebbekig strukturált, a hálózat úgynevezett rejtett rétegeinek *f* lexibilis számához kapcsolódóan.
3. a rendszerek azon képessége, hogy egy meghatározott tanulási célt kvázi-automatikusan, egymást követő önoptimalizálással (a "backpropagation" nevű tanulási algoritmus segítségével) érjenek el.

Valóban állíthatjuk, hogy az ANN és általában a gépi tanulás jelenlegi fellendülése meglehetősen meglepő, tekintve, hogy a mesterséges intelligencia ezen úgynevezett konnekcionista megközelítésének technológiai alapjait már a számítástechnika és a kibernetika kezdete óta kutatják (vö. pl. McCulloch/Pitts 1943, Hebb 1949, Rosenblatt 1958). Néhány rövidebb időszakot leszámítva azonban

14 Andreas Sudmann
figyelemre méltó kivétel,

3 A netpolitikával kapcsolatos kutatások hosszú hagyományának áttekintéséhez lásd például Lovink (2002).

Az ANN-t többé-kevésbé zsákutcának tekintették az AI-kutatás történetében (Sudmann 2016, 2018a). Ez a megítélés ma valószínűleg gyökeresen más, még akkor is, ha a kommentátorok jelentős része az ANN (még) alapvető korlátaira mutat rá, vagy továbbra is fenntartja más megközelítések, például a szimbolikus és szabályalapú formák fontosságát az AI-kutatásban (vö. Pasquinelli 2017, Marcus 2018).

Van némi vita arról, hogy pontosan mikor kezdődött a jelenlegi mesterséges intelligencia boom. Egyes szakértők 2009 körül bizonyos fejlődési ugrásokat emelnek ki a természetes nyelvfeldolgozás (NLP) és a beszédfelismerés területén. Különösen fontos volt azonban a számítógépes látás (CV) területén elért fejlődés. A Torontói Egyetem kutatócsoportja 2012-ben megnyerte az ImageNet nevű képfelismerési versenyt, és több mint felére csökkentette a korábbi megközelítések hibaaarányát. Ez a teljesítménybeli ugrás azért vált lehetővé, mert az úgynevezett konvolúciós neurális hálókat (CNN), azaz a számítógépes látás feladatára optimalizált hálózatokat először sikerült következetesen és hatékonyan GPU-k, azaz gyors, párhuzamosan szerveződő számítógépes hardverek alapján képezni, ahogyan azokat jellemzően a modern játékkonzolokban implementálták (Sudmann 2016).

Mindenesetre a nagy IT-vállalatok is gyorsan fejlődést regisztráltak a számítógépes látás és az ANN területén, ami valóságos boomot eredményezett az induló vállalkozások felvásárlásában és finanszírozásában. Az egyik ilyen startup volt a DeepMind, amelyet a Google 2013-ban 650 millió dollárért vásárolt fel. Három évvel később a DeepMinds AI-rendszere, az AlphaGo képes volt legyőzni az emberi világbajnokot a Go társasjátékban. Az AlphaGo sikerével az AI-boom megérkezett a mainstreambe, azaz az AI gyorsan meghatározó diskurzussá vált a kultúra és a társadalom számos területén, beleértve a legtöbb tudományterületet (Sudmann 2018a, 2018b).

Ez utóbbi nem jelenti azt, hogy az ANN teljesen ismeretlen volt a bölcsészeti és társadalomtudományok területén a 2016 előtti években. Különösen az 1990-es évek eleje körül nőtt meg jelentősen az ANN iránti érdeklődés olyan területeken, mint a kognitív tudományok és az elmefilozófia, nem sokkal azután, hogy az ANN első ipari implementációira sor került, és köszönhetően a backpropagation tanulási algoritmusok 1980-as években való meghonosodásának (Sudmann 2018a, vö. még az *Alexander Waibel*rel készült interjúval ebben az antológiában). Aligha tagadható azonban, hogy számos tudományágban már akkor is meglehetősen korlátozott volt az ANN-nek szánt általános figyelem. Végül az ANN 1980-as évekbeli fellendülése meglehetősen rövidnek bizonyult, ezért egyes ob- szervek megerősítve érzik magukat abban a hitükben, hogy a következő AI-tél eljön - ez csak idő kérdése. Természetesen egy ilyen esemény újra bekövetkezhet, de jelenleg semmi sem utal erre, inkább az ellenkezője látszik.

Mindazonáltal a "mesterséges intelligencia forradalmáról" szóló mindenütt jelenlévő beszéd és a Szilícium-völgy techno-utópistáinak haladásról szóló retorikája önmagában is hatalmas provokációt jelent számos kritikus számára,

nemcsak a humán tudományok területén, hanem az akadémiai világon kívül is. Tagadhatatlan, hogy a mesterséges intelligenciáról szóló vita a kezdetektől fogva jellemzően a kar-

szkeptikus, utópisztikus vagy disztópikus narratívák alapján (vö. Sudmann 2016, 2018b).⁴ És még ma is ritkák az e pozíciók közötti óvatos közvetítések. Mint ilyen, a mesterséges intelligenciáról szóló számos vita a közeli és távoli jövő spekulatív horizontjára irányul. És az sem véletlen, hogy a mesterséges intelligenciát ironikusan úgy írják le, mint éppen azt a kutatási területet, amely azzal foglalkozik, hogy feltárja azt, amire a számítógépek *még nem képesek* (vö. Michie 1971). Más szavakkal: Amint egy számítógép elsajátít bizonyos képességeket, az ilyen rendszer már nem tekinthető mesterséges intelligenciának. Így a mesterséges intelligencia végleg az utópia (vagy disztópia) birodalmába kerül.

Ugyanakkor csak nemrégiben léptünk abba a történelmi szakaszba, amelyben az AI mint sci-fi vagy technikai utópia és az AI mint az empirikus világ létező technológiája közötti szakadék záródni látszik. Persze joggal lehet itt rámutatni, hogy például az önvezető autókat már az 1980-as években, sőt még azelőtt is tesztelték az utakon, ⁵vagy hogy az első gépi nyelvi fordítórendszereket már az 1950-es években kifejlesztették (vö. Booth/ Locke 1955), de ez nem változtat azon a tényen, hogy mindkét technológia csak a közelmúltban nyerte el vagy közelítette meg azt az alkalmazhatósági potenciált, amelyet a globális gazdaság elvár tőlük.

A mesterséges intelligencia ipari felhasználhatósága és az emberi képességek egyre nagyobb mértékű felülmúlása különböző alkalmazási területeken új jelenségnek tűnik. A számítógépek azonban már az első naptól kezdve a "mesterséges intelligencia" egy formája voltak, és mint ilyenek, olyan dolgokra voltak képesek, amelyekre az ember (egyedül) nem volt ugyanúgy képes, például az Enigma német titkosítógép kódjának feltörésére (vö. Kittler 2013, vö. Dotzler 2006).

Tekintettel az új innovációk gyors ütemére és az alkalmazási területek bővülésére, korántsem könnyű feladat annak meghatározása, hogy a mesterséges intelligencia hogyan alakítja át napjainkban az emberek, a technológia és a társadalom közötti kapcsolatot, és milyen hatással van arra, hogy miként vagyunk képesek megfelelő módon megragadni ennek az elmozdulásnak a politikai és történelmi dimenzióját.

Az erre a kérdésre való válaszkérés magában foglalja az AI-vitában a kezdetektől fogva tárgyalt problémák átgondolását, például a hagyományosan antropocentrikus fogalmak, mint az érzékelés, a gondolkodás, a logika, a kreativitás vagy a tanulás átvihetőségét az "intelligens gépek" vitájára. Valójában még mindig fontos az ember és a gép közötti antropológiai különbség kritikus kezelése, a mesterséges intelligencia attribúcióinak és önleíró gyakorlatának dekonstruálása, ahogyan azt *Anne Dippel* és *VN Alexander is* bemutatják hozzászólásaikban. *Anne Dippel* esszéjében három önálló kommentárt kombinál, amelyek mindegyike a mesterséges intelligencia egy-egy más aspektusával foglalkozik, és mindegyike más-más alapvetés körül forog.

4 Bernhard Dotzler német médiatudós már az 1980-as évek végén azt írta, hogy a mesterséges

intelligenciával kapcsolatos összes ismert előrejelzés már Turing írásaiban megtalálható (1989).

- 5 A Carnegie Mellon Egyetem úgynevezett Navlab csoportja például 1984 óta épít ro- bot járműveket. A Carnegie Mellon volt az első egyetem is, amely ANN-t használt az önvezető autók fejlesztéséhez.

a metafora: intelligencia, evolúció és játék. Az első kommentárja egy auto-etnográfiai vignetta, amely keretet ad a mesterséges "intelligenciáról" és a gépek állítólagos "gondolkodási" képességéről való elmélkedésnek; mindkettő - ahogy Dippel állítja - feminista szempontból nagyon problematikus metafora, tekintettel az intelligens emberi lények (túlnyomórészt) női munkájára. A második betekintést nyújt jelenlegi etnográfiai terepmunkájába, amelyet nagyenergiájú fizikusok körében végez, akik napi munkájuk során gépi tanulási módszereket használnak, és a darwinista metaforának engednek, amikor az evolúciós algoritmusok jelentőségét képzelik el az emberiség jövője szempontjából. A harmadik kommentár a "játészó" algoritmusokat vizsgálja, és az "idegen" kategóriáját tárgyalja, amelyet - bár az antropológia területén ellentmondásos - sokkal alkalmasabbnak tart a mesterséges intelligencia megértéséhez, mint a közvetlen megszemélyesítés, amely egy nem emberi lényt kelt életre. *VN Alexander* viszont azt hangsúlyozza szövegében, hogy nincs bizonyíték arra, hogy az AI-rendszerek valóban képesek lennének "bizonyítékokon alapuló" döntéseket hozni az emberi viselkedésről. A mesterséges intelligencia fejlett statisztikákat használhat az általánosítások finomhangolására; de a mesterséges intelligencia egy megdicsőült aktuáriusi asztal, nem pedig egy intelligens ágens. E szkeptikus beszámoló alapján megvizsgálja, hogy Alan Turing 1952-ben bekövetkezett halálakor hogyan vizsgálta a biológiai intelligencia és a mesterséges intelligenciáról alkotott kezdeti elképzelése közötti különbségeket. Ennek megfelelően tanulmánya ezekre a különbségekre összpontosít, és korlátokat szab a jelenlegi mesterséges intelligencia legitím felhasználásának.

A jelenlegi mesterséges intelligencia-diskurzusok kritikai elemzése és a központi kon- A mesterséges intelligenciával kapcsolatos kérdésekben ugyanilyen fontos megérteni a média, az infrastruktúra és a technológiák azon csoportosulásait, amelyek lehetővé teszik és alakítják a mesterséges intelligencia használatát. E kihívásnak való megfelelés érdekében megfelelően figyelembe kell venni az érintett heterogén technológiák és alkalmazások sajátos jellegét és történelmi kialakulását (vö. Mckenzie 2017). *Axel Volmar* hozzájárulása "Productive Sounds: Touch-Tone Dialing, the Rise of the Call Center Industry and the Politics of Voice Assistants" című írása például a hangalapú asszisztensek és okos hangszórók - mint például az Amazon Alexa, az Apple Siri, a Google Assistant, a Mi- crosft Cortana vagy a Samsung Viv - növekvő elterjedésére reflektál, amelyek szavai szerint "a mesterséges intelligencia pusztá tömeges expozíció általi demokrációját" jelentik. A hangalapú asszisztensek, pontosabban a társalgási mesterséges intelligencia-technológiák politikájával úgy foglalkozik, hogy azokat a "produktív hangok" működésén alapuló távoli rendszerek hangalapú ember-gép interakciójának nagyobb történetéhez kapcsolja - a Touch-Tone-jelzéstől a várakozási zenén és az előre felvett üzeneteken át az interaktív hangreakciós (IVR) rendszerekig. Ebben a történetben a Volmar a telefonhoz és a hanghoz kapcsolódó munka és munkagyakorlatok változó formáira, valamint a

telefonos vagy más módon mediált beszéd automatizálásából és elemzéséből származó értékkivonás különböző formáira összpontosít. Azt állítja, hogy míg a háztartási és potenciálisan professzionális irodai végfelhasználók a webes keresés és a napi rutinok tekintetében a kényelem és a hatékonyság miatt elfogadják a hangalapú asszisztenseket, addig a vállalatok, a technológiai vállalatok, a felügyeleti

az államok és más szereplők célja, hogy hozzáférjenek magához a felhasználók hangjához, amelyet rendkívül értékes adatforrásnak - "aranybányának" - tekintenek a mesterséges intelligencia alapú elemzéshez.

Egy másik érdekes terület, ahol a mesterséges intelligenciát és különösen a gépi tanulási technikákat egyre gyakrabban alkalmazzák, a pénzügyi piac és az algoritmikus kereskedés különböző formái. Amint *Armin Beverungen* cikkében bemutatja, a pénzügyi kereskedést már régóta az adatfeldolgozás és a számítáció rendkívül kifinomult formái uralják a "kvantok" uralmában. Az elmúlt két évtizedben azonban a nagyfrekvenciás kereskedés (HFT), mint az automatizált, algoritmikus kereskedés egy olyan formája, amely inkább a sebességre és a volumenre, mint az okosságra összpontosít, uralta a pénzügyi piacok fegyverkezési versenyét. Beverungen szerint a gépi tanulás és a mesterséges intelligencia ma megváltoztatja ennek a fegyverkezési versenynek a kognitív paramétereit, és elmozdítja a határokat a HFT "buta" algoritmusai és az algoritmikus kereskedés más formáiban alkalmazott "intelligens" algoritmusok között. Míg a HFT nagyrészt a pénzügyi piacokra jellemző adatokra és dinamikára összpontosít, addig a mesterséges intelligencia által lehetővé tett új algoritmikus kereskedési formák kibővítik a pénzügyi piacok ökológiáját olyan módokon, amelyekkel az automatizált kereskedés szélesebb körű adatokra (például társadalmi adatokra) támaszkodik olyan elemzésekhez, mint a hanguletelemzés. Beverungen szerint e változások politikájának megértéséhez érdemes a megismerésre, mint a pénzügyi piacok csataterére összpontosítani, mivel az AI és a gépi tanulás a megismerés további újraelosztásához és új időbeliségéhez vezet. A megismerés politikájának meg kell küzdenie a pénzügyi piacokon zajló algoritmikus kereskedés homályosságával és időbeliségével, amelyek korlátokat jelentenek a pénzügyek demokratizálásának és társadalmi szabályozásának.

A globális mesterséges intelligencia-infrastruktúrák politikai dimenziójának megvilágítása érdekében,

nem csak azt kell megvizsgálunk, hogy a technológiai óriások hogyan használják a mesterséges intelligenciát a magánszektorban, hanem azt is figyelembe kell vennünk, hogy a közszféra egyre inkább az adatvezérelt válaszra törekszik, jobb és személyre szabottabb szolgáltatások nyújtását, a bürokrácia hatékonyságának növelését és a polgárok felhatalmazását ígérve. *Lisa Reutter* és *Hendrik Storstein Spilker* például Norvégia példáján keresztül tárgyalja a közszférában a mesterséges intelligencia alapú szolgáltatások előállításával kapcsolatos korai kihívásokat, és megvizsgálja, hogy ezek a kihívások hogyan tükrözik azokat a bizonytalanságokat, amelyek a mesterséges intelligenciával kapcsolatos hype mögött rejlenek a közszolgáltatásban. A Norvégiai Munkaügyi és Jóléti Igazgatóság adattudományi környezetével való etnográfiai találkozón keresztül fejlesztük a gépi tanúlással kapcsolatos hétköznapi munkára és az adatgyűjtés és -szervezés folyamataira összpontosít. Amint azt bemutatják, a gépi tanulási modellekbe táplálandó adatokkal kapcsolatos döntések ritkán egyszerűek, hanem a hozzáférési korlátozások, a kontextusfüggőségek és az elégtelen jogi keretek kezelését is magukban foglalják. Amint Reutter és Spilker bemutatják, az adatvezérelt norvég

közsféra így sok szempontból egy jövőbeli képzeletbeli elképzelés, gyakorlati jelenbeli iránymutatások nélkül.

A különböző mesterséges intelligencia-jelenségek sajátosságainak kritikus kezeléséhez elengedhetetlen a megfelelő utak, fogalmak és kritikai szintek feltárása. Kant óta a kritika a jelenségek megkérdőjelezését jelenti funkciójuk tekintetében.

és lehetőségük feltételei. Foucault szerint a kritika úgy is felfogható, mint az az erőfeszítés, sőt művészet, hogy megtaláljuk a módját annak, hogy "ne így és ilyen áron kormányozzanak" (Foucault [1997/1978]: 45). A kritika egy további koncepciója viszont arra törekszik, hogy a társadalom idealisztikus elképzeléseit a valós viszonyokkal összevetve vizsgálja meg, és feltárja, hogy miért és mennyiben maradhatnak el (szükségszerűen) ezek a társadalmi ideálok (vagy nem). Marx számára a kritikának ez a formája annak elemzését vonta maga után, hogy miért szembesülünk az illúzió és a hamis tudat szükségszerű előállításával, és ennek a fókusznak érezte magát Adorno és Horkheimer is a *Felvilágosodás dialektikája* (1944/1972) kritikai elemzésében.

Természetesen ezek csak néhány lehetséges kritikai gondolkodási irányok közül, amelyek hasznosak az egyre inkább mesterséges intelligencia által vezérelt világ mélyreható vizsgálatához. Ezen túlmenően szem előtt kell tartanunk, hogy a mesterséges intelligencia új konstellációkat és konfigurációkat biztosít a társadalmi-technológiai összeállítások számára, amelyeket nem feltétlenül lehet megfelelően vizsgálni a régi kritikai koncepciók szemüvegén keresztül, ahogy Geert Lovink az internet és a közösségi média technológiáival kapcsolatban kifejtette (2011: 88).

Ezért fontos, hogy megkérdőjelezzük a kritikai elemzés azon fogalmait, amelyeket a digitális kultúra megértéséhez használunk. *Tobias Matzner* szövege például a mesterséges intelligencia jelenlegi alkalmazásaival kapcsolatos néhány kiemelkedő kritikai állásponttal foglalkozik. Különösen azokat a megközelítéseket tárgyalja, amelyek a szubjektivitás változásaira mint a kritika útjára összpontosítanak, nevezetesen Wendy Chun és Antoinette Rouvroy. Míg Rouvroy általános ítéletet mond az általa "algoritmikus kormányzásnak" nevezett jelenség ellen, Chun azt javasolja, hogy a szubjektivitás konfigurációit a digitális technológia segítségével "lakjuk be". *Matzner* szövege e két álláspont közötti középutat célozza meg az érintett szubjektumok konkrét helyzetének kiemelésével. Ehhez a célhoz Linda Martin Alcoffnak a habituálissal mint szituált szubjektivitással kapcsolatos munkája kapcsolódik a médiaelmélet reflexióival. Végezetül a szituált szubjektumok e perspektívája a mesterséges intelligencia technológiák demokratikus konfigurációjának kérdéséhez kapcsolódik.

A mesterséges intelligencia kritikájának kérdése aligha érinti kevésbé a megfelelő méretezés problémáját. *Jonathan Roberge*, *Kevin Morin* és *Marius Senneville* fejezetében a szerzők azt állítják, hogy a mesterséges intelligencia kultúrájával kapcsolatos makroszintű kérdések és a mélytanulási technikákon belüli feltárhatatlanság mikroszintű kérdéseinek összekapcsolásához egy harmadik elemzési szintre van szükség. Ezt a mezzo-szintet "kormányozottságnak" nevezik, vagyis azt tárgyalják, hogy a hatalmi viszonyokat és a hatalom eloszlását a területen belül hogyan alakítja sajátosan a szervezetek és intézmények struktúrája. A montreali központot esettanulmányként véve - és a 2016-2018-as etnográfiai munkájuk alapján - két, egymással összefüggő kérdésre

összpontosítanak: a) a mélytanulásban implikált magán-állami partnerség újradefiniálására, valamint b) a jelenleg divatos "nyílt tudományos modell" következményeire.

Figyelembe kell vennünk továbbá, hogy az intelligens gépek legújabb fejlesztései tükrözhetnek néhány általános változást és folytonosságot az ember-gép kapcsolatok infrastrukturális és környezeti formálásában. A "Redukció

and Participation" című könyvében *Stefan Rieger* például egy figyelemre méltó stratégiával foglalkozik a médiakörnyezetben. Ez a test holisztikus felfogása és az összes érzékszervet - még az alsóbbrendű érzékszerveket is - bevonó megközelítés felé való elmozdulás. Rieger szerint mindenekelőtt ezek az érzékek játszanak döntő szerepet a mindenütt jelenlévő természetesség v á l á s során. Ennek következménye a technológiai evolúció és annak ellenállhatatlan sikere, amely az aug- mentáció és a terjeszkedés jól ismert topikjaitól eltérő történetet követ. Szembetűnő a technikailag lehetséges magas komplexitás szándékos csökkentése. Befolyásolják az internetpolitika, a demokratizálódás és annak a kérdésnek a szempontjai, hogy kinek kell egyáltalán hozzáférnie (és milyen módon) a médiakörnyezetekhez. A "redukció és részvétel" megfelel a más fajok és létformák bevonására irányuló igényeknek. Az ilyen igények célja, hogy bővítsék azok körét, akik cselekvőképességre és episztemikus relevanciára tesznek szert, ami magukat az algoritmusokat is érinti, ahogy Rieger érvel.

Az ügynökség és az episztemikus relevancia kérdése arra emlékeztet bennünket, hogy magának a mesterséges intelligencia kritikájának is fontos története van, amelyet figyelembe kell vennünk. A mesterséges intelligencia fejlődését ugyanis mindig is kísérte a politikai, társadalmi vagy gazdasági dimenziókra és ellentmondásokra vonatkozó kritikai reflexió. És gyakran maguk az informatikusok és mérnökök voltak azok, akik a kritika e különböző formáit megfogalmazták.

Norbert Wiener kibernetikus például már 1950-ben megjegyezte:

Ne feledjük, hogy az automata gép, bármit is gondoljunk az érzelmeiről, amelyekkel esetleg rendelkezik vagy nem rendelkezik, a rabszolgamunka pontos gazdasági megfelelője. Minden olyan munkának, amely a rabszolgamunkával versenyez, el kell fogadnia a rabszolgamunka gazdasági következményeit. Teljesen világos, hogy ez olyan munkanélküliségi helyzetet fog előidézni, amelyhez képest a jelenlegi recesszió és még a harmincas évek depressziója is kellemes viccnek fog tűnni. Ez a depresszió számos iparágat fog tönkretenni - talán még azokat is, amelyek kihasználták az új lehetőségeket. (Wiener [1988/1950]: 162)

Az egyik legintenzívebben tárgyalt AI-téma napjainkban az a spekulatív kérdés körül forog, hogy a robotok és intelligens gépek által vezérelt automatizálás mennyire vezet a munkaerőpiac felbolydulásához, és milyen mértékű munkahelyvesztést okozhat. Az AI-szakértők, például Kai Fu Lee úgy vélik, hogy a következő években a világ munkahelyeinek 40%-át az AI és a robotok helyezhetik át (Reisinger 2019; vö. még Frey/Osborne 2017). Az ilyen előrejelzések, bármennyire is nagy számban keringenek manapság, mindenekelőtt egy dolog: hol több, hol kevesebb jól megalapozott vagy megalapozott spekuláció. Hogy milyen lesz a világ 15 év múlva, azt sem okos tudósok, sem intelligens gépek nem tudják megjósolni. Mindazonáltal Norbert Wiener idézete legalább azt illusztrálja, hogy a kritika és a spekuláció kéz a kézben jár, akkor is

Hasonlóképpen, Joseph Weizenbaum számos kritikái pontja, amelyet a *Computer Power and Human Reason* (1976) című alapvető művében fogalmazott meg, reneszánszát éli a mesterséges intelligenciáról szóló jelenlegi vitákban. Weizenbaum könyve esetében kritikái beavatkozása kettős volt: Egyfelől az ember és a gép és/vagy a gondolkodás/ítéletealkotás és a számítás közötti alapvető különbségek hangsúlyozása motiválta, beleértve annak bizonyos alapvető korlátainak kiemelését, hogy mire lehet képes a mesterséges intelligencia; másfelől Weizenbaum figyelmeztetett arra, hogy vannak olyan feladatok, amelyeket egy számítógép képes lehet elvégezni, de nem szabadna. Számos, Weizenbaum által tárgyalt téma és javasolt érv konkrétan visszaköszön és továbbfejlődik a "mesterséges intelligencia etikájáról" folyó jelenlegi vitákban (vö. Cows/Floridi 2018; Tad-deo/Floridi 2018). Weizenbaummal ellentétben azonban, akinek kritikái reflexiói lényegében a klasszikus szimbolikus mesterséges intelligenciára épültek, a mai mesterséges intelligencia etikai vita azzal a kihívással néz szembe, hogy megfelelően megértse a gépi tanulási rendszerek és a mesterséges neurális hálózatok médiumait, technológiáit és infrastruktúráit, amelyek működési logikája szignifikánsan különbözik attól, amit néha "jó öreg mesterséges intelligenciának" neveznek (Sudmann 2018b). Ez pedig különösen nehéz feladat, hiszen az ANN marginális státusából adódóan a mesterséges intelligenciának ezen a sajátos területén nincs mélyreható szakértői hagyomány, sem a bölcsészet- és társadalomtudományok számos diszciplínájában, de még a természet- és műszaki tudományokban sem (vö. még az *Alexander Waibelrel* készült interjú ebben a kötetben).

Ezen túlmenően, az AI-boom kezdete óta számos vezető rekeresők feladták egyetemi oktatói vagy alkalmazotti állásukat, vagy szabadságot vettek ki, hogy startupokat alapítsanak, vagy a nagy technológiai óriásoknak dolgozzanak. Egyrészt csábítóak a hatalmas fizetési lehetőségek (akár alkalmazottként, akár egy start-up alapítójaként), másrészt sok tudós azért is vállal munkát a nagy technológiai cégeknél, mert feltételezik, hogy a kutatásaik feltételei az üzleti életben lényegesen jobbak, mint az egyetemen (például a tanulási adatokhoz vagy a nagy teljesítményű számítógépekhez való hozzáférés, az újrakereséshez szükséges forrásokhoz való hozzáférés tekintetében).

A legtöbb vállalat és különösen az elmúlt években a mesterséges intelligencia boomja nyomán alapított számtalan start-up vállalkozás is folyamatosan panaszkodik a terület szakértőinek hiányára, amit a további innovációk jelentős fékező tényezőjének tartanak. Számos intézmény felismerte ezt a problémát, és milliárdokat fektet be az AI képzésébe, kutatásába és fejlesztésébe. Mindazonáltal felmerül a kérdés, hogy milyen szempontok szerint, milyen célokkal és milyen feltételek mellett történik ez a finanszírozás. Ennek fényében elengedhetetlen, hogy a mesterséges intelligencia magán- és állami finanszírozása magában foglalja a kritikus kutatások támogatását is. Ez utóbbi mindenekelőtt a bölcsészet-

és társadalomtudományok feladata. Ahhoz azonban, hogy ezt a feladatot megfelelően el tudják látni, a "kemény tudományokkal" folytatott párbeszédre és együttműködésre van szükségük.

Van azonban egy másik oka is annak, hogy a jelenlegi mesterséges intelligencia-technológiák kutatása, különösen a politikai dimenziójukat illetően, olyan nagy kihívást jelent, amelyet még a szakértők sem tudnak könnyen leküzdeni. Amint arról az elmúlt években sokat vitatkoztunk, különösen az ANN-t a mesterséges intelligencia alapvetően átláthatatlan technológiájának tekintik. Míg az informatikusok valójában képesek megfigyelni és mérni az egyes neuronok és kapcsolataik aktivitását, függetlenül azok számától, addig az ANN aktivitását nem vagy csak korlátozottan tudják megérteni vagy megmagyarázni (vö. a kötethez való hozzájárulásomat is). Nyilvánvaló, hogy ennek a sajátos "fekete doboz" problémának komoly politikai-etikai következményei és hatásai vannak. Az például egy dolog, hogy a mesterséges intelligencia technológiákat például középkori kézírások felismerésére vagy bizonyos termékek fogyasztóknak való ajánlására használják. Amikor azonban a mesterséges intelligencia technológiákat arra használják, hogy egy személy hitelképességét értékeljék, vagy hogy megjelenése és viselkedése alapján eldöntsék, hogy egy adott személy elkövethet-e egy bizonyos bűncselekményt, a helyzet nyilvánvalóan más.

Ahogy *Dan Mcquillan* érvel esszéjében, a mesterséges intelligencia politikai technológia, és mint ilyen, a megszorítások fenntartására használják, de politikáját technikai átláthatatlansága és az etikai narratíva elhomályosítja. A mesterséges intelligencia konkrét műveletei a statisztikai regresszió és az optimalizálás révén gondolkodás nélküliséget és episztemikus igazságtalanságot eredményeznek. Eközben a mesterséges intelligencia előrejelző osztályozásai a bürokratikus kormányzati mentalitást kiterjesztik a jövőre, amelyet meg akar előzni. Az AI azonban törékeny, és csak azt oldja meg, amit Bergson "kész problémának" nevezett. Mcquillan szerint úgy kell megközelítenünk az AI-t, hogy a statisztikai valószínűségekkel szemben a lehetséges oldalára álljunk. Cikke egy feminista és egy helyhez kötött megközelítést mutat be a nem elnyomó mesterséges intelligencia kifejlesztésére, valamint az ehhez szükséges kollektív közösségi és munkahelyi struktúrák formáit. *Yvonne Förster* hasonlóképpen problematizálja, hogy különösen a jelenlegi mesterséges intelligencia-alkalmazások fekete dobozként működnek anélkül, hogy képesek lennének számot adni a mögöttes okokról, és maguk a mögöttes oksági folyamatok is átláthatatlanok maradnak. Esszéjében a láthatatlanság és az átláthatatlanság fogalmát fenomenológiai szempontból tárgyalja, és a tapasztalat és az érzékelés technológiához való viszonyát vizsgálja.

A mesterséges intelligencia demokratizálása

A mesterséges intelligencia kritikájának hosszú hagyományához képest a "mesterséges intelligencia demokratizálása" viszonylag új diskurzus. Alapvetően azóta alakult ki ez a diskurzus, amióta széles körben ismertté vált, hogy a mesterséges intelligencia a globális kultúra és társadalom minden területére

30 Andreas Sudmann

beavatkozik. E diskurzus kialakulásához és elterjedéséhez többek között a következő szempontok járultak hozzá:

1. a mesterséges intelligencia technológiák széles körű kritikája társadalmi, gazdasági és politikai következményeik, megnyilvánulásaik és hatásaik tekintetében.
2. a mesterséges intelligenciáról alkotott disztópikus képzelgések hosszú hagyománya
3. a nagy technológiai vállalatok adatosítási és adatelemzési gyakorlata és hegemon szerepük a mesterséges intelligencia jelenlegi és jövőbeli fejlődésében
4. az ANN mint az információfeldolgozás és az adatelemzés alapvetően átláthatatlan technológiájának értékelése

Az olyan kifejezéseket, mint a demokratizálódás és a demokrácia, néha úgy használják, mintha mindig pozitívan vagy pozitívan lehetne rájuk hivatkozni. Ugyanakkor a de- mokrácia elméletei folyamatosan emlékeztetnek minket arra, hogy a demokrácia eszméje, amely a "nép uralmát" jelenti, mindenkor jelentős kizárásokat feltételezett. Az ókori görög poliszokban csak a szabad polgárok - de nem a nők, a rabszolgák vagy azok, akiknek nem volt földjük - szavazhattak és politizálhattak. A kirekesztésnek ez a hagyománya még sokáig fennmaradt. John Locke felfogása szerint, amely meghatározó volt az angol parlamentarizmus kialakulásában, a választójogot még mindig csak a birtokosok kapták meg, és persze nem szabad elfelejtenünk, hogy a demokratikus társadalmakban a nők még jóval a 20. században sem vehettek részt a választásokon. Még ma is előfordul, hogy azok az emberek, akik hosszú évek óta élnek egy adott országban, bár elvileg annak minden törvénye alá tartoznak, ki vannak zárva az országos választásokból, hacsak nem rendelkeznek a szükséges állampolgársággal.

Ahogy az elején már felidéztük, az AI-technológia már segített a politikusok megválasztásában. Ennek fényében kézenfekvő feltenni a kérdést, hogy vajon a gépek maguk is szavazhatnak-e majd, és ha igen, mikor, vagy általánosabban, hogy vajon az emberhez hasonlóan bizonyos jogokkal rendelkező entitásokként fogják-e őket felfogni, és ha igen, mikor. Elég figyelemre méltó, hogy bár a gépek (még) nem szavazhatnak, máris megválaszthatók - ahogy ez történt, amikor Japánban (Tama City, Tokió) 2018, egy mesterséges intelligenciával működő rendszer indult a polgármester-választáson. A szóban forgó AI-rendszer azt ígérte, hogy statisztikai módszereinek köszönhetően hatékonyan tudja értékelni a polgárok kéréseinek előnyeit és hátrányait; azt állította, hogy igazságos döntéseket hoz, konszenzusra törekszik az érdekellentétekben, és az adók felhasználása tekintetében is a teljes átláthatóságra összpontosít. A szavazatok összeszámlálásakor kiderült, hogy a mesterséges intelligencia rendszer az utolsó helyen végzett az összes jelölt közül. Az eredmény talán még a technológia megszállott Japánban sem meglepő. Az ottani emberek, akárcsak más országokban, elfogadják a mesterséges intelligencia-rendszereket és a robotokat eszközként, szolgálként vagy játékként, de nehéznek tűnik elképzelni a gépek politikai képviselőtét, hacsak nem nagyon disztópikus forgatókönyvekben.⁶

-
- 6 Bár nem csak Japánban, hanem Európában vagy az Egyesült Államokban is a gép jelenléte a normálisa kormányokban (vö. Agar 2003).

Valószínűleg éppen az a tény, hogy a mesterséges intelligenciáról alkotott kulturális képzeletvilágot a disztópikus narratívák oly nagymértékben alakították, még mindig arra készíti az embereket, hogy féljenek az intelligens gépekkel való együttéléstől, vagy legalábbis mélységes kényelmetlenséget érezzenek. Ennek fényében a mesterséges intelligencia demokratizálására irányuló, a következőkben ismertetett legújabb erőfeszítések valóban úgy értelmezhetők, hogy az emberek elképzelése szerint az emberek és a gépek közös jövőjének ilyen disztópikus elképzelése ellen dolgoznak.

Itt azonban fontos megjegyezni, hogy a mesterséges intelligencia demokratizálásának igénye elkerülhetetlenül magában foglalja, hogy az ilyen technológiák önmagukban nem demokratikusak, vagy legalábbis erős tendenciájuk vagy lehetőségük van arra, hogy összeegyeztethetetlenek legyenek a demokratikus értékekkel és gyakorlatokkal. És a mesterséges intelligenciának erre a felfogására jó okai vannak. Ha az intelligens gépek fejlesztésének célja az ember helyettesítésére vagy felülmúlására irányul, vagy ha az AI-t a gazdasági növekedés hajtóerejének és a hegemon geopolitikai hatalom biztosításának feltételének tekintik, akkor mindezekben az esetekben a technológiának prima facie semmi köze a demokratikus értékek, például az empatikus értelemben vett egyenlőség megteremtéséhez és védelméhez. Hasonlóképpen, az algoritmikus torzításokról szóló jelenlegi viták az egyenlőtlenség és a különbözőség alapvető problémáira mutatnak rá, amelyek a mesterséges intelligencia-rendszereknek a társadalom minden területén történő nagyszabású alkalmazásával kapcsolatosak.

Alexander Monea fejezete például azt vizsgálja, hogy a számítógépes látórendszerek sötétebb bőrszínnel rendelkező felhasználók számára való hozzáférhetővé tételére tett kísérletek hogyan vezettek a fenotípusos faji jellemzők, különösen az olyan morfológiai jellemzők, mint a hajszerkezet és az ajak mérete, vagy a faj láthatatlanságához. A kritikai fajelméletre és a faji reprezentáció problémás történetére támaszkodva bemutatja, hogy a faji előítéletek mennyire elterjedtek azokban a vizuális adathalmazokban, amelyekben számos kortárs számítógépes látó algoritmust képeznek, lényegében keményen kódolva ezeket az előítéleteket a számítógépes látási technológiákba, mint például a Google Photos. A leggyakoribb ipari reakció ezekre a keményen kódolt faji előítéletekre az, hogy a faji hovatartozást láthatatlanná teszik a rendszerben, ahogyan azt a Google Photos esetében is tették. Továbbá bemutatja, hogy a faj láthatatlansága a számítógépes látásban hogyan vezet a "színvakság"jól ismert problémáihoz, csak az új médiában fejeződik ki. A szerző amellet érvel, hogy ezek alapvető problémákat jelentenek a mesterséges intelligencia potenciális demokratizálása szempontjából, és felvázol néhány konkrét lépést, amelyeket megtehetünk annak érdekében, hogy erőteljesebben követeljük az egalitárius számítógépes látórendszereket.

Mindazonáltal legalábbis egyesek úgy vélik, hogy a mesterséges intelligencia önmagában is képes lehet a szabadság, egyenlőség és testvériség új utópikus horizontját megnyitni, sőt, akár a világbéke biztosítására is produktívan

felhasználható (vö. Valladão 2018). Thomas Hobbes államelméletében a Leviatán (mint a fiktív társadalmi és kormányzati szerződés megtestesítője) az emberek közötti békés együttélés lehetőségének szükséges feltételeként fogalmazódik meg. Enélkül az emberiség visszazuhanna a természeti állapotba, a mindenki mindenki ellen háborúba. A modern állam azonban, mint azt a Hobbes óta eltelt időszak történelme megmutatta, rendkívül bizonytalan, törékeny en-

titás, amely képtelen tartós védelmet nyújtani *minden tagjának*. Ennél is fontosabb a szomorú igazság, hogy mondjuk a francia forradalom óta soha nem volt olyan állam vagy demokrácia, amely teljes mértékben képes lett volna megfelelni a szabadság, az egyenlőség vagy a szolidaritás követelményeinek, azok hangsúlyos értelmében.

Ennek fényében nem tűnik teljesen abszurdnak az az utópisztikus (és egyesek számára bizonyára disztópikus) elképzelés, hogy a politikai felelősséget és a társadalom irányítását teljes egészében a gépekre ruházzák. De ha elképzelhető, hogy az emberiség úgy dönt, hogy inkább felsőbbrendű gépek kezébe helyezi a sorsát, akkor még az is elképzelhető, hogy az emberek is kiállhatnak vagy ki fognak állni egy adott ponton egy valóban jobb globális társadalom megvalósításáért - a mesterséges intelligencia mint politikai entitás vagy "békegép" segítségével (Honkela 2017).

A mesterséges intelligencia demokratizálására vonatkozó jelenlegi elképzeléseknek azonban kevés közülük van a társadalom alapvető átalakítását célzó kritikához. Mindazonáltal a demokratikus mesterséges intelligencia koncepciója, mint a jelen projektje, még mindig nagyon szorosan kapcsolódik az utópisztikus elképzelésekhez és motivációkhoz, mivel hasonlít a történelem során az ízlés, a művészet, a média, a technológia vagy a társadalom egésze kapcsán kialakult demokratizálási stratégiák és koncepciók sokaságára.

A mesterséges intelligencia demokratizálására vonatkozó jelenlegi elképzelések erős hasonlóságot mutatnak a kibertérrel, a virtuális valósággal és természetesen az internettel kapcsolatos utópikus-politikus elképzelésekkel, mivel ezek különösen az 1990-es évek eleje óta terjedtek el (vö. Eglhoff 2002). Az az elképzelés, hogy a kibertér és/vagy az internet (a fogalmak nem azonosak, mégis gyakran szinonimaként használják őket) önmagában egy emancipációs tér, amelyet régebben "cyber-utópiának" neveztek, és legkésőbb 1995a Critical Art Ensemble kritikájának tárgya. Ezzel szemben még ma is sok tudós, művész és net-aktivista ragaszkodik ahhoz az elképzeléshez, hogy akár az internet és/vagy a kibertér valójában a szabadság, a felforgatás és az ellenállás terét jelöli, amelyet minden heterogén ellentmondása és problémája ellenére meg kell védeni.

A demokratizálás utópisztikus-idealista dimenziója a fogalomnak a nagy technológiai vállalatok által a mesterséges intelligenciával kapcsolatban történő jelenlegi használatában is megmutatkozik. A "demokratikus mesterséges intelligencia" koncepcióját mindenekelőtt az egyetemes, mindenre kiterjedő hozzáférhetőség, részvétel és átláthatóság nagy ígéretként mutatják be. A Microsoft számára például az AI demokratizálása lényegében azt jelenti, hogy a technológiát "minden ember és minden szervezet" kezébe adják (vö. Microsoft News Center 2016; vö. Johnson 2017a).

Ami a technológiai óriások hivatalos programját illeti, jelenleg különböző stratégiákat követnek ennek a célnak az elérése érdekében: Először is, az egyik általános elképzelés az AI egyszerűsítésének, szabványosításának és

automatizálásának előmozdítása, hogy a vállalatokon és egyetemeken belül és kívül a nem szakemberek is egyre inkább használhassák a megfelelő technológiákat (például az ANN-t) saját céljaikra és alkalmazásaikra. Másodszor, a nagy informatikai vállalatok nyílt hozzáférést kívánnak biztosítani a felhasználóknak, tudósoknak és vállalatoknak a különböző felhőszolgáltatásokhoz, a számítási erőforrásoktól kezdve (mint például a Google

Tensor Processing Units, azaz speciális chipek a gépi tanulási műveletek gyorsítására), programkönyvtárak és keretrendszerek, mint a Scikit, PyTorch, Keras vagy Tensor-Flow, képzési adathalmazok, mint az MNIST vagy az ImageNet, valamint különféle egyéb szoftvereszközök, amelyek a mesterséges intelligencia szélesebb körű elterjedését és fejlesztését segítik.

Nem meglepő, hogy a vállalatok nem nyújtják szolgáltatásaikat egyéni érdekek nélkül, vagy anélkül, hogy valamilyen ellenszolgáltatást várnának: A Google például újra megköveteli, hogy az erőforrásait használó kutatók saját kutatási eredményeiket és esetleg kódjukat is nyílt forráskódúvá tegyék (Johnson 2017b). Ezen túlmenően feltételezik, hogy eszközeik nyílt forráskihelyezése azt is eredményezheti, hogy a független fejlesztők jelentős költségek nélkül járulnak hozzá azok fejlesztéséhez (vö. Lomonaco/Ziosi 2018). Ráadásul az olyan nagyvállalatok, mint a Microsoft, profitálnak abból, hogy maga a nyílt forráskódú ötlet nagy hírnévnek örvend a technológiai kutatói közösségben, és a kutatóknak érdekükben áll, hogy munkájuk jól látható és széles körben elismert legyen (vö. Bostrom 2017).

Marcus Burkhardt szövege kritikusan foglalkozik azzal az ígérettel, hogy a fejlesztők számára hozzáférést biztosít a megjelenő gépi tanulási technológiákhoz, és lehetővé teszi számukra, hogy alkalmazásaikat okossággal vagy intelligenciával ruházzák fel, és felteszi a kérdést, hogy a gépi tanulás és az AI mint a technológiai fejlesztés és innováció területei hogyan strukturálódnak önmagukban. A GitHubon a gépi tanulás és a mesterséges intelligencia kódolási kultúráinak első feltérképezésével amellet érvel, hogy a gépi tanulási technológiák társadalmi és kulturális következményeinek kritikus perspektíváinak kialakításához fontos, hogy közelebbről foglalkozzunk a kódkönyvtárak és a programozási keretrendszerek eddig nagyrészt elhanyagolt infrastruktúrális rétegeivel.

A mesterséges intelligencia megnyitásának különböző intézkedéseihez kapcsolódó bizonyos előnyökön túlmenően számos kutató, intézmény és vállalat hajlamos hangsúlyozni, hogy a problémák megoldása ezen a területen olyan kollektív erőfeszítés, amelyet nem lehet egyénileg megvalósítani, ezért szükséges az ötletek és módszerek minél szélesebb körű és nyílt megosztása. A gazdasági és tudományos verseny problémáit és ellentmondásait azonban ritkán vitatják meg. A felszínen úgy tűnik, hogy a mesterséges intelligencia kutatását alapvetően a következők vezérik a kötetlen idealizmus. A valóság azonban az, hogy a területet valóban a tehetségekért, a tőkékért és más "erőforrásokért" folytatott kemény nemzetközi verseny jellemzi. A nagy technológiai cégek napirendjének középpontjában pedig az a szívós küzdelem áll, hogy elsőként győzzék le a mesterséges intelligencia megoldatlan problémáit és/vagy érjék el a végső célt, az *általános mesterséges intelligencia*, az úgynevezett erős mesterséges intelligencia, azaz egy olyan gép létrehozását, amely képes az emberhez hasonló vagy annál jobb feladatokat megoldani vagy megtanulni.

Ez vonatkozik az úgynevezett nonprofit szervezetre is, amely még a márkanevébe is beépítette a "nyitottság" gondolatát: OpenAI. Amint azt a kötethez írt saját hozzájárulásomban bemutatom, az OpenAI valahogyan a jelenlegi "AI demokratizálási" hype avantgárdja volt, mégpedig azért is, hogy előtérbe helyezte a demokratikus értékek, például a hozzáférés, a részvétel és az átláthatóság iránti elkötelezettségét. De ha ex-amináljuk a szervezet eddigi tevékenységét, az OpenAI befektetése

inkább a mesterséges intelligencia alapvető technológiai problémáinak megoldásában való előrelépésről szólt, ahelyett, hogy arra összpontosított volna, hogyan lehetne a nyílt, demokratikus mesterséges intelligencia koncepcióját technológiailag és koncepcionálisan értelmes módon továbbfejleszteni.

Egyelőre úgy tűnik, ha kritikusan vizsgáljuk az olyan vállalatok retorikáját, mint a Google vagy a Microsoft, mintha a demokratikus mesterséges intelligencia ígérete már teljesült volna a hozzáférhetőségével. Különösen a technológia esetében azonban a demokratizálódás nemcsak a használathoz való hozzáférést jelenti, hanem az ellenőrzés lehetőségét is (vö. Lomonaco/Ziosi 2018). Hogy egy ilyen folyamatot lehet-e és hogyan lehet ésszerű módon megszervezni és alakítani, például állami felügyelet vagy más intézkedések révén, az még nyitott kérdés, és talán nem is lehet általánosságban megválaszolni. A döntő pont azonban az, hogy azoknak a vállalatoknak, amelyek a "mesterséges intelligencia demokratizálását" szorgalmazzák, *legalábbis elvben* hajlandónak kell lenniük arra, hogy korlátozzák szuverenitásukat és/vagy elfogadják más külső entitások beavatkozását.

Ez utóbbi azonban valószínűleg nem áll a nagy technológiai vállalatok érdekében. Sőt, az a tény, hogy a tech-óriások ilyen teljes mértékben magukévá teszik a "demokratikus mesterséges intelligencia" gondolatát, határozottan jelzi, hogy ez a koncepció mennyire nem fenyegeti gazdasági vagy kulturális hatalmukat, sőt, épp ellenkezőleg.

Mindazonáltal a mesterséges intelligencia demokratizálása, ahogyan azt a nagy technológiai csoportok szorgalmazzák, nem csak a hozzáférés, az átláthatóság és a részvétel ellentmondásos fogalmairól szól. A koncepció továbbá magában foglalja azt a célt is, hogy a "jó célokat", azaz a világ kisebb-nagyobb problémáinak megoldását szolgálja. A Microsoft "AI for Earth" kezdeményezése például a klímaváltozás elleni küzdelmet vagy az egészségügyi ellátórendszerben tapasztalható egyenlőtlenségek felszámolását célozza. Egy ilyen menetrend alapján természetesen kínos, hogy a Microsoftot nemrégiben azzal vádolták, hogy az ország Központi Katonai Bizottsága által ellenőrzött kínai Nemzeti Védelmi Technológiai Egyetem kutatóival együttműködve olyan mesterséges intelligencia-problémákon dolgoztak, amelyekről a kommentátorok úgy vélték, hogy az állami megfigyelési technológiák számára felhasználhatók. A Microsoft azzal utasította el ezeket a vádakokat, hogy a szóban forgó kutatási munkáknak ugyanannyi vagy kevesebb közük volt a megfigyeléshez, mint a WiFi-nek vagy a Windows operációs rendszernek. Emellett a vállalat rámutatott, hogy a nemzetközi együttműködés ilyen formái nagyon is jellemzőek az AI-kutatás területén.

Bármilyen is a helyzet ebben a konkrét esetben, egyértelmű, hogy különösen a mesterséges intelligencia területén mindig is nehéz volt különbséget tenni a katonai és a polgári felhasználás között.⁷ Például a mesterséges intelligencia más alkalmazási területeihez hasonlóan a gépi fordítás és a természetes nyelvi feldolgozás területén végzett kutatások jelentős része (a mesterséges intelligencia

e területének politikai vitájához lásd az Alexanderrel készített interjút).

-
- 7 Néha a vállalaton belül is viták tárgyát képezi, hogy a katonaságtól érkező parancsokat el kell-e fogadni. Vö. a Google alkalmazottainak közelmúltbeli tiltakozását az úgynevezett "Maven projekt" ellen (vö. vö. Shane/Wakabayashi 2018). A mesterséges intelligencia katonai felhasználásáról szóló friss vitát lásd még Ernst/Schröter/Sudmann (2019).

Waibel ebben a kötetben), a hadsereg finanszírozta, különösen a DARPA által támogatott programok. Ez az elkötelezettség nem véletlen. Különösen a hidegháború idején volt nagy igény az oroszról angolra történő fordításokra (és fordítva a szovjetek részéről az angolról oroszra történő fordításokra). Továbbá a globális katonai műveletek és a katasztrófavédelmi erőfeszítések mindig is általános érdeklődést keltettek a nagy mennyiségű idegen nyelvű szövegek gyors fordítása iránt. Végezetül meg kell jegyezni, hogy a gépi fordítás területe kezdettől fogva alapvető matematikai és kriptológiai ismereteken alapult, amelyeket a második világháború alatt a katonai és titkosszolgálati kutatók fejlesztettek ki.

Ahogy a mesterséges intelligencia katonai célokra való felhasználása mutatja, a "nyitottság" és az "átláthatóság" önmagában nem számíthat pozitív értéknek. Nick Bostrom (2017) szerint a biztonsági intézkedésekkel kapcsolatos nyitottság vagy a célokkal kapcsolatos nyitottság lehet jó, de a forráskóddal kapcsolatos nyitottság nem feltétlenül kell, hogy az legyen. Ennek megfelelően Bostrom a "nyílt mesterséges intelligencia" differenciált megközelítése mellett érvel: Amikor olyan technológiák fejlesztéséről van szó, amelyek jelentős károkat okozhatnak, akkor természetesen nem szabad nyilvánosságra hozni őket.

Különösen az ANN-technológiával kapcsolatban merül fel az alapvető kérdés, hogy az átláthatóság és a nyitottság követelményei milyen mértékben valósíthatók meg egyáltalán, mivel az ANN konnekcionista megközelítését alapvetően átláthatatlannak kell tekinteni (Sudmann 2017). Mindazonáltal az úgynevezett "megmagyarázható mesterséges intelligencia" különböző megközelítései legalábbis megpróbálják csökkenteni a jelenlegi mesterséges intelligencia-rendszerek átláthatatlanságát.

Amint Schieferdecker, Großmann és Schneider a kötethez való hozzájárulásukban hangsúlyozzák, a különböző feladatokra mesterséges intelligencia módszereket alkalmazó szoftveralapú rendszereket alapvetően a "kritikusságuk" jellemzi, ami alatt a biztonság és a védelem szempontjából kritikus területeken való felhasználásukat értik, mint például a közlekedés és az autópálya, a bankok és a pénzügyek, az egészségügy, a kiberbiztonság vagy az ipari automatizálás. Amint azt a szerzők kifejtik, a mesterséges intelligencia-alapú rendszerek kritikussága szigorú és hatékony minőségfejlesztést igényel a telepítés előtti fázisokban és a futásidőben. Cikkükben a szerzők áttekintik a mesterséges intelligencia-alapú rendszerek úgynevezett "verifikációs és validálási módszerekkel" történő védelmének jelenlegi állását, különös tekintettel a mesterséges intelligencia-alapú rendszerek fő funkcionális összetevőire és azok kiterjesztett minőségi követelményeire. Mivel minden mesterséges intelligenciát elsősorban szoftverben fejlesztenek, ezért áttekintik a szoftveralapú rendszerek minőségmérnöki tervezésének elvi megközelítését általánosságban. Schieferdecker, Großmann és Schneider szerint a tesztelés a legismertebb és leghatékonyabb V&V módszer, és valószínűleg a mesterséges intelligencia alapú rendszerekkel való foglalkozás alapját is képezi: Használható a mesterséges

intelligencia-alapú rendszerek eredményeinek megerősítésére vagy tanúsítására, digitális központtá válhat azok összehasonlításához és benchmarkingjához, és így hozzájárulhat a mesterséges intelligencia közös tudásalapjához.

Az eddig vázolt jelenségek fényében nyilvánvaló, hogy a mesterséges intelligencia politikai gazdaságtana nagy kihívást jelent a döntéshozók számára. Amint azt *Frank Pasquale* esszéjében bemutatja, az úgynevezett centralizátorok a nagyon nagy cégek adathalmozását ösztönzik, míg ezzel szemben a decentralizátorok a decentralizáltabb innovációt szeretnék. Bár mindkettőnek nagyon eltérő elképzelései vannak a hosszú távú öko- nomikus fejlődésről, mindkettő segíthet ellensúlyozni a digitális kapitalizmus híveinek korlátlan törekvéseit (és kiábrándító mindennapi valóságát). Hozzájárulnak ahhoz is, hogy megértsük, amikor az óriáscégek megpróbálják megoldani azt, amit Friedrich Hayek "tudásproblémaként" azonosított - és amikor a homály és a ködösítés révén súlyosbítják azt. Ha a konglomeráció és a vertikális fúziók valóban olyan mesterséges intelligenciát mozdítanak elő, amely valós problémákat old meg - gyorsabb közlekedést, jobb élelmiszert, magas színvonalú egészségügyi ellátást és egyebeket -, akkor a hatóságoknak hagyniuk kellene, hogy folytassák. Pasquale szerint az ipari nagyság segít abban, hogy jobban megértsük és irányítsuk a természeti világot. Ugyanakkor amellet érvel, hogy az államoknak meg kellene akadályozniuk az alkupozíció és a tőkeáttétel pusztá felhalmozását, még akkor is, ha az az AI fejlesztését szolgálja. A politikai döntéshozóknak meg kell találniuk a módját, hogyan kezeljék a technológia szabályozásával kapcsolatos ellentmondásokat és eltérő felfogásokat. Az egyik fontos feladat itt az, hogy a politikai döntéseket olyan törvényekké alakítsák, amelyek a gyakorlatban megfelelőek, de figyelembe veszik e technológiák kritikáját is. De milyen szerepet játszhatnak a törvények a mesterséges intelligencia demokratizálásában? Ezzel a kérdéssel foglalkozik *Christian Djefall* fejezete. Szövege rávilágít a mesterséges intelligencia nyitottságának dimenzióira, és bemutatja, hogy a mesterséges intelligencia előnyös és hátrányos is lehet a demokrácia számára. Az alkotmányjog valójában a mesterséges intelligencia demokratizálására szólít fel. Az AI-rendszerekre való támaszkodás és az AI-rendszerekre való delegálás demokratikus egyensúly helyreállítását igényli. A fejezet ezután azt vizsgálja, hogyan lehet a mesterséges intelligenciát demokratizálni. Három réteget határoz meg, amelyek egy sor választási lehetőséget írnak le: a technikai, a társadalmi és az irányítási réteget. A technikai szinten számos választási lehetőség van; egy olyan konkrét koncepció, mint a tervezhetőség, segíthet a demokratikus kormányzást lehetővé tevő választások meghatározásában. A mesterséges intelligencia rendszerek befolyása gyakran nem a technológiában gyökerezik, hanem a társadalmi döntéseken keresztül tulajdonítják a mesterséges intelligenciának. A közigazgatási jogban az automatizált döntéseket a jog erejével ruházzák fel. A kormányzási réteg megmutatja, hogy a technológiákat hogyan befolyásolhatják átfogó döntések. Ez történhet például a keretek és a szervezés révén. Az összes réteget együttvéve bőven van hely a mesterséges intelligencia alkalmazásainak demokratikus meghatározására.

Talán vita tárgyát képezi, hogy a gépi tanulási algoritmusok milyen mértékben mint "kulturális gépek" (Finn 2017) már most is hatással vannak mindennapi életünkre, és megváltoztatták a világban szerzett szociokulturális

tapasztalatainkat. A gépi tanulás és az ANN kulturális hatásáról szóló vita az utóbbi időben felerősödött annak a kérdésnek a körüljárása körül is, hogy a mesterséges intelligencia hogyan tekinthető kreatívnek, és talán még a művészetről (gyakorlatokról) alkotott felfogásunkat is megváltoztatja. Az erről szóló nyilvános vitákat a Christie's aukciója táplálta, ahol egy mesterséges intelligencia-rendszer segítségével "létrehozott" festményt magas áron adtak el. Érdekes módon a művészeti kollektíva, amely a

ez a festmény azt állította, hogy "a művészetet keresztül akarják elmagyarázni és demokratizálni a mesterséges intelligenciát". Valószínűleg előre látható volt, hogy az AI-t a művészeti világ is gyorsan magáévá teszi. A "kreatív mesterséges intelligenciáról" szóló legújabb viták azonban hajlamosak figyelmen kívül hagyni, hogy a gépek állítólagos kreativitásának problémája korántsem új keletű, ahogy azt *Jens Schröter is* bemutatja cikkében. Valójában már az 1960-as években, az úgynevezett "információs esztétikában" hasonló kérdéseket vitattak meg. Esszéjében Schröter tehát historizálja a jelenlegi vitákat, és amellett érvel, hogy a kérdésre, hogy lehetséges-e a gépi kreativitás vagy a gépi művészet, nem lehet az "ember" és a "gép" (AI) absztrakt szembeállításával válaszolni.

A művészet, a kreativitás és az intelligens gépek közötti kapcsolat azt mutatja, hogy a mesterséges intelligencia politikájáról és demokratizálásáról szóló vita nem korlátozódhat bizonyos területekre (gazdaság, katonaság) vagy a szereplők bizonyos csoportjaira (pl. "A nagy technológiai óriások"). Ehelyett figyelembe kell vennünk, hogy a mesterséges intelligencia kritikáját és a demokratizálása iránti elkötelezettséget számos civil szervezet, tudományos intézmény, jogász vagy politikus is támogatja; olyan szereplők, akiknek erőfeszítései kétségtelenül megérdemlik a saját portréjukat. Ez a könyv tehát csak egy kis hozzájárulás egy olyan ellentmondásos vitaterülethez, amelynek körvonalait, viszonyait és feltételeit még fel kell tárni.

Hivatkozások

Algar, Jon (2003): *A kormánygépezet. A számítógép forradalmi története.*

Cambridge, MA: MIT Press.

Birch, Anthony H. (1993): *A modern demokrácia fogalmai és elméletei.* London et al.

Routledge.

Bostrom, Nick (2017): Bostrom: "A nyitottság stratégiai következményei a mesterséges intelligencia fejlesztésében". In:

Global Policy, 1-12. (<https://nickbostrom.com/papers/openness.pdf>).

Bublies, Pia (2017): Bublies: "Az adatok új gazdaságot teremtenek." In: *The Economist*, május 6. (<https://www.economist.com/briefing/2017/05/06/data-is-giving-rise-to-a-new-economy>).

Cowls, Josh és Floridi, Luciano (2018): "Prolegomena to a White Paper on an Ethical Framework for a Good AI Society". Június (<https://ssrn.com/abstract>).

=3198732 vagy <http://dx.doi.org/10.2139/ssrn.3198732>).

Diamond, Larry Jay és Plattner, Marc F., szerk. (2012): *Felszabadító technológia: Social Media and the Struggle for Democracy.* Baltimore, MD: Johns Hopkins University Press.

Dotzler, Bernhard (1989): Dotzler, Dotzler: "Tudás/élmény: Versuch über die Verortung der Künstlichen Intelligenz". In: *MaschinenMenschen.* Katalog zur Ausstellung des Neuen Berliner Kunstvereins, 17.-23.07. Berlin: NBK: 127-132.

46 Andreas Sudmann

- (2006): *Diskurs und Medium. Zur Archäologie der Computerkultur*. Band 1.
München: Fink.

- Egloff, Daniel (2002): *Digitale Demokratie: Mythos oder Realität*. Wiesbaden: Westdeutscher Verlag.
- Ernst, Christoph, Schröter, Jens és Sudmann, Andreas (2019): "AI and the Imagination to Overcome Difference." *spheres - Journal for Digital Cultures* 6 (2019): megjelenés alatt.
- Finn, Ed (2017): *What Algorithms Want. Képzlet a számítástechnika korában*. Cambridge, MA: MIT Press.
- Foucault, Michel (1997 [1978]): "Mi a kritika?" In: *What is Critique? Semiotexte*: Los Angeles, CA, 41-81.
- Frey, Carl Benedikt és Osborne, Michael A. (2017): "A foglalkoztatás jövője: Technological Forecasting and Social Change, Elsevier, vol. 114(C): 254-280. (<https://ideas.repec.org/a/eee/tefoso/v114y2017icp254-280.html>).
- Goodfellow, Ian, Bengio, Yoshua és Courville, Aaron (2016): *Deep Learning*. Cambridge; London: MIT Press.
- Hebb, Donald Olding (1949): Hebb Hebb: *The Organization of Behavior. A Neuropsychological Approach*. New York: John Wiley & Sons.
- Honkela, Timo (2017): *Rauhankone*. Gaudeamus.
- Horkheimer, Max és Adorno, Theodor W. (1944/1972): *A felvilágosodás dialektikája*. New York: Herder and Herder.
- Hughes, Chris (2019): "It's Time to Break Up Facebook". In: *New York Times*, május 9. (<https://www.nytimes.com/2019/05/09/opinion/sunday/chris-hughes-facebook-zuckerberg.html>).
- Johnson, Khari (2017a): "A mesterséges intelligencia demokratizálása a technológiai óriásokon múlik". In: *VentureBeat*. (<https://venturebeat.com/2017/12/28/ai-weekly-ai-democratization-depends-on-tech-giants/>).
- Johnson, Khari (2017b): "Google unveils second-generation TPU chips to accelerate machine learning". In: *VentureBeat* (<https://venturebeat.com/2017/05/17/google-unveils-second-generation-tpu-chips-to-accelerate-machine-learning/>).
- McCulloch, Warren S. és Pitts, Walter (1943): "A Logical Calculus of the Ideas Immanent in Nervous Activity". In: *The Bulletin of Mathematical Biophysics* 5.4: 115-133. DOI: <https://doi.org/10.1007/BF02478259>.
- Locke, William N. és Booth, A. Donald, szerk. (1955): Booth, Booth, Booth, Booth: *Machine Translation of Languages*. Cambridge, MA et al: The Technology Press of The Massachusetts Institute of Technology, John Wiley & Sons, John Wiley & Sons, and Chapman & Hall.
- Lomonaco, Vincenzo és Ziosi, Marta (2018): "A mesterséges intelligencia demokratizálódásának mítoszáról". (<https://medium.com/ai-for-people/on-the-myth-of-ai-democratization-a472115cb5f1>).
- Lovink, Geert (2003): *Lovink: Dark Fiber-Tracking Critical Internet Culture*. Cambridge, MA: MIT Press.

- Lovink, Geert (2011): *Networks Without a Cause. A közösségi média kritikája*. London: London: Polity Press.
- Marcus, Gary (2018): "Deep Learning. A Critical Appraisal." (<https://arxiv.org/abs/1801.00631>).
- Mackenzie, Adrian (2017): *Machine Learners. Az adatgyakorlat archeológiája*. Cambridge, MA: MIT Press.
- Mckenzie, Adrian (2017): *Machine Leaners. Egy adatgyakorlat régészete*. Cambridge, MA: MIT Press.
- Mensch, Kirsten és Schmidt, Jan C., szerk. (2003): *Technik und Demokratie. Zwischen Expertokratie, Parlament und Bürgerbeteiligung*. Opladen: Leske+Budrich.
- Michie, Donald (1971): Findler & B. Meltzer (szerk.): "Formation and Execution of Plans by Machine", in N.V. Findler & B. Meltzer, szerk: *Artificial Intelligence and Heuristic Programming*: New York: American Elsevier: 101-124.
- Microsoft News Center (2016): (<https://news.microsoft.com/features/democratizing-ai/>)
- Mitchell, Thomas (1997): *Gépi tanulás*. New York: McGraw-Hill.
- Pasquinelli, Matteo (2017): "Logikát morfondírozó gépek: Neural Networks and the Distorted Automation of Intelligence as Statistical Inference". In: *Üveggyöngy folyóirat*, 1. oldal, "Logikai kapu: A mesterséges elme politikája".
- Rapoza, Kenneth (2019): "Demokráciák válságban: Feladta-e a Nyugat a demokráciát?" In: *Forbes*, január (9.<https://www.forbes.com/sites/kenrapoza./2019/01/09/demokraciak-valsagban-a-nyugat-feladta-a-demokraciat/#425fb6db1242>).
- Reisinger, Don (2019): "A.I. Expert Says Automation Could Replace 40% of Jobs in 15 Years." In: *Fortune*. (<http://fortune.com/2019/01/10/automation-replace-jobs/>).
- Rockhill, Gabriel (2017): *A jelen ellentörténete. Korszerűtlen kihallgatások a Globalizáció, technológia, demokrácia*. Durham: Duke University Press.
- Rosenblatt, Frank (1958): "The Perceptron. A Probabilistic Model for Information Storage and Organization in the Brain". In: *Psychological Review* 65.6: 386-408.
- Samuel, Arthur (1959): "Some Studies in Machine Learning Using the Game of Dáma." In: *IBM Journal of Research and Development*, Vol. 3: 210-229.
- Shane, Scott és Wakabayahsi, Daisuke: "'The Business of War': Google Employees Protest Work for the Pentagon". In: *The New York Times* 20184., április. (<https://www.nytimes.com/2018/04/04/technology/google-letter-ceo-penta-gon-project.html>).
- Srnicek, Nick (2016): *Platform Capitalism*. Cambridge, UK; Malden, MA/USA: Polity Press.
- Sudmann, Andreas (2016): "Wenn die Maschinen mit der Sprache spielen." In: *Frankfurter Allgemeine Zeitung* No. Nov256., 2: N2.
- Sudmann, Andreas (2018a): "Zur Einführung. Medien, Infrastrukturen und Technologien des maschinellen Lernens". In: *Machine Learning. Medien, Infrastruk-*

- turen und Technologien der Künstlichen Intelligenz*. Ed. Christoph Engemann és Andreas Sudmann. Bielefeld: transcript: 9-23.
- Sudmann, Andreas (2018b): "Szenarien des Postdigitalen. Deep Learning als MedienRevolution" In: Engemann/Sudmann, 55-73.
- Taddeo, Mariarosaria és Floridi, Luciano (2018) "Hogyan lehet a mesterséges intelligencia a jó erő",
Tudomány (3616404): 751-752.
- Valladão, Alfredo Da Gama e Abreu (2018): "Mesterséges intelligencia és politikatudomány". *Policy Paper*, szeptember, 1-29.
- Webb, Amy (2019): *A nagy kilenc. Hogyan a technológiai titánok és gondolkodó gépek hogyan tudnák Warp Humanity*. New York: Hachette Book Group.
- Weizenbaum, Joseph (1976): Weizenbaum: *Computer Power and Human Reason. Az ítélőképességtől a Számítás*. Oxford, Anglia: Freeman & Co.
- Wiener, Norbert (1988 [1950]): *Wiener: The Human Use of Human Beings: Cybernetics and Soci-ety*. London. Free Association.
- Zuboff, Shoshana (2018): *A felügyeleti kapitalizmus kora. A harc az emberi Jövő a hatalom új határán*. New York et al: Public Affairs.

Metaforák, amelyek szerint élünk¹

Három kommentár a mesterséges intelligenciáról és az emberi állapotról

Anne Dippel

Előjáték

A következő esszében három önálló kommentárt szeretnék összegyűjteni, amelyek mindegyike a mesterséges intelligencia egy-egy más aspektusával foglalkozik, és mindegyike egy másik metafora körül forog: intelligencia, evolúció és játék. Az első kommentár egy auto-etnográfiai vignetta, amely keretet ad a mesterséges "intelligenciáról" és a gépek állítólagos "gondolkodási" képességéről való elmélkedésnek; mindkettő nagyon problematikus metafora a feminista szemszögből, amely a (túlnyomórészt) női munkára, az intelligens emberi lények születésére és felnevelésére vonatkozik. A második betekintést nyújt jelenlegi etnográfiai terepmunkámba, amelyet olyan

nagyenergiájú fizikusok körében végzek, akik napi munkájuk során gépi tanulási módszereket használnak, és a darwinista metaforának engednek, amikor az evolúciós algoritmusok jelentőségét képzelik el az emberiség jövője szempontjából. A harmadik kommentár a "játészó" algoritmusokat vizsgálja, és a beszélgetésbe bevonja az "idegen" sokat vitatott antropológiai kategóriáját, amely - ahogy érvelésem szerint - sokkal relevánsabb a mesterséges intelligencia megértéséhez, mint a közvetlen megszemélyesítés, egy nem emberi lény életre keltése.

Új, nem mesterséges intelligens élet születik

Egy újonnan született emberi lényt látok. Napról napra társaságom van velem, amint egyre bonyolultabb testi mozgásokat gyakorol, érzékeli a körülötte lévő többi test belső érzelmeit, vagy reagál a tengernyi megkülönböztethetetlen hangra, annak ellenére, hogy egyetlen szó jelentését sem érti. Míg ő a saját

saját reflexeimet, egy életet megváltoztató eseménynek vagyok tanúja: egy minden más, csak nem mesterséges intelligencia megjelenésének. Lassan a motoros tevékenységek egyre kontrolláltabbá válnak, az izomzat fokozatosan felépül, a tekintet pedig mintha kissé tudatosan, egy adag kíváncsisággal és áhítattal követné az érdeklődési pontokat. Egy fiatal ember tanul.

Egy új élet kialakulását látva radikálisan újragondolom a mesterséges intelligencia és a gépi tanulás fogalmát, és még inkább a nyelv jelentőségét, amelynek hatalmában áll a politikai valóságot alakítani.

Tudnak-e gondolkodni a gépek, kérdezte Alan M. Turing majdnem hetven évvel ezelőtt (1950). Provokatív metaforája a mai napig meghatározza azt a módot, ahogyan az informatikusok az algoritmusok adatfeldolgozási képességét és "intelligens" (vagy inkább érthető) eredményeit érzékelik. Az intelligens gép képe a közvéleményben erősen megerősödött. Ma már "intelligens" infrastruktúrákról, okos tévékről, okos otthonokról, sőt okos városokról beszélünk; mind az úgynevezett "okossági mandátumot" példázzák (Halpern, Mitchel, Gheoghegan 2017).

Tudnak-e tanulni a gépek? Ez már nem kérdés, hanem feltételezés és módszer, amelyet szinte minden olyan tudományágban használnak, amely a nagy adatokra támaszkodik, a fizikától kezdve a piacon és a pénzügyeken át a mezőgazdaságig. A gondolkodás és a tanulás, amelyek eredendően emberi tulajdonságok, a gépekre vonatkoztatva kevés értelmet nyernek. Gyakran ártatlan metaforákként utasítják el őket. A szavaknak azonban hatalmuk van. Nemcsak a környező valóságot írják le, hanem a gondolkodásunkat és a cselekvésünket is alakítják (Lakoff és Johnson 1980). Ebben az értelemben a gépi "intelligencia" sokkal több, mint retorikai eszköz. Befolyásolja a felfogásunkat, mint (nem)emberi tulajdonságot.

Az intelligencia fogalma egy nagyon specifikus és szűk értelmezésből ered, amely szerint mit jelent intelligens lényként viselkedni. A mesterséges

intelligencia úttörőjének tartott Christoph von der Malsburg, aki eredetileg részecskefizikusnak készült, az intelligenciával kapcsolatos neurobiológiai kutatásaiban elsősorban a vizuális megismerésre és a memóriára összpontosított (Malsburg 1990). Nem nehéz párhuzamot vonni a gépi tanulási algoritmusok mai felfogásával, amelyeket gyakran dicsérnek azért, mert az emberi képességet meghaladóan képesek mintákat felismerni a gigantikus adathalmazokból. Egy olyan antropológus számára, aki a különbség antropocentrikus kritériumait alapvetően gyanúsnak tartja, ez a leegyszerűsített ember-gép metaforikus összehasonlítás naivitásában kissé kiábrándítóan tűnik, ha nem is hátborzongatóan. Von der Malsburg diadalmasan állította, hogy az emberi agy memóriakapacitása nem haladja meg az egy gigabájtot. Az embereket azonban nem nyers adathalmazokkal etetik. A gépek pedig az emberekkel ellentétben nem feltétlenül rendelkeznek a tapasztalatok palimpszesztikus biológiai memóriájával, hanem inkább egy kiterjesztett memóriával, hogy von der Malsburg metaforájával játsszunk tovább, amely egy tágas adattároló tartály metaforája.

Mindenekelőtt az emberi intelligencia és memória nem egydimenziós...

al egymáshoz való viszonya. Az intelligencia egy megtestesült folyamat, amely nagymértékben függ a kapott figyelemtől és törődéstől. Elég egy gyors pillantást vetni egy újonnan született emberre, hogy elutasítsuk a számítástechnika vak lelkesedését a mesterséges élet létrehozására. Ebben az összefüggésben a gépi tanulás üres, testetlen metaforának tűnik. A test (a csecsemő és az anyja) az, ami központi szerepet játszik az intelligencia fejlődésében. Egy újszülött esetében a fizikai és a pszichológiai elválaszthatatlanok. A test és az elme még nincs szétválasztva, a karteziánus dualizmusnak alávétve. Nem léteznek különálló entitásokként, vagy inkább kölcsönös ölelésben léteznek. Minden megtestesült, és minden tudatos. Az arckifejezés, a gesztusok és a hang a szociokulturális cenzorok nélkül működik. Arcuk lassan megtanul megvetni, eleinte véletlenszerűen, később koncentráltabban. Úgy tűnik, mintha a csecsemő tudata fokozatosan egy teljesen kifejezett "én"-né húzódná össze. Eleinte apró szálaknak tűnik, majd kitágul, növekszik és megnyílik, hogy tudatos lényé váljon. De mielőtt ez megtörténne, a baba egyszerűen csak létezik. A csecsemők intelligenciája a környezettel való bánásmódban fejlődik. Követelik, hogy észrevegyék és érzékeljék őket, bár még nem képesek megérteni, hogy mi is az a figyelem valójában.

Mindazok a mindennapi megfigyelések, amelyeket feminista anyaként és antropológusként gyűjtöttem, arra a meggyőződésre juttattak, hogy az emberi és a mesterséges intelligencia bármilyen összehasonlítását bizarrnak, ha nem teljesen értelmetlennek kell tartanunk. Egy olyan csecsemő szociális és érzelmi komplexitásának megfigyelései, akinek a feje a testsúlya egyharmadát teszi ki, és akinek nincs nyelve, de egyszerre több is tud lenni, mint nyelv, világossá tették számomra, hogy a differenciálatlan intelligencia mint olyan fogalma a legveszélyesebb szempont a mesterséges intelligenciáról szóló politikai vitában. A mesterséges intelligenciával kapcsolatos kutatások középpontjában a fogalom rendkívül leegyszerűsített és testetlen értelmezése áll, amely nemcsak túlbecsüli a gépi intelligenciát és alábecsüli az ember biológiai komplexitását, hanem azzal a veszéllyel is jár, hogy teljesen elveti a felelős emberi ágens jelentőségét.

Miközben a neuro-számítógépes tudósok az önmásoló algo-ritmikus

intelligenciáról álmodoznak, megszámlálhatatlan női test táplálja és ápolja a még meg nem született emberi intelligenciát. Miközben a tudomány továbbra is az embert mint testet öltött metaforát használja fel, hogy a mesterséges életet dicsőítse, a női test képes a teremtés igazi csodájára, a (túlnyomórészt férfi) technotudományos világ csodálkozó tekintete alig érinti. A megtestesült gondoskodás politikáját (Hamington 2001) vagy a gondoskodás politikáját a technotudományban (Martin, Myers, Viseu 2015) kell visszahozni a mesterséges intelligenciáról és annak ahhoz való viszonyáról, hogy mit jelent embernek lenni, folytatott szélesebb körű társadalmi beszélgetésbe.

A mesterséges intelligencia prométheuszi álma a fizikában

Szokásos antropológiai terepmunkám során nem csecsemőket tanulmányozok, hanem olyan tudósokkal ülök szemben, akik mesterséges intelligenciával foglalkoznak; pontosabban nagyon specifikus gépi tanulási algoritmusokkal, amelyek képesek a részecskék bomlásának végtelen adatai között szitálni. Az Európai Nukleáris Kutatási Központban (CERN) jó néhány nagyenergiájú részecskefizikus dolgozik, akik "a számítógép segítségével játszó isteneknek" tekintik magukat. A CERN-ben a kutatók mindennapi munkájuk során egyre inkább a felügyelt gépi tanulásra támaszkodnak. A CERN-ben már az 1980-as években bevetették az úgynevezett MVA-t (multivariable analysis), a gépi tanulás egy formáját (Galison 1997).

Eleinte a nagyenergiájú részecskefizikusok a számítógépes tudományos szakértelemtől függetlenül fejlesztettek ki algoritmusokat a ritka szubatomi ütközések mintáinak felismerésére. A fizikusok és az informatikusok közösségei nem voltak mindig olyan szoros kapcsolatban egymással, mint ma. A "részecskegyorsító Nagy Hadronütköztető" (LHC) létrehozásával azonban ez a két látszólag egymástól távol álló közösség összeolvadt. A nagyenergiájú fizika tudásbázisának fokozatos "informatizálódása" következett be, amely az adatsűrűség tárolására és az események előzetes meghatározásához és az elméletek fizikai méréseken alapuló teszteléséhez szükséges Monte Carlo-elemzések elvégzésére alkalmas nagy teljesítményű számítógépektől függ.

Az elmúlt 15 évben egyre több informatikus lépett be a mindennapi kutatási gyakorlatba, amint azt a CERN éves statisztikái is mutatják, akik támogatják a fizikusokat a kódolásban és a kísérletek szimulálásában (CERN Annual Statistics Website 2019). A CERN informatikusokba és a számítógépes re-keresés különböző területeibe fektet be, a gépi tanulási algoritmusoktól a kvantumszámításig. A "betanított" algoritmusok tengernyi adatot gyűjtenek, detektálnak és elemeznek. A kortárs nagyenergiájú fizikusokat "kódvarázslóknak" (Chun 2013) nevezhetjük, akik a világot pszeudo-véletlenszerű algoritmusok szemüvegén keresztül értelmezik. Így nem meglepő, hogy az emberiség jövőjéről alkotott elképzeléseiket oly mélyen kondicionálják a körülöttük "élő" algoritmikus infrastruktúra logikái. A legtöbb fizikus azonban elutasítaná ezt a feltételezést. Hajlamosak az algoritmusokat mediált eszközöknek tekinteni, amelyek képesek lehetnek arra, hogy kiterjesszék elménket, ugyanakkor teljes mértékben a fizikusok irányítják és szelídítik őket. A két kategória, az ember és a gép, világosan elkülönül egymástól, mindegyiknek más-más szerepe és hierarchiája van a kísérletben. A fizikusok meg vannak győződve az ember felsőbbrendű helyzetéről az algoritmusokkal szemben, legyenek azok bármilyen intelligensek is. Ha a fizika szemiotikai-materiális elemzésével kapcsolatban kétségek merülnek fel, azok általában a szakterületen kívül hangzanak el, például a médiatudományokban vagy a filozófiában, vagyis olyan tudományágakban, amelyek a természettudományok kortárs ismereteinek "közvetítettségét" tükrözik.

54 Anne Dippel

A fizika pártatlan bírónak tekinti magát, akit nem érintenek a médiumok logikái. Más szóval, azt, hogy a megfigyelő hogyan és mit lát, nem befolyásolja az a berendezés, amelyet a megfigyelt látására találtak ki.

Ugyanakkor a szinte steril ember-eszköz szétválasztásról szóló meggyőződéseket a jövő kiborgjáról szóló spekulációk kísérik, a mesterséges intelligenciával feljavított és attól szinte elválaszthatatlan holnap emberéről. Az ilyen kiborgikus víziókat sok fizikus osztja, különösen azok, akik a CERN spekulatívabb és jövőorientáltabb kutatásokkal foglalkozó részlegeiben dolgoznak, például a biológiai evolúció (szaporodás, mutáció, rekombináció, szelekció) elvei által inspirált úgynevezett evolúciós algoritmusokkal. Itt található az olyan informatikai látványok, mint Rodrigo Suarez, egyik informátorom. Ő a gépekben az intelligencia kontinuumát látja, amely az egyes sejtektől a teljes értékű emberig fejlődik, és a számítógépben éri el végső állapotát. Még ha nincs is teljesen meggyőződve arról, hogy a mesterséges intelligencia elérheti az emberhez hasonló állapotot, arról álmodik, hogy egy napon az emberek fejlődhetnek és örökké élhetnek, félelemtől és betegségektől mentesen, mesterséges intelligenciával feljavított kiborgokként. Rodrigo Suarez nem lát különbséget egy biológiai sejt, egy számítógép vagy egy emberi lény intelligenciafogalma között. Beszélgetésünk során érvélelésnek peremére sodrom, de Rodrigó Suarez (és sok más informatikus) számára ezek csak az örök élet előnyei léteznek, még akkor is, ha a halhatatlanság álmát a legerősebb keveseknek kell elérni. Az evolúció elve nem számol be a mindenki számára való méltányosságról vagy igazságosságról. Úgy tűnik, hogy egy durva darwinista vélemény van beágyazva az algoritmikus koncepciókba, amelyek a mesterséges intelligencia jelenlegi kutatási politikáját irányítják. Míg az informatika az embert visszahozza a középpontba, addig a természeti kultúra kutatása decentralizálja őt. A felvilágosodás nagy "M"-mel írt alakja (Tsing 2015) visszaköveteli hatalmi pozícióját. Az evolúciós algoritmusok, amelyek még korai fejlődési szakaszban vannak, az "episztemológia és az ontológia" (Bruder 1532018,), valamint az elme és a test technológiával való egyesítésének álmán nyugszanak, hozzájárulva a *homo automata sapiens* felemelkedéséhez.

Egyesek számára ez talán csak egy nárcisztikus álmom a termelésről és a reprodukcióról.

(méhirigység?), talán még az ókori görög értelemben vett önhittség is, Prométheusz vagy Éva játékmódja, amely megpróbálja ellopni a f lamét vagy az almát (Dippel 2011). Úgy tűnik, nehéz megtalálni az egyensúlyt a techno-optimizmus és a techno-pesszimizmus között, különösen egy olyan tudós számára, aki a technológia egyik új sámánjaként dolgozik. Ettől függetlenül a mesterséges intelligenciával kapcsolatos politikának figyelembe kell vennie az embereket.

A mesterséges intelligencia mint idegen a játékban

A prométheuszi álmom akkor látszik a legjobban illusztrálódni, amikor a gépek és az emberek egy játékasztalnál "néznek szembe" egymással, közvetlen ludikus

konfrontációban. A kibernetika legújabb kori történetében több sorsdöntő játszma is lezajlott, például Mac Hach VI az amerikai sakkszövetség játékosa ellen (1967) vagy az IBM ikonikus szuperszámítógépe, a Deep Blue Garri Kaszparov ellen (1996, 1997). Mindkét esetben az ember győzött

a számítás pusztá erejével. Egy 2015, egészen más versenyző lépett a globális szintre. Az Alpha Go, egy számítógépes program, amely képes a Go játékot (amely stratégiailag sokkal összetettebb, mint a sakk) játszani, győzött egy emberi játékos ellen. Az első győzelem után a professzionális Go játékost, Lee Sedolt is legyőzte. Az AlphaGo egy Monte Carlo fakereső algoritmust használ (ugyanazt a módszert, amelyet a CERN-ben a nagyenergiájú fizikában használnak), hogy új optimális lépéseket találjon.

Ezek a példák azt mutatják, hogy az ember-gép összehasonlítás iránti vágy milyen mélyen beágyazódott a technológiai fejlődés történetébe. Az ember a mérce, amely a technológia számára az intelligencia szempontjából a fő kritériumként szolgál. A Lee Sedol és az AlphaGo közötti játék is felvetette az "idegenség" kérdését - vajon a mesterséges intelligencia másképp játszik, mint az emberek? Használhatjuk-e egyáltalán a "játék" kategóriáját egy algoritmusra vonatkoztatva? Játszanak-e a számítógépek? A fenti kérdések mind összetettebbek, mint amilyenek látszanak, különösen, ha figyelembe vesszük, hogy az AlphaGo olyan lépéseket választott, amelyek extrém kockázatra való hajlandóságukban szinte embertelennek tűntek. Ahogy a Deep Mind csapata hangsúlyozza: "Az AlphaGo stratégiája a flexibilitás és a nyitottság szellemét testesíti meg: az előítéletek hiánya lehetővé teszi számára, hogy megtalálja a leghatékonyabb játékmenetet" (DeepMind.com). A mesterséges intelligencia hajlamos jól kezelni egy potenciálisan káros áldozatot hozó elképzelést, ha az páratlan kompenzációhoz vezet a játékban. Általánosabb filozófiai szinten azt mondhatnánk, hogy nincs tudata, és nem is érti a saját lehetséges "halálát". Ez egy egészen másfajta játékeret nyit meg, ahol minden döntés olyan kockázatos lehet, amennyire a csekkék és a balansz logikája megengedi.

A mesterséges intelligencia továbbra is nem egzisztenciális kapcsolatban áll semmivel.

ami az emberek számára fontos (vö. Dippel 2018). Végül is a gépeket pontosan azért hozták létre, hogy enyhítsék vagy megkönnyítsék az emberiség egzisztenciális állapotát (vö. Giedion 1982). Antropológiai szemszögből azt állíthatnánk, hogy az ember - a "nagy M betűs fickó, aki az antropocént létrehozta" (Tsing 2015) - megteremtette önmaga "metaforikus megfelelőjét" (Lévi-Strauss 2381973,); a másság diszpozícióját olyan időkben, amikor az ember és más élőlények közötti hagyományos határszabályok kérdésessé váltak. Két fő utat látok tehát a mesterséges intelligencia vízióiban. Egyrészt megfigyelhetjük a mesterséges intelligencia "metaforikus megfelelőjeként" való létrehozását, hogy Claude Lévi-Strauss antropológus és az ember és a madarak közötti összehasonlítást továbbgondoljam. Mindkét faj sok más hasonlóság mellett kapcsolatokat alakít ki és fészket épít, de van egy dolog, amit mi, emberek nem tudunk - repülni. Ebben az értelemben a madarakat metaforikus megfelelőnek tekintjük, amelyben a korlátozott képességeink repülésének és kiterjesztésének álma tárolódik. A mesterséges intelligencia olyan, mint egyfajta madár. Lehetővé teszi számunkra, hogy lássuk, mik vagyunk és mik nem vagyunk; amivé álmodunk válni, de talán soha nem lehetünk azok. Másrészt a mesterséges

intelligencia inkluzív változata, amely a "harmadik természet" (Richter & Rötzer 2018), a kiborgok fogalmain alapul.

(Haraway 1991) és a természet-kultúra (Gesing, Knecht, Flitner & Amelang 2019), amely a politikai szférától és a társadalmi következményektől függetlenül létezik.

A mesterséges intelligencia első koncepciói, ahogy Norbert Wiener híres megfogalmazása szerint, a modern rabszolgák megteremtéséről szóltak (1972, 72). Az úr és szolga közötti kapcsolatokkal kapcsolatos régi félelmek tükröződnek a mesterséges intelligencia politikájával kapcsolatos vitákban, már a kezdetek óta (Winner 1977). Ahelyett, hogy egy olyan rendet keresnének, amely lehetővé tenné egy jobb társadalom kialakulását, a jelenlegi koncepciók vakon reprodukálják a meglévő uralmi és posztkolonialista viszonyokat. A mesterséges intelligencia víziója ma többnyire neoliberais és pozitívista világnézet enged, a soha meg nem szűnő automatizált munka eszményét erőltetve (Gregg 2018). A munkásosztály végét jelentő, osztályalapú álmokat táplálva túlnyomórészt elitista fantáziákat szolgál. Nem veszi figyelembe a fenntartható környezet megteremtését, amely lehetővé teszi az embereknek, hogy megtalálják helyüket a természetben. Ehelyett a természetet "a másikként" ápolja, amelyet a technológia segítségével kell uralni.

A technológia azonban hajlamos beláthatatlan irányokba elkalandozni, ami termékeny talajt biztosít az ideológiák számára (Latour 2006). A közösségi média körüli aktuális kérdések nagyon találó példaként szolgálnak erre. A barátok és családok világméretű összekapcsolására készültek, de a politikai szférában bomlasztó és manipulatív eszközökké váltak, mélyen befolyásolva az emberi képességet arra, hogy komplex szövegeket megértsen, vagy hogy hosszabb ideig fenntartsa a figyelmet. Ez a talán triviális példa csak azt mutatja, hogy manapság kiemelkedően fontos, hogy a mesterséges intelligenciát ne csak kifejezetten technikai szempontból vizsgáljuk, hanem szélesebb társadalmi-kulturális és politikai kontextusban. Kutatóként és állampolgárként ébernek kell maradnunk.

"Fed" az emberek által és az emberekért

A mesterséges intelligenciát annak kell tekinteni, ami valójában: egy technológiai idegen lénynek. Ha figyelmen kívül hagyjuk a mesterséges intelligenciának ezt az "idegenségét" vagy másságát, akkor félreértjük azt a képességét, hogy más politikák utópikus lehetőségeihez vezessen. Valójában csak azáltal, hogy a mesterséges intelligenciát a technológiailag másként kezeljük, láthatjuk úgy, mint valami olyasmit, ami "kiter az én és a kultúra rendje elől, ugyanakkor kihívást jelent számukra" (Leistle 2015). A status quo megkérdőjelezéséhez pedig a hatalmi metaforák tudatos használatával vagy kritikájával kezdhethetjük, emberi képességeket tulajdonítva a mesterséges intelligenciának, vagy egy meghatározott társadalmi-politikai keretbe (jelen esetben neoliberais és pozitívista keretbe) ágyazva azt. A Fehér Ház jelentése a mesterséges intelligenciáról a néhai Obama-adminisztráció idején így hangzik: "A gépi intelligencia fejlesztése és tanulmányozása segíthet jobban megérteni és értékelni emberi intelligenciánkat.

Átgondoltan használva a mesterséges intelligencia növelheti intelligenciánkat, segíthet nekünk jobb és bölcsebb utat kijelölni" (Technology Council Committee 2016, 7, 39). Az ilyen grandiózus politikai felvetéseket azonban be kell ágyazni egy új társadalmi valóságba, ahol minden polgárnak

nyílt hozzáféréssel rendelkeznek a mesterséges intelligencia által vezérelt áruhoz. A kutatóknak, a politikusoknak, a magánszektorban és a közvéleménynek el kell jutniuk a mesterséges intelligencia közösségivé tételéhez és az emberek felhatalmazásához, ami a jelenlegi politikai és gazdasági rendszerben nehezen elképzelhető. Ebben az értelemben a mesterséges intelligenciának az emberek tulajdonában kellene lennie, mivel túlnyomórészt az emberek "táplálják", például a digitális technológia használatának mindennapi gyakorlatában, és így lehetővé téve a technológiai vállalatok számára, hogy adatainkat összegyűjtsék, hogy a vállalati titoktartási megállapodások mögé bújtatva táplálják algoritmusait. Az emberiség és a mesterséges intelligencia jövője nem engedhet a darwinista vízióknak. Ebben az utópisztikus kontextusban a mesterséges intelligencia valódi médium és közvetítő lehetne - nem pedig egy sötét, privatizált Leviatán, amelyet azok számára manipulálnak, akik szeretnek háborút vezetni, hatalmat tartani és erőforrásokat felhalmozni. Ahhoz, hogy egy ilyen vízió valóra váljon, egy nagyobb társadalmi párbeszédre van szükség, amely túlmutat a gyors számítástechnika és az automatizált munka optimalizálási logikáján. Olyan emberekre van szükség, akik a *vita activa-t* gyakorolják és felelősséget vállalnak, ahelyett, hogy arról álmodoznának, hogy azt egy techno-istennek adják ki.

Ezzel a megjegyzéssel szeretném lezárni ezt az esszét, mert egy sokkal igényesebb teremtmény várja, hogy tápláljuk, nem nyers adatokkal, hanem tejjel, figyelemmel és gondoskodással. Az ő intelligenciája még sok-sok évet fog igénybe venni, hogy kifejlődjön, a szuperszámítógép számítási teljesítményétől és a Monte Carlo keresési algoritmusoktól függő mértékben. A fiam táplálása sokkal többet igényel, mint hogy "elég tartalma legyen" (Stokel-Walker 2019). Ez egy olyan szeretetmunka, amelyet az emberiség kezdete óta nemzedékről nemzedékre adnak át nők és férfiak. Olyan, amelynek nincs szüksége "metaforikus megfelelőjére" a technológiában.

Bibliográfia

- Barr, Alan J., Andy Haas és Charles W. Kalderon (2016): "'That looks weird'- evaluating citizen scientists' ability to detect unusual features in ATLAS images of LHC collisions". ATL-COM-OREACH-2016-017, arXiv:1610.02214v1.
- Bruder, Johannes (2018): Bruder: "Ahol a nap soha nem süt. A megvilágosodás utáni megismerés kialakuló paradigmái". In: Digitális kultúrák és társadalom 4/1, pp. 133-157.
- A CERN éves statisztikai honlapja (2019): (<https://cds.cern.ch/collection/CERN%20Annual%20Personnel%20Statistics?ln=de>) [Utolsó hozzáférés: 2019.4.24].
- Chun, Wendy Hui Kyong (2013): Programozott látomások: Szoftver és memória. Cambridge, Mass.: MIT Press.
- DeepMind.com: (<https://deepmind.com/research/alphago>). [Utolsó elérés: 2019.4.24].
- Dippel, Anne (2017): "Das Big Data Game. Zur spielerischen

62 Anne Dippel

Konstitution kollabo-

rativer Wissensproduktion in der Hochenergiephysik am CERN". In: *NTM 4* (2017), pp. 485-517.

- Dippel, Anne (2018): Dippel: In: Feige, Daniel M., Ostritsch, Sebastian, Rautzenberg, Markus (Eds.): *Philosophie des Computerspiels. Theorie - Praxis - Ästhetik*. Stuttgart: Metzler, pp. 124-148.
- Dippel, Anne (2011): Ironisches Prolegomenon für einen "Entartungsschutz des Menschen" zum vernünftigen Wesen vom Homo Sapiens Sapiens zum Homo Sapiens Optimus. In: *Ist Technik die Zukunft der menschlichen Natur?* Göttingen: Wehrhahn, pp. 104-114.
- Executive Office of the President National Science and Technology Council Committee on Technology (2016): Felkészülés a mesterséges intelligencia jövőjére (https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf). [Utolsó hozzáférés: 2019. 4. 24.].
- Galison, Peter (1997): Galison: *Image and Logic. A mikrofizika anyagi kultúrája*. Chicago: Chicago: Chicago University Press.
- Gesing, Friederike, Knecht, Michi, Flitner, Michael, Amelang, Katrin (szerk.) (2019): *NaturenKulturen. Denkräume und Werkzeuge für neue politische Ökologien*. Bielefeld: transcript.
- Giedion, Sigfried (1982): Giedion: *Die Herrschaft der Mechanisierung. Ein Beitrag zur anonymen Geschichte*. Frankfurt a. M.: Athenäum.
- Gregg, Benjamin (2018): Gregg: Az eljövendő politika. In: *Digitális kultúrák és társadalom* 4/1, pp. 157-180.
- Halpern, Orit, Mitchel, Robert, Geoghegan, Bernard Dionysius (2017): The Smart- ness Mandate: Notes Toward a Critique. In: *Grey Room*, pp68., 106-129.
- Hamington, Maurice (2001): Jane Addams és a megtestesült gondoskodás politikája. In *The Journal of Speculative Philosophy* 15/2, pp. 105-121.
- Haraway, Donna (1991): A Cyborg Manifesto. Tudomány, technológia és szocialista feminizmus a huszadik század végén. In: *Simians, Cyborgs and Women: The Reinvention of Nature*. New York: Routledge, pp. 149-181.
- Lakoff, George, Johnson, Mark (1980): *Metaphors We Live By*. Chicago: Chicago University Press.
- Latour, Bruno (2006): Latour Latour: Ethnografie einer Hochtechnologie: Das Pariser Projekt "Aramis" eines automatischen U-Bahn-Systems. In: Rammert, Werner (szerk.): *Technografie. Zur Mikrosoziologie der Technik*. Frankfurt a.M. , New York: Campus, pp. 25-60.
- Leistle, Bernhard (2015): A masszázs mint paradigma az antropológiában. In: *Semiotica: Journal of the International Association for Semiotic Studies* 204, pp. 291-313.
- Lévi-Strauss, Claude (1973): Das wilde Denken. Frankfurt a.M: Suhrkamp.
- Martin, Aryn., Myers, Natasha, Viseu, Ana (2015): *The Politics of Care in Technoscience*. In: *Social Studies of Science* 45/5, pp. 1-17.
- Richter, Steffen, Rötzer Andreas (2018): *Dritte Natur Technik1. Kapital Umwelt*. Berlin: Matthes & Seitz.

- Stokel-Walker, Chris (2019): Az algoritmusok etetése teljes munkaidős feladat. BBC (<http://www.bbc.com/capital/story/20190307-the-hidden-armies-that-power-the-internets-new-stars>) [Utolsó elérés: 2019.4.24].
- Tsing, Anna (2015): Az antropocén feminista megközelítése: Earth Stalked by Man. A Barnard College for Women-ben tartott előadás. (https://www.youtube.com/watch?v=ps8J6a7g_BA) [Utolsó elérés: 2019.4.24.].
- Von der Malsburg, et al. (1990): *Pattern segmentation in associative memory*. Neural Computatio 2, pp. 94-106.
- Wiener, Norbert (1950): *Wiener: The Human Use of Human Beings*. Boston: Houghton Mifflin.
- Winner, Langdon (1977): *Autónóm technológia. Az ellenőrzésen kívüli technika mint a politikai gondolkodás témája*. Cambridge, Mass: MIT Press.

AI, sztereotipizálás szteroidokon és Alan Turing biológiai fordulata

V. N. Alexander

Bevezetés

A mesterséges intelligencia (AI) tervezői az emberi agy képességeit próbálják utánazni "öntanuló" neurális hálózatokkal, amelyeket tömeges kiválasztási folyamatok vagy más "felügyelet nélküli" kiválasztási folyamatok segítségével képeznek ki. Feltehetően a bemeneti adatok logikája hasonlóan íródik bele a mesterséges hálózat szerkezetébe, mint ahogyan a bemeneti adatok az emberi agyat formálják. Mégis, évtizedek múltán a mesterséges intelligenciával képzett csevegőrobotok és fordítóalkalmazások még mindig nem tudják átugrani a Turing-teszt alacsony lécét. Egyre világosabbá válik, hogy a mesterséges intelligencia nem képes jeleket értelmezni fluid kontextusokban. A biológiai számítás minőségileg különbözik a mai gépi számítástól? Alan Turing halála idején azt vizsgálta, hogy a biológiai reakció-diffúziós folyamatok hogyan hoznak létre mintákat, amelyek viszont korlátozzák a sejtválaszokat és differenciáltan indítják el a fejlődést. Ma úgy gondoljuk, hogy hasonló mechanizmusok biztosítják az időbeli és térbeli korlátokat az idegsejtek csoportjai számára, amelyek lehetővé teszik számukra az érzékszervi kötődést, valamint az emlékek kialakítását és felidézését. Ha megélte volna, hogy folytassa munkáját, Turing talán átirányította volna a mesterséges intelligencia kutatását, hogy jobban kezelje a kontextuális korlátok létrehozásának kihívását, ami talán szükséges ahhoz, hogy az emberi ítélőképességnek nevezett, előre nem látható és szinte csodaszerű válaszok létrejöjjenek. A jelenlegi, statisztikailag szervezett, emberi ügyekben alkalmazott mesterséges intelligencia csak sztereotipizálásra jó, ami természetesen aláássa az egyéni demokratikus szabadság alapfeltételét.

Egy "okos" gép, akárcsak egy organizmus, képes megkeresni egy tárgyat, leolvasni egy kódot, megtalálni egy tárgyat.

egy mintát és általánosításokat tesznek. Egy szervezethez hasonlóan egy gépet is lehet úgy tervezni, hogy önfenntartó célokat kövessen. Azt azonban nem mondhatjuk, hogy a gépek jelenleg humanoid intelligenciával rendelkeznek. A mesterséges intelligencia robotok nem tudják megérteni az embereket, mert nem értenek a nyelvhez. Nem értik az iróniát, az új metaforákat, metonimákat, szójátékokat vagy vicceket. A nyelv alapvetően utalásos, nem szó szerinti, ahogy Turing egyszer egy barátjának írt levelében bebizonyította:

Turing úgy véli, hogy a gépek azt hiszik,
 hogy Turing hazudik az embereknek.
 Ezért a gépek nem gondolkodnak (1952b).

Az a tény, hogy Siri nem érti ezt a viccet, nem azért van, mert nincs elég világ és idő a hálózat betanítására; ez a mesterséges intelligencia és a biológiai intelligencia (BI) közötti alapvető különbségre utal. Az olyan szelekciós folyamatok, mint amilyeneket a mesterséges intelligencia-hálózatok képzésére használnak, nem képesek valódi intelligenciát kifejleszteni. Azért tehetek ilyen merész kijelentést, mert az olyan szelekciós folyamatok, mint amilyenek a neodarwinisták állítása szerint az állati intelligenciát fejlesztették ki, valójában nem képesek erre a feladatra. A mesterséges intelligencia chatbotok kudarca a közmondásos döglött kanári, amely egy sokkal nagyobb problémára utal: az intelligens cselekvés evolúciójának és fejlődésének túlzottan lebecsült felfogására. A mesterséges intelligencia jobb lehet a BI-nél az olyan mechanikus, szabályokhoz kötött cselekvésekben, mint az autózvezetés, de nem képes meghatározni, hogy az ember mit akar vagy mit szándékozik tenni. A nyilvánosságot nem szabad arra kérni, hogy bizzon a mesterséges intelligenciában, elfogadva, hogy annak működése csak egy rejtély. Ez az írás célja, hogy lerántja a leplet a mesterséges intelligencia varázslójáról, felfedve, hogy ez az állítólag emberfeletti intelligencia valójában csak egy eszköz, egy nagyon erős eszköz, amelyet néhányan arra használnak, hogy sokakat irányítsanak.

Az innováció huszonegyedik századi evolúciós elmélete

Az elméleti biológusok számára egyre világosabbá válik, hogy bár a genetikai anyagban bekövetkező kis véletlenszerű változások természetes szelekciója szerepet játszik az evolúciós folyamatokban, az ilyen szelekció eredménye egy faj stabilizálódása és a diverzitás csökkenése. Az innováció, ma már úgy véljük (lásd Turing, 1952a; Margulis & Sagan, 2002; Reid, 2007; Shapiro, 2011; Noble, 2016), valószínűleg olyan nagy, egymással összefüggő mutációs eseményeknek köszönhető, mint a hibridizáció, génduplikáció, laterális génátvitel, transzpozonok, szimbiózis és, ami e vita szempontjából fontos, a Turing által felfedezett termodinamikus ön- vagy önorganizáló szemiotikai folyamatok. Az ilyen mechanizmusok általában új, kész eszközöket (nem véletlenszerűen összeállított dolgokat) hoznak létre, amelyek funkciói aztán szelektálhatók vagy nem szelektálhatók. Ez nem az apád evolúciós elmélete.

A mesterséges intelligenciát azon a feltételezésen alapul, hogy az adaptív tanulás az 1950-es évekbeli, fokozatos szelekcióval járó neodarwinista modell követi. A mesterséges intelligencia a természetes szelekcióhoz hasonlóan az alkalmasság statisztikai definícióján alapuló általánosításokat végez: a

leggyakrabban előforduló mintákat választja ki. A mesterséges intelligencia pozitív és negatív megerősítéssel tanul. A BI is tanulhat így, de lehetnek megvilágosodások is.

Algoritmusok kontra szemiotikai szokások

Hogy megpróbáljam megmagyarázni, miért nincs humorérzéke a mesterséges intelligenciának, azzal kezdem, hogy míg a számítógépek *digitális kódokat* használnak és *algoritmusokat* fejlesztenek ki a kontextustól függetlenül, addig az élő sejtek *analóg jeleket* használnak és önerősítő *szemiotikai szokásokat* alakítanak ki a kontextuson belül. Ez a tanulmány a Biosemiotika, egy újonnan fejlődő, transzdiszciplináris terület, amely a kibernetika, a komplex rendszertudomány és a biokémia területéhez kapcsolódik, szémszögéből vizsgálja a mesterséges intelligencia és a biológiai intelligencia közötti különbségeket. A bioszemiotika szerint, míg egy kód az egyik forma pontos lefordítását igényli egy másikba, addig egy jel többféle formára fordítható le, más jelek és átalakítók viszonylagos hasonlóságától és/vagy közelségétől függően. Lehet, hogy a biológiai jeleknek ez a flexibilitása lehetővé teszi, hogy a jelátvitel könnyen folyjon, szinkronosan és koherensen kommunikálódjon a szomszédos sejtek felé, még akkor is, ha a jel nem egészen a helyes vagy hagyományos. Lehet, hogy ez a különbség a mesterséges intelligencia és a biológiai intelligencia között magyarázatot adhat arra, hogy a mesterséges intelligencia nem képes megfelelően lefordítani a jeleket a kontextusokban.

Ha a mesterséges intelligencia öntanuló algoritmusai úgy tűnik, hogy néha jól működnek az emberi cselekedetek előrejelzésében, akkor ez azért van, mert a sztereotípiák gyakran igazak. A mesterséges intelligenciát jelenleg az amerikai bírósági rendszerekben használják az ítéletek meghatározásának segítésére, felnagyítva a strukturális társadalmi előítéleteket az AI-hálózatba táplált adatokban. A bűnöző újbóli elkövetésének valószínűségét úgy jósolják meg, hogy típusba sorolják. Az eredmény az, hogy a feketék keményebb büntetést kapnak, mint a fehérek hasonló adatpontok esetén (Angwin et al., 2016). Az AI a sztereotipizálás szteroidokon.

A mesterséges intelligenciát a lakosság irányításában is alkalmazzák. Andrew Hallman, az amerikai Központi Hírszerző Ügynökség digitális innovációért felelős igazgatóhelyettese szerint az internetfelhasználókról gyűjtött összes adatnak köszönhetően az ügynökség a mélytanulás segítségével ma már jobban "előre tudja jelezni a társadalmi zavargások és a társadalmi instabilitás kialakulását... három-öt napra előre" (Konkel, 2016). Ez aggaszt, hogy egy bűnözés előtti egység állt fel és működik. Nem lehet igaz, hogy a magánéletünk feláldozásával biztonságban leszünk. A tömeges megfigyelés és a nagy adatgyűjtés csak az ellenvélemény elhallgattatását és a status quo fenntartását szolgálhatja, nem pedig a tényleges bűncselekmények megelőzését. Az összetett rendszerek, mint az emberek, hajlamosak nem lineárisan viselkedni: az egyéni viselkedés előrejelzésének képessége nem javul arányosan az előrejelzésekhöz felhasznált adatmennyiség növekedésével.

Bár a neurális hálók tervezői visszacsatolást és előreccsatolást használnak a nemlineáris biológiai folyamatok utánzására, nem tartalmaznak kreatív

mechanizmust, és az így létrejövő intelligencia olyan névtelen, arctalan bürokráciákra hasonlít, amelyek sok generáción keresztül halmozták fel a polgárokkal való bánásmódra vonatkozó eljárásokat, és amelyek nemcsak konzervatívak, hanem minden egyes iterációval egyre inkább szűkítik a lehetőségeket.

Az első lépés a mesterséges intelligencia demokratizálása felé az, hogy leleplezzük ezt az állítólagosan az embernél jobb bírót. A mesterséges intelligencia nem ügynök, hanem egy erős és potenciálisan hasznos eszköz.

amelynek sokak kezében kellene lennie, nem csak néhány ember kezében. Michael Kwet (2018) meggyőzően érvelt amellett, hogy a demokratikus digitális társadalomhoz kompromisszumok nélküli adatvédelemre, nyílt forráskódú szoftverekre, valamint decentralizált személyes felhőrendszerekre van szükség, amelyek lehetővé teszik az információk közvetlen megosztását. Ha a személyes adatok gyűjtését így leállítjuk, a bíróságok és a felügyeleti szervek nem lesznek képesek arra, hogy a mesterséges intelligenciát arra használják, hogy az egyéneket a különböző csoportokhoz való tartozásuk vagy a velük való társulásuk alapján ellenőrizzék. Remélem, hogy a mai mesterséges intelligencia és a BI közötti különbségekről szóló elemzésem meggyőzheti a nyilvánosságot, hogy szkeptikusabban viszonyuljon a mesterséges intelligencia előrejelzéseinek feltételezett bölcsességéhez vagy pontosságához.

Mit tudott Turing a BI-ről

Turing feltalálta a legpraktikusabb eszközt, amelyet az emberek valószínűleg valaha is használni fognak. Mérnöki sikereit mégis az emberi intelligencia természetének megértésére irányuló, nem gyakorlatias vágy vezérelte. Mi is az ő példáját követjük, amikor megpróbáljuk megérteni, hogyan használjuk a legjobban a számítástechnikai eszközöket a huszonegyedik században.

Az 1950-es években, miután javaslatot tett az "öntanuló" számítógép modelljére, Turing gondolkodása egy olyan fordulatot vett, amelyet utólag úgy nevezhetnénk, hogy másodrendű kibernetika vagy mesterséges élet. Kísérleteket és tanulmányokat kezdett végezni a matematikai biológia területén. Míg Andrew Hodges (1983), Turing fő életrajzírója a növényi és állati morfogenezis iránti érdeklődését az intelligencia utánzása iránti érdeklődésétől való eltérésnek tekintette, Jack Copeland (2004), aki a Turingtudományának meghatározó kommentárját adja, rámutat, hogy Turing világossá tette, hogy ez az új munka az intelligens számítás további vizsgálata, még akkor is, ha a zsiráfminták, Fibonacci-spirálok és a levélgenerálás felé fordult a figyelme.

Turing fedezte fel azokat a spontán folyamatokat, amelyek révén a nem szervezett rendszerek beavatkozás, külső szelekció nélkül szerveződnek. C. H. Waddington (1940) azt javasolta Turingnak, hogy a fejlődés egyszerűen valahogy rendeződik, a legkisebb ellenállás útját követve folyik. Turing volt az, aki azt javasolta, hogy egy instabilitás, egy véletlen minta - nem pedig egy kifejezetten erre a funkcióra tervezett induktor - indíthatja el a kevesebb rendből a több rendbe, a kaoszából a differenciálódásba (vagy, Gregory Batesonra utalva, egy különbségbe, amely különbséget tesz) való átmenetet. Ilya Prigogine, aki Nobel-díjat kapott a kapcsolódó kutatásért, Manchesterben találkozott Turinggal, és 1952 megvitatták az elméletet (Hodges, 1983: 587). Prigogine csak a *Physic Today*-ben megjelent tanulmányában 1972 ismerte el Turing hozzájárulását.

Turing azt állította, hogy az instabilitási ponttól távolabb diffundáló reakciók

az úgynevezett morfogenetikai mezőket eredményezik, amelyek differenciáltan határozzák meg a génaktivitást, ahogyan azt Waddington leírta. Míg a biológusokat az érdekelte, hogy ez mit jelent az embriológia számára, addig Turing arra vonatkozó ismereteket keresett, hogy a neuronok hogyan differenciálódhatnak és szerveződhetnek hasonló módon. Egy 1951-es, J. Z. Young neurofiziológusnak írt levelében Turing megjegyzi: "Az agy szerkezetének olyannak kell lennie, amelyik képes

a genetikai-embriológiai mechanizmus révén érhető el" (Copeland, 2004: 517). Bár nem tudjuk, hogy Turing a fejlődést analógnak gondolta-e a tanulással, kiderül, hogy igen. Édesanyjának, Sara Turingnak talán nem volt teljesen igaza, amikor úgy vélte, hogy fia halálakor egy "korszakalkotó felfedezés" küszöbén állt (qtd in Hodges, 1984: 624).

AI a BI-vel összehasonlítva

Amikor Turing először megtervezte öntanuló hálózati számítógépét, azt feltételezte - hallgatólagosan a neodarwinizmust követve -, hogy az emberek *véletlenszerű* találgatásokat tesznek, amikor nem ismerik a probléma megoldására szolgáló eljárást. Az 1948 "Intelligens gépek" című művében azt állítja, hogy "egy emberi gyermek kiképzése nagyrészt a jó és rossz találgatásokért járó jutalmak és büntetések rendszerétől függ" (Copeland, 2004: 425). Turing egy sakkozó programot tervezett, amely véletlenszerűen kipróbálható választható lépésekkel rendelkezik. Ha egy lépés végül kudarchoz vezetett, azt nem erősítették meg. Turing neurális hálózatát úgy tervezte, hogy szervezetlenül induljon, és megfelelő "interferenciával" szervezetté váljon, imitálva a nevelést. Hasonló típusú *kapcsolati* megközelítéseket használnak ma a legtöbb öntanuló algoritmusban. Az "utasítástábla", ahogyan Turing a programot nevezte, a hálózatban testesül meg, ahogyan azt a jutalmazás és a büntetés megváltoztatja. A visszajelzés lehet programozó által kezelt, vagy az interneten tömegesen gyűjtött. Bár ezt a megközelítést öntanulónak vagy önszerveződőnek nevezik, ahogy Turing megjegyezte, az ilyen megközelítésekhez még mindig szükség van külső "beavatkozásra".

A mesterséges intelligencia legújabb fázisát "felügyelet nélküli" tanulásként nevezik. Egy vizuális felismerő hálózatot például több millió véletlenszerű, címkézetlen képnek tettek ki az interneten. Végül felismert néhány közös mintát, kitaláltad, macskaarcokat, ismeretlen módon, több száz szint mélységben elsajátítva pályákat és torzításokat (Le et al., 2013). A programozók nem mondták meg a hálózatnak, hogy mit találjon, de a hálózat most már használható macskaarcok megtalálására. Ezek az új, felügyelet nélküli hálózatok nem sokban különböznek Turing 1948-as, öntanuló hálózatról alkotott elképzelésétől. A fő különbség az a pont, ahol a programozó beavatkozik, a képzési folyamat során, hogy egy előre meghatározott mintát célozzon meg, vagy miután a hálózat felismerte a programozót érdeklő mintát. Az utóbbi esetben a kiválasztás egysége a teljes hálózat, nem pedig a hálózaton belüli egyes kapcsolatok.

Az állatok leggyakrabban "felügyelet nélküli" helyzetekben tanulnak, különösen a nem emberi állatok, és ritkábban tanítják, szándékosan jutalmazzák és büntetik őket. Ez a majom látja, majom csinálja megközelítés. De a beavatkozás, vagyis a szelekció még mindig működik. Egy eljárás újra és újra történő megtapasztalása valójában megváltoztatja a neuronális kapcsolatokat. Az

együtt tüzelő neuronok összekapcsolódnak, ahogy Donald Hebb (1949) oly híressé tette...

szerk. A bemagolással történő tanulás, a kapcsolatok idővel történő megerősítése statisztikai jellegű. Ami a *legtöbbször* előfordul - akár "helyes", akár nem -, az kiválasztódik és megerősödik.

Az ismétlés az egyik módja a neuronok kapcsolatainak kialakulásának, de nem az egyetlen.

Az emberek (és valószínűleg más állatok is) jobban fel tudnak idézni részleteket összefüggésekben, ha azok tetszőlegesen hasonló vagy tetszőlegesen közeli dolgokhoz kapcsolódnak. A rímek, ritmusok, hangzások és más költői eszközök, például a metafora és a metonímia általában segítik az emlékezést, még akkor is, ha a kapcsolatokat nem erősítik meg ismételten. A neuroncsoportok félig-meddig otikus szokásait ritka sztochasztikus rezonanciák (azaz pusztán véletlenszerű minták) indíthatják el, amelyek önszerveződéshez vezetnek. Az emberi logika és nyelvhasználat kiszámíthatatlanságának forrása, a neuronsejtekben és a neuronsejtek között zajló jelcselekvésnek ez a költői típusa uralja a tudatalatti folyamatokat. A hipnózisban lévő alanyok keresztmodális észlelést tapasztalnak - például elkezdnek színeket hallani -, ami arra utal, hogy amikor a tudatos észlelést megkerüljük, a tudatalatti költői működése jobban megfigyelhető (Alexander és Grimes 2017). A szinesztéziás emberek jobban fel tudnak idézni tetszőleges tényeket, mivel a számokhoz vagy betűkhöz egyedi színek, textúrák és formák társulhatnak (Harvey 2013). Az önkényesen hasonló/közeli tényezőkön alapuló összefüggések nem redukálhatók statisztikai leírásra; a tényezők *száma* nem annyira lényeges az eredmények szempontjából, mint a tényezők más tényezőkkel szembeni tulajdonságai.

A bioszemiotika formalizálása

Modellezheti-e egy számítógép azt, ahogyan a természet szerveződik azáltal, hogy összekapcsolja az egymáshoz hasonló/közeli dolgokat? Turing felfedezte a nemlineáris egyenleteket, amelyek képesek számítógép által generált zebracsíkokat, invaginációt, metakronális hullámokat és más természetes emergens mintákat létrehozni. Bár Turingmunkája évekig nem volt bizonyított, és sokan úgy vélték, hogy az egyenletei által generált és a természetben található minták közötti hasonlóságok csupán véletlenek, Sheth (2012) és Raspopovic (2014) végül kimutatták, hogy a Turing-mechanizmus valóban leírja azt a folyamatot, amelynek során a fejlődő embriókban ujjak jönnek létre. A biológusoknak időbe telt, amíg azonosították azokat a tényleges kémiai jeleket, amelyek megfelelnek azoknak a kapcsolatoknak, amelyeket Turing elképzelése szerint meg kellett volna kapniuk, ha az önszerveződés a differenciálódás és a fejlődés mechanizmusa lenne. Turingegyenletei összetettek, de elég, ha csak annyit mondunk, hogy a diffúziós sebességek, a reakciósebességek és az ezen sebességek változásának módjait jelző változókat tartalmazzák. A reakciók jellemzően több morfogént foglalnak magukban, például X és Y reakciója Z-t eredményez; Z és A reakciója 2Y-t eredményez. Az első reakció kimeríti Y-t; a

második növeli Y-t. Másképp fogalmazva, a folyamat magában foglalhat egy olyan aktivátort, amely katalizálhatja saját termelését és saját inhibitorának termelését, amely bizonyos esetekben gyorsan diffundálhat, megteremtve a terepet az utazó hullámminták kialakulásához. Van ellentmondás vagy

paradoxon ezekben a folyamatokban, amelyek egyszerre önteremtőek és önkorlátozóak, egy kicsit úgy, mint Turing szillogizmusa, amelyet a tanulmány elején mutattunk be.

Hadd próbáljam meg megvilágítani az ilyen típusú folyamatok bioszemiotikus elemeit egy nagyon leegyszerűsített, mindössze két elemet tartalmazó vizuális modellel. A biológiai számítás szemléltetésére anyagi tulajdonságokkal rendelkező alakzatokat használok szimbólumként, mert a biológiai jelek és receptorok (jelolvasók) kötődése gyakran alakfüggő. Legyen mondjuk, hogy van egy \mathbb{L} típusú molekulánk és egy \mathbb{T} típusú molekulánk. Ezek különböző irányok, pl. \uparrow és \downarrow . Sem a \mathbb{L} -k, sem a \mathbb{T} -k nem lépnek kölcsönhatásba önmagukkal. Tehát

$$1. \mathbb{L} + \mathbb{L} = \mathbb{L}\mathbb{L} \text{ és}$$

$$2. \mathbb{T} + \mathbb{T} = \mathbb{T}\mathbb{T}$$

A $\mathbb{T}\mathbb{S}$ és $\mathbb{L}\mathbb{S}$ együttesen bizonyos orientációkban szintén nem eredményez változást: például,

$$3. \mathbb{L} + \mathbb{T} = \mathbb{L}\mathbb{T} \text{ és}$$

$$4. \uparrow + \downarrow = \uparrow\downarrow$$

Amikor azonban a \mathbb{T} -k és a \mathbb{L} -k más módon találkoznak, kölcsönhatásba léphetnek egymással és megváltozhatnak, például egy \mathbb{T} átalakulhat egy \mathbb{L} -vá. Az átalakulások attól függenek, hogy a \mathbb{T} nyitott horizontális része találkozik-e a \mathbb{L} nyitott vagy zárt horizontális részével. Például,

$$5. \mathbb{L} + \downarrow = \mathbb{L}\downarrow$$

$$6. \uparrow + \mathbb{T} = \mathbb{T}\uparrow$$

Ezek azok az egyszerű helyi szabályok, amelyek korlátozzák a kölcsönhatásokat. Az [5] és [6] összefüggéseiben azt mondhatjuk, hogy a \downarrow metaforikusan olyan, mint egy \mathbb{L} , és a \uparrow metaforikusan olyan, mint egy \mathbb{T} .

Mivel a molekulák mindig hőmozgásban vannak, a találkozásuk módja véletlenszerű. Statisztikailag nézve az új $\mathbb{L}\mathbb{S}$ vagy az új $\mathbb{T}\mathbb{S}$ keletkezése egyformán valószínű. Azt gondolhatnánk, hogy ezek a reakcióforgatókönyvek együttesen hajlamosak arra, hogy

hogy a véletlenszerű keveredés fenntartásával átlagolódjon, de - ahogy Turing egy hasonló kísérletben megállapította - ehelyett differenciálódás következhet be. A mi kísérletünkben mondjuk $\mathbb{L}\mathbb{S}$ csomó alakul ki egy területen, ahogy az megtörténhet, mivel a véletlenszerűség nem tökéletesen nem ismétlődő. Egy csomóban nem keletkezik új \mathbb{T} , mert egy \mathbb{T}

több $\mathbb{T}\mathbb{S}$ előállításához szükséges. Még több $\mathbb{L}\mathbb{S}$ keletkezhet a csomó szélein, ha a $\mathbb{L}\mathbb{S}$ ok véletlenül érintkeznek a megfelelő orientációjú \downarrow sokkal. A csomó önmagát növeli. Nincs szükség külső beavatkozásra. Azt mondhatjuk, hogy e \downarrow

és \mathbb{L} jelek anyagi tulajdonságai (azaz a jelek viszonylagos hasonlósága és közelsége) vezetnek a kollektív aktivitáshoz, egy emergens foltmintázathoz. A \mathbb{L} -k képesek értelmezni (reagálni, kölcsönhatásba lépni, lefordítani) a \downarrow jeleket, és

több ilyen jelet termelni.

önmagunk, több \mathbb{L}_s .

Egy ilyen keverékből készült levesből néhány \mathbb{L} csomót és néhány $\overline{\mathbb{L}}$ csomót kapnánk, amelyek mind a \mathbb{L} -k, mind a $\overline{\mathbb{L}}$ -k véletlenszerű keverékében lebegnek. Ha \mathbb{L} = fekete és $\overline{\mathbb{L}}$ = fehér, akkor szürke alapon fekete és fehér foltok jelennek meg, mint az ausztrál szarvasmarha-kutya bundáján. A

Az állati szőrminták kialakulásának tényleges folyamata sokkal bonyolultabb, de ez egyszerű vizuális illusztrációként szolgál a spontán önszerveződésre, amely a természetben mindenütt előfordul, különösen az agyban.

A "The Chemical Basis of Morphogenesis" (Turing, 1952a) és "A Diffusion Reaction Theory of Morphogenesis in Plants" (Turing & Wardlaw, 1952) című munkáiban Turing bemutatja, hogy a nemlineáris egyenletek képesek leírni azt, ahogyan a minták szponzívan alakulnak ki a szerves anyagból. Megmutatja, hogy a géneknek nem kell teljes mértékben meghatározniuk a szervezet komplex szerkezetét. A kódoló gének főként a sablonokat adják meg az anyagok előállításához, a megfelelő sorrendben és mennyiségben, de nem tartalmazzák az utasításokat arra vonatkozóan, hogyan kell az anyagokat összerakni.¹ Nem is kell. A fizika és a kémia törvényei, valamint az anyagok tulajdonságai (mint például a π s és \mathbb{L} s) az átalakítási szabályok és a korlátozások, amelyek segítik a gén által termelt anyagok önszerveződését. Számítógép-programozóként Turing nagy csodálattal tekintett volna a természet leleményességére és gazdaságosságára. Nem kellett fizikailag rögzítenie a DNS-ben a fejlődési eljárást. Ehelyett a Természet spontán önszerveződő programokat használt ki.

A bioszemiótikus alkalmazkodás ebben a rendszerben akkor lehetséges, ha például egy \mathbb{L} megtörténik.

egy új molekulával, az L-molekulával, *mintha az egy \mathbb{L} lenne*. (L tehát a \mathbb{L} téves jeleként működik.) A biológiai rendszerek által felfedezett minden új jelnek egy készen álló, már létező jelölvasó rendszernek kell működnie. *Nem lehetnek pusztán véletlenszerűek*, mint a neo-Dar-

winian elmélet. Az L és \mathbb{L} kötődés eredménye egy új molekula lehet, amely differenciáltan váltja ki az új kombináció által érintett sejteket.

Waddington (1940) adta Turingnak az *epigenetikai tájképet*, amely a fejlődést irányító fizikai erők (vagy általában a sejtválaszok) vizuális metaforája volt, és amely Turing elméletét inspirálta. Waddington azzal érvelt, hogy mielőtt egy sejt differenciálódna, instabil állapotban van, mint egy labda, amely egy hegy tetején ül, alatta pedig különböző völgyyszerűségekkel. Turing felismerte, hogy bármilyen kis fluktuáció az instabilitás ezen pontjáról az egyik vagy másik völgypálya felé lökheti. Ezek az elképzelések René Thom korai bioszemiótikus "katasztrófaelméleteként" váltak ismertté (lásd Favareau, 2009: 337-376). Waddington azt találgatta, hogy az alternatív útvonalak "versengő" autokatalitikus reakciók lehetnek, amelyek ugyanazoknak a molekuláknak egy részét különböző folyamatokhoz használják. Az L kötődhet a \mathbb{L} -hez, és elindíthat egy útvonalat, vagy a \mathbb{L} kötődhet a \mathbb{L} -hez, és elindíthat egy másik útvonalat.

Az önszerveződés szelekciós folyamata az elemek *formai* tulajdonságain, tulajdonságain alapul, nem csak az elemek *számán*, mint a statisztikai szelekció esetében. Turing fedezte fel azt a folyamatot, amelynek során a differenciáló hullámok, a morfo- genetikai mezők spontán, külső szelekció nélkül keletkeznek. *Ez a fajta kompu- táció valóban öntanuló.*

I A génműködés megértésének történetét lásd Keller (2002).

Emergens agyi minták

A huszadik század nagy részében úgy vélték, hogy az agyhullámok öncélúak, mint a motor hangja, amely nem járul hozzá a motor működéséhez. Most már azt kell figyelembe vennünk, hogy ezek a hullámok egyfajta emergens programként szervezhetik a neuronok működését. A szakirodalom alapos áttekintésében Kelso et al. (1991), Uhlhaas et al. (2009) és De Assis (2015) arról számolnak be, hogy sok idegtudós a munkamemória és a figyelem mögött álló mechanizmusokat emergens agyhullámok formájában érti, amelyek távoli neuronokat szinkronizálnak, virtuális neuronegyütteseket hozva létre (De Assis, 2015; Postle, 2006). Úgy tűnik, hogy a hullámok biztosíthatják "a "kontextusokat" az elvi sejtek hálózatai által hordozott "tartalom" számára" és "a pontos időbeli struktúrát, amely szükséges ahhoz, hogy a neuronegyüttesek meghatározott funkciókat hajtsanak végre, beleértve az érzékszervi kötődést és a memóriaképzést" (Buzsáki & Chrobak, 1995). Ezen túlmenően az emergens hullámmintázatok határozhatják meg azt is, hogy mely adatok kapnak figyelmet, azaz tudatosságot (lásd Thompson & Varela, 2001), ami viszont befolyásolja a további érzékszervi feldolgozást.

A tanulásnak ez a jelterjedési elmélete, amely önszerveződő jeleket (nem kódokat) használ, segíthet megmagyarázni, hogyan képesek az emberek rugalmasan alkalmazkodó kategóriákat kialakítani és használni, és kezelni a komplex, változó környezetet. A helyi fluktuációk lehetővé teszik

a sztochasztikus rezonancia (mint az L_s és L_s esetében), a hasonló és közeli lehetséges állapotok, ami viszont lehetővé teszi az egyformaság terjedését, azonnali szerveződést. A természetes szelekció nem "lát", hogy ezeket a helyi kölcsönhatásokat kiválassza (nincs is rá szüksége, mivel ezek a kölcsönhatások spontán módon egyszerűen a legalacsonyabb energiaállapotba áramlanak). Ami a fitness szempontjából szelektálható, az a lokális kölcsönhatásokból kialakuló globális mintázatok hatása (vö. Rocha, 1998).²

A mesterséges neurális hálózatokat vagy a mélytanuló hálózatokat nem úgy tervezték, hogy utánozzák az önszerveződés és a természetes szelekció közötti fluid kölcsönhatást. A mesterséges intelligencia designátorok inkább a szigorúan szelekcionista, azaz konnekcionista megközelítések mellett kötelezték el magukat. Bár a tanulás így is megvalósítható, ez automatákat hoz létre, akárcsak a szabványosított tantervek és a könnyörtelen tesztelés, jutalmazás és büntetés.

A Google Translate még a legutóbbi ünnepelt frissítéssel (Levis-Kraus, 2016) is rosszul bánik a szójátékokkal, viccekkel és versekkel. Jung-Beeman és munkatársai (2004) pszichológusok szerint a belátás - az irodalmi témák és metaforák megértése és a viccek megértése - megköveteli a távoli agyterületek azonnali szinkronizálását a gamma-hullámokon keresztül. Ahhoz, hogy olyan számítógépeket tervezzünk, amelyek képesek az utalásos nyelvet, az embereket megérteni, talán egy fluidabb közegre van szükség az utazó hullámok megjelenéséhez. A Stieg et al. (2014) szerinti atomi kapcsolóhálózatok

ígéretesnek tűnnek; ezeket olyan emergens minták létrehozására használták, amelyek egyszerű természetes rendszereket utánoznak. Kísérleti kémiai

2 Hasonlóképpen, az önző gén hipotézissel ellentétben, a természetes szelekció nem "látja" a géneket *önmagukban*.

csak a termékeiket.

A reakció-diffúziós számítógépek már több mint egy évtizede léteznek (Adamatsky et al., 2005), de bár emergens mintázatokat hoznak létre, az állandóbb kapcsolatokról lemondanak. Úgy tűnik, agyunk mindkettőt használja.

Lehet, hogy végül a reakció-diffúziót fogjuk használni, hogy több humanoid mesterséges intelligenciát hozzunk létre, de már most is nyolcmilliárd emberi számítógép van összekapcsolódva az Inter-hálózatban, mint megannyi neuron, amelyek készen állnak a szerveződésre. A tudás látványos fejlődésének lehetősége a kezünkben van, ha csak mi irányítanánk a mesterséges intelligencia algoritmusait, nem pedig mi irányítanánk őket. Ha több információval rendelkezünk a mesterséges intelligencia természetéről, mint a BI-ről, jobban tudnánk dönteni arról, hogy mennyire vagy mennyire vagyunk hajlandóak hagyni, hogy a mesterséges intelligencia helyettünk gondolkodjon.

Hivatkozások

- Adamatzky, A., Costello, B. és Asai, T. (2005): *Reakció-diffúziós számítógépek*. Amsterdam: Elsevier.
- Alexander, V. N. & Grimes, V. A. (2017): Folyékony bioszemiotikus mechanizmusok a tudatalatti szokások hátterében. *Biosemiotics*, 10(3), 337-353.
- Angwin, J., Larson, J., Mattu, S., Kirchner, L. (2016): Machine Bias. *ProPublica*. Retrieved from <http://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> (hozzáférés: 2019., július).
- Copeland, J. (2004): *Turing: Az esszenciális Turing*. Oxford: Oxford University Press.
- De Assis, L. (2015): Neural binding, consciousness, and mental disorder: Komplexitás mint közös elem. *Journal for Neurocognitive Research*, 57 (3-4), 110-121.
- Favareau, D. (2009): *Essential readings in Biosemiotics: Anthology and commentary*. Amsterdam: Springer.
- Gilbert, C. & Sigman, M. (2007): Brain states: Top-down hatások az érzékszervi feldolgozásban. *Neuron*, 54 (5), 677-696.
- Harvey, J. (2013): Harvey Harvey, Harvey: Érzékszervi észlelés: A szinesztézia tanulságai: Using synesthesia to inform the understanding of sensory perception. *Yale Journal of Biology and Medicine*, 86(2), 203-216.
- Hebb, D.O. (1949): Hebb: *A viselkedés szerveződése*. New York: Wiley & Sons.
- Jung-Beeman, M., Bowden, E.M., Haberman, J., Frymiare, J. L., Aradmbel-Liu, S., Greenblatt, R., Reber, P. J., & Kounios, J. (2004): Idegi aktivitás, amikor az emberek belátással oldanak meg verbális problémákat. *PLOS Biology* 2 (4), 500-510.
- Keller, E. F. (2002). *A gén évszázada*. Cambridge: Harvard UP.
- Kelso, J., Bressler, S.L., Buchanan, S., DeGuzman, G.C., Ding, M., Fuchs, A., & Holroydet, T. (1991). Fázisátmenet az emberi agyban és viselkedésben.

Physics Letters A 169 (3), 134-144.

Konkel, F. (2016): A CIA azt állítja, hogy már 3napokra5 előre meg tudja jósolni a társadalmi zavargásokat.

Egyes védelem. Retrieved from <http://www.defenseone.com/technology/2016/10/>

- cia-says-it-can-predict-social-unrest-early-3-5-days-out/132121 (hozzáférés: 2019., július)
- Kwet, M. (2018): *Mail & Guardian*: Break the hold of digital colonialism. 29. június. Retrieved from <https://mg.co.za/article/2018-06-29-00-break-the-hold-of-digital-colonialism> (accessed July 9, 2019)
- Le, Q. V., Ranzato, M., Monga, R., & Ng, A. Y. (2013): Magas szintű jellemzők építése nagy léptékű felügyelet nélküli tanulással. *Proceedings of the 29th International Conference on Machine Learning*. Edinburgh.
- Levi-Kraus, G. (2016): A nagy M.I. ébredés. *New York Times Magazine*. Margulis L. & Sagan, D. (2002): *A genomok megszerzése: A fajok eredetének elmélete*. New York: Basic Books.
- Noble, D. (2016): *Tánc az élet dallamára: Biológiai relativitás*. Cambridge: Cambridge University Press.
- Postle, B. (2006): Memory as an emergent property of the mind and brain. *Neuroscience* 139(1), 23-38.
- Prigogine, I., Nicolis, G. & Babloyantz, A. (1972): Termodinamika az evolúcióban. Az élő rendszerekben fenntartott funkcionális rend látszólag szembemegy a második törvénnyel; a nem egyensúlyi termodinamika leírja, hogy az ilyen rendszerek hogyan tudnak megbirkózni az entrópiával. *Physics Today* 25 (11), 23.
- Raspopovic, J., Marcon, L., Russo, L., & Sharpe, J. (2014): Digit patterning is controlled by a Bmp-Sox9-Wnt Turing network modulated by morphogen gradients. *Science* 345(6196), 566-570.
- Reid, R. G. B. (2007): *Biological emergences: Evolution by natural experiment*. Cambridge: MIT Press.
- Rocha, L. M. (1998): Válogatott önszerveződés és az evolúciós rendszerek szemiotikája. *Evolúciós rendszerek: Biológiai és ismeretelméleti perspektívák a selektivitás és önszerveződés témakörében*. S. Salthe, G. Van de Vijver és M. Delpo (szerk.). Kluwer Academic Publishers; 341-358.
- Shapiro, J. (2011): *Evolution: A 21. század nézőpontja*. FT Press.
- Sheth, R., Marcon, L., Bastida, M. F., Junco, M., Quintana, L., Dahn, R., Kmita, M., Sharpe, J. & Ros, M. (2012): A Hox gének egy Turing-típusú mechanizmus hullámhosszának szabályozásával szabályozzák az ujjpercek mintázódását. *Science* 338(6113), 1476-1480.
- Stieg, A., Avizienis, A., Sillin, H., Aguilera, R., Shieh, H.-H., Marin-Olmos, C., Sandouk, E., Aono, M. Gimzewski, J. (2014): Önszerveződés és dinamikus struktúrák kialakulása neuromorfikus atomkapcsoló hálózatokban. *Memristor Networks*. A. Adamatsky és L. Chua (szerk.) New York és London: Springer; 173-209.
- Thompson, D. W. (1942): Thompson: *On Growth and Form*. Cambridge: Cambridge University Press.
- Thompson, E., Varela, F. (2001): Radical embodiment: Neurális dinamika és tudatosság. *Trends in Cognitive Sciences* 5(10), 418-425.

- Turing, A. (1952a): A morfogenezis kémiai alapja. *Philosophical Transactions a londoni Királyi Társaság. B sorozat* 237(641), 37-72.
- Turing, A. (1952b): Turing: Levél Norman Routledge-nek. Letöltve: <http://www.turingarchive.org/viewer/?id=167&title=1a> (hozzáférés: 2019. július 9.)
- Turing, A., Wardlaw, C. (1952): A növények morfogenezisének diffúziós reakcióelmélete. Retrieved from <http://www.turingarchive.org/browse.php/C/7> (accessed July 9, 2019)
- Uhlhaas, P., Pipa, G., Lima, B., Melloni, L., Neuenschwander, S., Nikolic, D & Singer, W. (2009): Neurális szinkronitás az agykérgi hálózatokban: történelem, koncepció és jelenlegi helyzet. *Frontiers in Integrative Neuroscience* 3(17), 1-19.
- Waddington, C. (1940): Waddington: *Szervezők és gének*. Cambridge: Cambridge University Press.

Termékeny hangok

Az érintőhangos tárcsázás, a Call Center ipar felemelkedése és a virtuális hangalapú asszisztensek politikája

Axel Volmar

Az okostelefonokban, okos hangszórókban és járműfedélzeti rendszerekben egyre inkább elterjedő virtuális hangalapú asszisztensek, mint például az Amazon Alexa, az Apple Siri, a Google Assistant, a Microsoft Cortana vagy a Samsung Viv, a mesterséges intelligencia demokratizálódását jelentik a pusztán tömeges megjelenés révén.¹ A hangalapú asszisztensek, amelyeket általában intelligens virtuális asszisztenseknek (IVA) vagy intelligens személyi asszisztenseknek (IPA) neveznek, a szoftveres ügynökök azon osztályába tartoznak, amelyek szóbeli parancsok és megkeresések alapján képesek válaszolni a felhasználók kérdéseire és feladatokat végrehajtani, ha hangalapú felhasználói felülettel (VUI) vannak felszerelve. A technológiai vállalatok a hangalapú intelligens asszisztenseket a kortárs mesterséges intelligencia csúcsteljesítményeként, valamint az ember és a gép közötti zökkenőmentes együttműködés új formájaként reklámozzák, amelyek a digitálisan hálózatba kapcsolt és felhőalapú technológiák irányításának és navigálásának intuitívabb módját kínálják. A társalgási mesterséges intelligencia közelgő elterjedése azonban számos alapvető kérdést vet fel az algoritmikus vezérléssel, valamint a hangalapú ember-gép interakció természetével és történetével kapcsolatban. Hogyan strukturálódnak a hangalapú együttműködés e kialakulóban lévő formái, és hogyan változtatja meg a hangvezérlés a szoftvertechnológiához való viszonyunkat és annak kritikai értékelését? Milyen következményei vannak annak, hogy a mesterséges intelligencia-technológiák nagyrészt felhőalapú számítástechnikán alapulnak, és így a felhasználói adatokat felhőszerverekre küldik feldolgozásra?

Tekintettel a legtöbb kereskedelmi forgalomban kapható mesterséges intelligencia technológia "fekete doboz" jellegére,

természetesen meglehetősen nehéz részletes információkat szerezni arról, hogy az egyes hangalapú asszisztensek mesterséges intelligencia algoritmusai pontosan hogyan működnek. A politikájuk megértéséhez azonban nem szükséges minden részletükben megérteni, hogyan működnek algoritmikusan; elegendő annak tanulmányozása, hogy mire használják őket, és hogyan jelennek meg a különböző érdekeltek és szereplők előtt. Ezért az intelligens személyi asszisztenseket - mobiltelefonokon, operációs rendszereken és különösen az intelligens személyi asszisztenseknél - úgy fogalmazom meg, hogy

I A Canals piacelemző cég friss jelentése (2019) azt jósolja, hogy a világméretű intelligens hangszóró telepítési bázisa az egymillió 114eladott darabról több mint egymillióra200 nő 82.4százalékkal a következő év végére. 2019.

a médiatudós Tarleton Gillespie értelmében vett *platformok*. Nagy visszhangot kiváltó tanulmányában Gillespie amellel érvel, hogy a platformok politikája nyomom követhető annak vizsgálatával, hogy hogyan

az olyan online tartalomszolgáltatók, mint a YouTube, gondosan pozicionálják magukat a felhasználók, az ügyfelek, a hirdetőik és a politikai döntéshozók felé, stratégiai állításokat fogalmaznak meg arról, hogy mit tesznek és mit nem tesznek, és hogyan kell értelmezni a helyüket az információs tájképben. Különösen egy kifejezés, a "platform" mutatja meg ennek a diszkurzív munkának a körvonalait. (Gillespie 2010: 347)

Hasonlóképpen, ebben a tanulmányban kevésbé fogok a gépek belső működésére összpontosítani, mint inkább a hangalapú interfészek közvetlen környezetükhöz való különböző kapcsolataira, valamint a különböző szereplők, például a felhasználók, a call center ügynökök, a vállalkozások, a nagy technológiai vállalatok és a megfigyelő államok céljaira. Mindazonáltal jelentős történelmi kitérőt teszek annak érdekében, hogy a társalgási mesterséges intelligenciát a hang- és hangalapú ember-gép interakció tágabb történelmébe illesszem, és hogy hangsúlyozzam az ideiglenes hangalapú asszisztensek és a hálózati szolgáltatások korábbi hang- és hangalapú felhasználói interfészei közötti folytonosságokat és cezúrákat. E megközelítés mellett szól az is, hogy a hangalapú asszisztensek körüli jelenlegi felhajtás ellenére a hang- és beszédalapú ember-gép interfészek korántsem új keletűek. Stanley Kubrick *2001: Űrodüsszeia* (1968) című filmjének pszichológiai problémákkal küzdő HAL táblaszámítógépe óta az ember-számítógép interakcióra szolgáló beszédalapú interfészek állandó helyet foglalnak el az iparosodott társadalmak kulturális képzeletvilágában.

Bár az olyan kifinomult mesterséges intelligencia rendszerek, mint a HAL, még mindig science fiction maradnak, a hang és a beszéd valóban az egyik legrégebbi interfész a távoli rendszerekkel való interakcióhoz. A korai alkalmazások azonban nem a számítógépiparban, hanem a távközlési ágazatban jelentek meg. Röviddel az AT&T *a2001 DIVA* (digitális lekérdezés/hangos válasz) rendszerekben való lekérdezéshez használt Touch-Tone telefonjait népszerűsítette, amelyek lehetővé tették a számítógép által vezérelt és Touch-Tone parancsok által kiváltott hangüzenetek formájában történő információlekérdezést. Az elosztott szolgáltatásokkal hang és beszéd útján történő interakció telefonos gyakorlata még az 1940-es és 1950-es évekre nyúlik vissza, mielőtt a növekvő call center iparban továbbfejlődött volna. A gépekkel való beszéd mai gyakorlata tehát újraértelmezi a telefonhoz kapcsolódó, elfeledett vagy elvetett felhasználói tapasztalatokat. Ennek érdekében csatlakozom Jonathan Sterne (2012) *médiatudós*nak a telefon és a hangtechnológiák központi szerepének hangsúlyozásához a digitalitás történetében:

A telefonálást a megszokotthoz képest gyakran érdektelen dolognak tartják,

a huszadik századi médiatörténet esztétizáltabb témái, mint például a mozi,

televízió, hangfelvétel, rádió, nyomtatott sajtó és számítógép. A telefonálás és annak infrastruktúrájának sajátos jellemzői azonban az elmúlt mintegy 130 év legtöbb hangtechnológiájának hangzásában központi szerepet játszanak. A telefónia intézményi és technikai protokolljai segítettek kialakítani a kommunikáció ma is használt definícióit, az információ alapgondolatát, amely az "algoritmikus kultúra" egészét áthatja a csomagkapcsolt kapcsolástól a DVD-kig és a játékokig, valamint a digitális technológiák protokolljait és rutinjait, amelyeket nap mint nap használunk. (2-3)

Míg Sterne a telefonrendszer történetét, és különösen a jeltömörítési módszerek és a perceptuális kódolás fejlődését használta fel az mp3 formátum mint "kulturális artefaktum" kibontására (Sterne 2006), addig én a beszéddel kapcsolatos mesterséges intelligencia-alkalmazásokat a távoli telefonszolgáltatások és a távközlési és ügyfélszolgálati ipar (fél)automatizálási folyamatainak hosszabb története, különös tekintettel a call centerekre, háttérben tárgyalom. Az automatizálás a kezdetektől fogva a call centerek mozgatórugója, ha nem is a lehetőség feltétele. A legtöbb ilyen kísérlet alapja az, amit *produktív hangoknak* szeretnék nevezni, azaz olyan hangoknak, amelyek egy (fél)ig automatizált rendszeren belül meghatározott célokat szolgálnak, vagy akár szó szerint munkát végeznek, például kapcsolási vagy algoritmikus folyamatokat indítanak el.² Az olyan produktív hangok, mint a Touch-Tone jelek, a várakozási zene és a rögzített hangüzenetek állnak annak az átalakulási folyamatnak a középpontjában, amelyben a telefontársaságok a telefonrendszert a hangátvitelre szolgáló speciális célú alkalmazásból általános célú információs hálózattá kívánták bővíteni (vö. Lipar- tito 2003). A hangok szintetizált hangok formájában a társalgási mesterséges intelligencia és a digitális személyi asszisztensek esetében a hangok a korábban a számítógép-felhasználók által végzett általános célú kézi feladatok speciális célú helyettesítőjeként váltak produktívvá.

Médiaelméleti szempontból ezt az átmenetet úgy érthetjük meg, hogy a fogalom - a produktív hangmédiát nem a *kommunikáció* médiumaként, hanem - Erhard Schüttpelz német médiaelméletíró szavaival élve - az *együtműködés* potenciálisan erős médiumaként (Schüttpelz 2017: 14; vö. Volmar 2017). Például a telefon együttműködési médiumként emlegetni azt jelenti, hogy nem pusztán társalgási médiumként, hanem a logisztikai, bürokratikus, problémamegoldó és egyéb hétköznapi, személyes, munkával kapcsolatos "infrastrukturális" feladatok megkönnyítésére szolgáló univerzálisabb eszközként fogjuk fel (Star/Bowker 2002). Most, amikor az ilyen logisztikai feladatokat lazán az internethez, online platformokhoz, mobilalkalmazásokhoz vagy intelligens hangszórókhoz társítjuk, érdemesnek tűnik emlékeztetni arra, hogy a hálózatba kapcsolt információs szolgáltatások mögöttes narratívája valójában régebbi, mint maga az internet, és hogy egykor mélyen összefonódott a

2 Bár én a "produktív hangok" kifejezést ebben a konkrét értelemben használom, az általános

fogalmat Alix Hui-tól és Joeri Bruyninckx-tól veszem át, akik a "Produktív hangok a mindennapi terekben" című workshopjukon vezették be a kifejezést: *Sounds at Work in Science, Art, and Industry, 1920-Present* című előadásukban a Max Planck Tudománytörténeti Intézetben április 27-28-án, 2018.

áramköri kapcsolású távközlési infrastruktúra. Amellett érvelek, hogy a hangközponitú mesterséges intelligencia-alkalmazások a call centerekben (ma már általában "contact center"-nek nevezik őket) és a háztartási környezetekben a kooperatív hangmédia és a körülöttük forgó gyakorlatok automatizálására tett különböző kísérletek történetének jelenlegi eszközállományának tekinthetők.

Az együttműködés mindig különböző szereplők és csoportok által és között végzett gyakorlatokat foglal magában. A kooperatív gyakorlatok fejlődésének kiemelése érdekében a hangautomatizálás történetében különös figyelmet fordítok a telefonhoz és a hanghoz kapcsolódó munka- és munkagyakorlatok formáira. Míg a média- és technikatörténet kutatói kiterjedten tanulmányozták a telefonosok munkáját (pl. Green 1995; Lipartito 1994), én Sumanth Gopinath média- és hangkutatónak a csengőhang-iparral foglalkozó munkáját követem (Gopinath 2013), és a hang- és telefonos munka jelentőségére összpontosítok az ügyfélszolgálati ipar infrastrukturális keretein belül, hogy nyomon kövessék a hálózatos, beszédalapú ember-gép interakciók kialakulását. Ennek érdekében azt vizsgálom, hogy a call center ügynökök és az ügyfelek, valamint az emberek és a gépek közötti változó munkamegosztások és -delegálások hogyan alkotják a *távoli együttműködés* infrastruktúráit, amelyek egyes részeivel a jelenlegi digitális asszisztensekben is találkozhatunk.

Az első, részben visszalépek egy kicsit, hogy újra megvizsgáljam az AT&T által az 1960-as évek elején bevezetett nyomógombos telefon következményeit. A nyomógombos telefonok eredetileg az operátorok helyettesítésére szolgáltak a telefonhívások kezdeményezésének és kapcsolásának további automatizálásával, és a nyomógombos telefonok új tárcsázási módszere a kéthangú, többfrekvenciás (DTMF) jelzés volt, amely a "sávon belüli", azaz hallható vezérlőjelek alapján működött - azok a tárcsahangok, amelyeket a vezetékes és mobiltelefonokban még mindig hallunk, amikor a billentyűzeten lévő gombokat nyomogatjuk. Néha a hangokat még az okostelefonokon is szimulálják, például a messenger-alkalmazásokban. Azt állítom, hogy miközben a többfrekvenciás jelzések automatikussá és hatékonyabbá tették a telefonos kapcsolást, egyúttal a telefonos munkák delegálásának és kiszervezésének gyakorlatához is vezettek az operátoroktól az automatikus rendszerek és az ügyfelek felé.

Ennél is fontosabb, hogy az MF-jelzés lehetővé tette a hangkódolt alfanumerikus információk továbbítását a telefonhálózaton keresztül, és így a modern call centerekben az automatikus telefonos információs rendszerek kialakulásának alapvető feltételét képezte. A fejezetben 2. felidézek néhány ilyen technológiai újítást, különösen az automatikus híváselosztókat (ACD-k) és az interaktív hangválaszrendszereket (IVR), amelyek mindkettő alapvetően hozzájárultak a call center-iparág felemelkedéséhez. Ezután azt vizsgálom meg, hogy ezek hogyan járultak hozzá a telefonhívások félig-meddig-meddig-közepesedéséhez és a hang és a hangmunka további újraelosztásához azáltal, hogy a telefonbeszélgetéseket közös kérdésekre és szekvenciákra bontották, és hogy mind a call center ügynököknek, mind a hívóknak hogyan kellett

alkalmazkodniuk ezekhez a sztenderdizált "határobjektumokhoz" (Star/Griesemer 1989), hogy az automatizált rendszerek működőképeseek legyenek.

A 3. szakaszban azt mutatom be, hogy a mesterséges intelligencia hogyan lépett a színpadra a kontaktközpontban, ahogyan azt korábban nevezték, a beszédfelismerés, -megértés és -szintézis formájában. Azzal érvelek, hogy a telefonhívások félautomatizálásának és az ügynökök és az ügyfelek automatizált rendszerekhez való igazításának évtizedei különösen fogékonyá tették a contact center-t a mesterséges intelligencia technológiájára az iparágon belül. A társalgási mesterséges intelligencia megvalósítása az IVR-ekhez hasonló logikán alapul, mivel a telefonbeszélgetéseket főként korlátozott számú kategóriára vagy egységre bontja, például bizonyos kulcsszavakra vagy feltételezett érzelmi állapotokra. Ugyanez a logika van jelen a kortárs otthoni hangalapú asszisztensekben is. Azzal, hogy a kortárs hangalapú asszisztenseket a félautomatizálás és a kooperatív telefonos gyakorlatok tágabb történetébe helyezem, amely a produktív hangokon és a call center iparban végzett hangmunkán alapul, végső soron arra töreksem, hogy az internet és a digitális kultúra meglévő történetét (pl. Haigh et al. 2015) kibővítsen azzal, hogy a telefonalapú telekommunikáció fejlődését a digitálisan hálózatos média és a kooperatív gyakorlatok koncepciójának, tesztelésének és általános elterjedésének fontos területének tekintem.

1. Nyomógombos telefonok és érintőhangos tárcsázás: Innováció az általános célú infrastrukturális technológiákban

A huszadik század első felében a Bell-rendszerben a telefonhívások kezelése nagyrészt emberi telefonkezelők kezében maradt, annak ellenére, hogy az automatikus kapcsolás számos megoldása, például a Strowger-kapcsoló, már rendelkezésre állt. Míg a kézi kapcsoláshoz való ragaszkodás fő oka technikai kérdések és a Bell System vezetőinek vonakodása volt az automatikus kapcsolásra vonatkozó külső szabadalmak licencelésétől (Green 1995; Lipartito 1994), addig az automatikus kapcsolás ellenzői azzal érveltek, hogy a kapcsolat létrehozása olyan technikai munkát jelent, amelyet a telefonszolgáltatás részeként kell nyújtani, és ezért az operátoroknak kell elvégezniük. Harris F. Hopkins, a *Bell Laboratories Record* egyik cikkének szerzője így fogalmazott: "Az ellenzők úgy vélték, hogy az automatikus kapcsolás az ügyfél szempontjából helytelen. 'A közönség nem fogja eltűrni, hogy maga végezze a működtetést' - mondták" (Hopkins 1960: 83). A második világháború után azonban egyre inkább elterjedtek a helyi telefonhívások indítását automatizáló forgó tárcsás telefonok. Ez az önműködtetésre való áttérés azt mutatja, hogy az automatizálás központi logikája túlmutatott a munka gépekkel való egyszerű helyettesítésén, és általában a munka delegálására vagy újraelosztására terjedt ki, ebben az esetben a szolgáltatóktól az ügyfelek felé. A munka gépekre és ügyfelekre történő kiszervezése a munkaköltségek megtakarítása érdekében, amely a ma digitális kultúra jellegzetességét képezi, már a háború utáni távközlési ágazatban is

gazdasági hajtóerő volt.

Novemberben a Bell újabb innovációt 18,1963,vezetett be a tárcsázási automation: a nyomógombos telefon, amely nem csak egy másfajta módon a

kézi tárcsázás, hanem a tárcsázójelek létrehozásának egy teljesen új módja. A tárcsázó telefonnal történő tárcsázás elektromos impulzusok sorozatát hozta létre, amelyek száma megfelelt a tárcsán feltüntetett számjegyeknek. A nyomógombos telefon gombjának megnyomása azonban két, elektronikus oszcillátorok által generált, hallható szinuszhangból álló, különálló hangpárt hozott létre. Ez az úgynevezett kéthangú többfrekvenciás (DTMF) tárcsázási módszer a Bell állomásfejlesztési osztályának L. A. Meacham által javasolt négyszer négy frekvenciás sémán alapult, bár kezdetben csak hét frekvencia (négy a spektrum alsó és három a felső tartományban) tíz egyedi hangpárt generált (Meacham et al. 1958).³

A módszer a legmodernebb szilárdtest-technológiát alkalmazta, és a Dahlbom et al. 1949; Hopkins 1960 között végzett számos terepi kísérleten alapult (1948-1960 Dahlbom et al. 1949; Hopkins 1960). Egy szám tárcsázásakor a kettős hangok hasznos akusztikus visszajelzést adtak a hívónak. Mindazonáltal a hangokat elsősorban nem az emberi fülnek címezték, hogy hallja, hanem elektronikus szűrőbankoknak, amelyeket a helyi kapcsolóállomásokon, az úgynevezett "hívásközpontokban" helyeztek el dekódolásra. Hogy megakadályozzák

3 A hangok párosítása egyszerű szerkesztési szabályt követett. Minden függőleges oszlophoz más-más hangot rendeltek az alacsony frekvenciatartományban (FA = Hz697, FD = Hz770, FC = Hz 852 és FD = Hz941), míg minden vízszintes sorhoz más-más magasabb frekvenciájú hangot rendeltek (FE = Hz1209, FF = 1336 Hz, FG = Hz1447 és FH = Hz1633). Így módon minden billentyűhöz egy magas és egy alacsony frekvenciájú hang különböző kombinációja van hozzárendelve. A szükséges hardver egy billentyűzetkódolóból és egy hanggenerátorból állt.

a beszélt nyelv, a zajok és egyéb hangok ne zavarják a DTMF-hangok átvitelét, a mikrofont egy gomb megnyomásakor kikapcsolták. Továbbá, mivel a tárcsázójelek hallható "sávon belüli" frekvenciák voltak, a Bell technikusai olyan frekvenciakombinációkat választottak, amelyek a mindennapi életben valószínűleg nem fordulnak elő, hogy a kapcsolóberendezések vevőkészülékeiben ne forduljanak elő hamis pozitív és hamis negatív jelek: "A használt frekvenciák minimalizálják a harmonikusok okozta interferenciát. Ez lehetővé teszi a pillanatnyi korlátozást mindkét frekvenciasávban, és kielégítően véd az esetleges hangzavarok ellen" (Hopkins 1960: 86). Ha valaha is elgondolkodott azon, hogy a nyomógombos hangok miért hangzanak inkább az elektronikus zene túlvilági zöreijéhez, mint a zenei hangszerek harmonikus hangjaihoz, akkor ez az oka.

Az 1964-es New York-i vilákiállításán a Bell bemutatta a DTMF-jelzést a nagyközönségnek Touch-Tone márkanev alatt. A Touch-Tone hívás segítségével az alkönyveltek első ízben tudtak közvetlenül távolsági hívásokat kezdeményezni anélkül, hogy emberi operátorra, mint közvetítőre lett volna szükségük. A nyomógombos telefon bevezetése tehát szorosan összefüggött az elektronikus kapcsolórendszerek (ESS) többé-kevésbé egyidejű bevezetésével a központi kapcsolóállomásokon. Az ESS-ek a digitális "tárolt programvezérlésen" (SPC), a telefonkapcsolás automatizált és számítógépesített felügyeleti módszerén alapultak, amelyet a 1954 Bell Labs matematikusa, Erna Schneider Hoover fejlesztett ki (Harr et al. 1964). Az elektronikus kapcsolás stabilabbnak és megbízhatóbbnak bizonyult, mint a mechanikus módszerek, és szinte teljesen kiküszöbölte az emberi kezelőszemélyzet szükségességét. Mivel a hangalapú tárcsázás létfontosságú volt a digitális kapcsolás bevezetéséhez, a hang használata is a digitalizálás történetének egyik alapvetően fontos lépése volt. A DTMF tárcsázás legfőbb előnye az volt, hogy a hangokat sokkal gyorsabban lehetett létrehozni és érzékelni, mint a rotációs telefonok által generált impulzusjeleket. A megnövekedett sebesség különösen a távolsági hívások és az egyes mellékállomásokra irányuló hívások esetében volt hasznos, például nagyobb szervezeteken belül, mivel ez jelentősen megnövelhette a tárcsázandó számjegyek számát, és ezért időt és türelmet igényelt a hívó féltől. Miközben a Bell a Touch-Tone tárcsázást a híváskezdeményezés kényelmesebb módjaként népszerűsítette ügyfelei körében, a módszert különösen arra szabták, hogy tehermentesítse a kapcsolóközpontokat, ahol a régi, lépésről lépésre történő kapcsolások komoly szűk keresztmetszetekké válhattak a kapcsolási folyamatban, különösen a csúcsforgalom idején. A Touch-Tone jelzéssel a kapcsolóközpontok sokkal több hívást tudtak kezelni sokkal kisebb idő alatt.

A sávon belüli jelátvitel bevezetésének azonban nem célja volt a tárcsázási folyamatot és a hívások kezelését, hanem a telefonhálózathoz kapcsolódó elektronikus, esetleg digitális rendszerekkel való együttműködés új módszereinek lehetővé tételét. Ahogy Hopkins (1960) rámutat, a Bell bízott abban, hogy ezt a "lehetséges jövőbeli szolgáltatást" kínálja, mert a Touch-Tone tárcsázás "az ügyfél számára egy potenciális (lassú sebességű) adatátvitelt" (87)

biztosítana. Az első széles körben elterjedt

nyomógombos telefon a Western Electric 1500-as modellje volt, amely a 0-tól 9-ig terjedő számjegyeknek megfelelő 10 gombot tartalmazott (lásd az 1. ábrát). A későbbi modellekhez a ma már mindenütt jelenlévő szám (#) és csillag (*) jelekkel ellátott gombokat adtak hozzá, hogy lehetővé tegyék és ellenőrizzék a szimbolikus adatok továbbítását. A potenciális távirányítónak vagy végberendezésnek alakított telefonkagylót alfanumerikus információk, például hitelkártyaszámok, vagy helymeghatározó parancsok, például függőleges szolgáltatási kódok (VSC) továbbítására lehetett használni. A VSC-k a csillag (*) és ritkábban a számjellel (#) kombinált számsorozatok. A telefon billentyűzetén vagy tárcsázó tárcsán tárcsázva a VSC bizonyos telefonos szolgáltatási funkciók, például a hívástartás, a hívásátírányítás, a folyamatos újratárcsázás vagy a hívásblokkolás engedélyezésére vagy letiltására használható. A "vertikális" kifejezés a helyi telefonos infrastruktúrán belüli magasabb szintű struktúrákra mutató parancsokra utal, nem pedig a hagyományos telefonszámokra, amelyek "horizontális" egy másik földrajzi helyre vagy kapcsolóközpontra mutatnak. Az AT&T az 1960-as és 1970-es években kezdte bevezetni a VSC-eket "Custom Local Area Signaling Services" (CLASS vagy LASS) kódok néven az előfizetők számára. A Touch-Tone segítségével a hang így az automatizált elektronikus és digitális rendszerekkel való interakciók akusztikus interfészévé vált.

A beszédaktus-elmélet (Austin 1975) szemszövegéből nézve a DTMF-hangok "hangi aktusokként" vagy "hangaktusként" fogalmazhatók meg, azaz olyan hangokként, amelyek nemcsak *reprezentálnak* valamit vagy információt tartalmaznak, hanem *hatnak is*, és következményeik is vannak. A DTMF-hangok, mint a hálózaton keresztül automatizált elektronikus rendszerekkel való kommunikációra tervezett hallható vezérlőjelek, szó szerint *produktív* hangokká váltak a telefonrendszerben, mivel kapcsolókat indítottak el, információkat továbbítottak és távvezérelt automatikus folyamatokat irányítottak. A Bell mérnökei tehát a produktív hangok alapján akarták felkészíteni a telefonrendszert az információs korszakra. Vagy másképp fogalmazva, a Bell mérnökei felismerték, hogy a saját infrastruktúrájuk optimalizálására kitalált technológiát arra is fel lehet használni, hogy új információs szolgáltatásokat fejlesszenek ki és kínáljanak mind üzleti, mind lakossági ügyfeleiknek. Ami a gyakorlatot illeti, a telefonszámok tárcsázásának és más szolgáltatások, például a VSC-k használatának növekvő szokása hozzájárult ahhoz, hogy az előfizetőket az *adatmunka* különböző formáira képezzék ki. Amint fentebb említettük, a Touch-Tone tárcsázás lehetővé tette a végponttól-végpontig tartó jelátvitelt, a vezérlőjelek továbbítását nemcsak a legközelebbi kapcsolóközpontba, hanem a hálózaton belül bárhol lévő kapcsolórendszerekhez is. Ezért a DTMF-módszert olyan infrastrukturális médiumnak kell tekinteni, amely alapvető szerepet játszott a telefon átalakulásában, amely a távolsági beszélgetésre szolgáló speciális célú technológiából a beszéd, az adatátvitel és a távvezérlés általános célú technológiájává vált.

2. Gépekhez beszélni, kódolva beszélni: A hívás felemelkedése Center Industry és a félig automatizált telefonbeszélgetések

Az AT&T az 1960-as évek közepén kezdett új, egyéni hívásslolgáltatást kínálni a Touch-Tone tárcsázás alapján. Ezek olyan új funkciókat tartalmaztak, mint a hívásvárakoztatás, a hívásvárakoztatás és a háromirányú szolgáltatás vagy konferenciahívás. Ezenkívül automatikus adatgyűjtő és információkereső rendszerekkel, mint például a digitális lekérdező/hanghívó rendszer (DIVA) (lásd a 2. ábra szövegdobozát), igyekeztek a megosztott együttműködés új formáit az üzleti világ és a hazai előfizetők számára elérhetővé tenni. A Bell mérnökei a hétköznapi "infrastrukturálás" (Star/

Bowker 2002) számos különböző területen, például a banki, a kiskereskedelmi és a lakossági felhasználásban (lásd a 2. ábrát). J. H. Soderberg számára, aki összefoglalta az érintéshang-alapú szolgáltatások néhány lehetséges kereskedelmi alkalmazását a 1969. kapcsolt telefonhálózatban, utat mutatott a hálózatba kapcsolt eszközök és elosztott szolgáltatások digitális jövője felé:

Az érintésvezérelt telefon vezérlésre való felhasználásának lehetőségei gyakorlatilag korlátlanok. A Touch-Tone telefon nem csak a számítógépes revolúciót hozhatja be az ország minden nappalijába vagy irodájába, hanem számos más, egyszerűbb vezérlési funkciót is el tud látni. Még az is elképzelhető, hogy a jövőbeni rendszerek lehetővé teszik, hogy bekapcsolja a légkondicionálót, hogy otthonában kellemes legyen a hőmérséklet, amikor hazatér egy utazásról, vagy hogy "telefonos bevásárlást" tegyen lehetővé - csupán a telefon néhány gombjának megnyomásával. Az eredmény a mindennapi feladatok drámai egyszerűsödése lehet. (Soderberg 1969: 203)

Ahogy Soderberg víziója mutatja, a Bell mérnökeinek és marketingeseinek meglepően világos elképzelései voltak a digitálisan hálózatba kapcsolt, félautomata szolgáltatásokban rejlő lehetőségekről a telefonos banki szolgáltatások, az elosztott könyvelés, az otthoni vásárlás és az intelligens otthoni alkalmazások terén. Nem utolsósorban a trösztellenes törvények miatt, amelyek tiltották a "továbbfejlesztett" távközlési szolgáltatások keresztfinanszírozását, és nagyrészt megakadályozták az AT&T-t abban, hogy számítógépes üzletágakba merészkedjen, számos ilyen lehetséges alkalmazás több mint egy évtizedig csak az maradt, technológiai lehetőség és jó reklám a Touch-Tone szolgáltatás számára. Jócskán eltartott egészen az 1980-as évekig, mire a nyomógombos telefonok jelentős telítettséget értek el.⁴ De ha megnézzük bármelyik, az 1980-as évek yuppie-kultúráját bemutató korabeli hollywoodi filmet, mindenütt Touch-Tone szolgáltatásokat látunk, és rájövünk: a telefonrendszer volt a szolgáltatások internete a szolgáltatások internete előtt.

Az érintésses telefonon alapuló szolgáltatások azonban rendkívül sikeresnek bizonyultak az ügyfélszolgálati ágazatban, és mélyen kapcsolódtak a call centerek elterjedéséhez. Az 1950-es évek végén és az 1960-as évek elején a

telefonársaságok irodáiban call centerek kezdtek kialakulni a saját ügyfél- és operátori támogatásukra. Két technikai újítás segítette elő a helyiségalapú call centerek elterjedését. Először is, a privát automatizált mellékközpontok (PABX), később privát au- tomated business exchange-nek is nevezett telefonközpontok bevezetése lehetővé tette az automatikus átirányítást egy nagyobb szervezeten belüli mellékszámra, és így felváltotta a telefonos recepciósok vagy ügyeletesek munkáját (lásd Bodin 2002: 20). Röviddel ezután az automatikus híváselosztók (ACD) kiterjesztették a PABX képességét a bejövő hívások gyűjtésére - például a központi

egy szervezet irodájába - és továbbítja őket az ügyfélszolgálati ügyintézők egy csoportjához.⁵ Abban az esetben, ha az összes ügynök foglalt volt, az ACD a bejövő hívást várakozó sorba helyezte, amíg egy ügynök szabaddá nem vált. Az ACD-k funkciói kifinomult algoritmusokon, például Erlang-számításokon alapulnak, amelyek megjósolják, hogy hány ügynökre van szükség, és hogyan lehet a legjobban beállítani a nagyszámú egyidejű hívást a sorba, illetve kiosztani. Az ACD-k tehát a call centerek alapjának tekinthetők, és a mesterséges intelligencia első fajtáját képviselik (a szó tágabb értelmében), mivel automatikus döntéshozatalt vezetnek be a hívások kezelésébe. Az ACD-k azonban nem mesterséges intelligencia a szó szűk értelmében, hanem "feltételes hívásirányítási megoldások, amelyek a ha-akkor feltételeken, vagy a szervezet által előre meghatározott szabályokon alapulnak" (Stanley 2018). Ennek ellenére az ACD-k a mai napig biztosítják, hogy a hívók a lehető leggyorsabban választ kapjanak, és hogy minden ügynök idejét egyenletesen és hatékonyan használják ki.

Mind a PABX-ek, mind az ACD-k csökkentették a központi telefonközpontokban az emberi operátorok és a recepciósok szükségességét, sőt, teljesen feleslegessé tették munkájukat. A kifinomult ACD-k ráadásul a híváskezdeményezés különböző szempontjairól jelentéseket adtak (Bodin 2002: 22-23). Az automatikus híváselosztók különösen értékesnek bizonyultak a nagy hívásmennyiséggel szembesülő szervezetek számára. Az automatikus házon belüli útválasztásnak azonban az algoritmikus eljárások miatt nyilvánvaló hátránya volt, hogy a hívók nem tudtak közvetlenül kapcsolatba lépni egy ügynökkel. Mivel a hívók nem valószínű, hogy kétszer ugyanazt az ügynököt kapják, ez megakadályozta, hogy kapcsolatot alakítsanak ki bizonyos ügynökökkel, és így sokkal kevésbé személyes hívásélményt eredményezett. Az AT&T által bevezetett díjmentes 1-800-as számok alapvetően országos szintű automatikus híváselosztást 1967hoztak létre - a szolgáltatás először egy országos vagy helyi hívásközpontba irányította át a hívásokat, ahol a helyhez kötött ACD-k továbbirányították a hívást a rendelkezésre álló ügynökökhöz.⁶ Az ingyenesen hívható számok az ügyfélszolgálati hívásforgalom soha nem látott mértékű növekedéséhez vezettek, és az anonim felhasználói élményt *de facto* szabványként megszilárdították. Az ACD-k váltak a nagyméretű, decentralizált és földrajzilag elosztott hívásközpontok alapjává. A korai, gazdaságilag sikeresnek bizonyult ACD-megoldások közül az amerikai Rockwell gyártó az egyik legelismertebb. A cég

Galaxy ACD-je, ahogy a készüléket nevezték, tette lehetővé a Continental Airlines számára, hogy 1973-ban telefonos foglalással kezdjen foglalkozni.

Az 1970-es években a DTMF-jelzésben rejlő lehetőségeket felismerték a gyártó- a call center berendezések gyártói. Az ügynevezett interaktív hangválasz (IVR) rendszerek nemcsak a hívások továbbítását automatizálták, hanem a tényleges telefonbeszélgetések bizonyos részeit is. Az ACD-k képesek voltak üdvözlő üzeneteket lejátszani, de a

-
- 4 Az automatikus híváelosztók (ACD-k) feladata a bejövő hívások szűrése, sorrendbe állítása és hozzárendelése a beérkező hívásokhoz.
a legjobb elérhető ügynök.
 - 5 Az ingyenesen hívható szám feltalálója egyszer azt állította, hogy minden, amit feltalált, valójában egy mutató volt a digitális könyvtár.

a hívó fél várakoztatásba helyezésén kívül más funkcióval nem rendelkezik. Az IVR-ekben az előre felvett üzenetek megkérdezték a hívó igényeit, akusztikusan végigvezették a hívót egy menüstruktúrán, és különböző szolgáltatások közül választhatott, amelyeket a hívó a nyomógombos telefon megfelelő gombjainak megnyomásával választhatott ki. Ezek a rendszerek jellemzően félautomata ember-gép rendszerek voltak, IVR-ekkel az elülső oldalon, és emberi ügynökökkel, akik előre meghatározott pontokon vagy akkor vették át az irányítást, amikor az automatizált rendszer korlátokba ütközött. Az emberek és a gépek közötti munkamegosztást úgy valósították meg, hogy a telefonbeszélgetéseket többé-kevésbé redundáns részekre bontották, és az előbbieket automatizálták. A legtöbb ügyfélkérdésre kategóriák és opciók rögzített csoportjai válaszoltak, amelyeket előre felvett üzeneteken keresztül adtak át, amelyekre a hívók DTMF-hangokkal válaszolhattak. Az ügyfél és az adott szervezet közötti kapcsolatot, amely egy IVR-rendszerrel együtt bontakozik ki, tehát úgy tekinthetjük, mint amit Susan Leigh Star és James R. Griesemer "konszenzus nélküli együttműködésnek" nevezett, amely egy közös technobeszélgetési "határobjektumon" alapul (1989).

Az IVR-ek önkiszolgáló funkciói lehetővé tették a helyettesítést, legalábbis részben,

nemcsak az operátori munka, hanem az ügyfélszolgálati ügyintézők által végzett tényleges hang- és tranzakciós munka is. A munkaerőköltségek nyilvánvaló megtakarításán kívül az automatikus call center rendszerek előnye volt, hogy lehetővé tették a kiszolgálási idő kiterjesztését. Az árnyoldala azonban az volt, hogy mivel a hívók már nem is beszéltek emberi ügynökökkel - legalábbis addig nem, amíg a rendszer nem kapcsolta őket hozzájuk -, az IVR-ek még anonimabbá tették a telefonos élményt, mint az automatikus híváselosztók által végzett, látszólag véletlenszerű kiválasztási folyamat. Az évek során a gyártók alternatív beviteli nyelvként hangfelismeréssel egészítették ki az érintéshangot. A hangvezérlés bevezetésének elsődleges célja az volt, hogy az IVR-szolgáltatásokat kiterjesszék a forgótelefonok tulajdonosaira, de az eredmény az lett, hogy a hangfelismerés révén bárki, aki a gépelés helyett a beszédet részesítette előnyben, most már beszélt nyelven keresztül is kapcsolatba léphetett az IVR-rendszerrel. Ez az a pont, ahol a mesterséges intelligencia technikák először lépnek a színpadra.

A legtöbb ilyen áramköri kapcsolt telefonos rendszert azóta felváltotta a csomagkapcsolt, IP-alapú technológia. Történetük ezért - legalábbis bizonyos mértékig - egyben egy félig-meddig elavult, emberi operátorokból és interaktív rendszerekből álló jövő médiakulturális vízióinak régészete vagy rekonstrukciója is. Együttal az analóg és digitális rendszerek hibrid jövőjére is utalnak, amely egyszerre volt analóg és digitális. George Lucas első játékfilmje, a *THX 1138* (1971) példaértékű a telefonos információs hálózatok e korszakának jövőképeit illetően. Lucas egy futurisztikus földalatti társadalom képét festi meg, amelyet kommunikációs és megfigyelési technológiák hatnak át, emlékeztetve George Orwell regényére *1984*. Az érintőkapcsolós telefonok és félautomata rendszerek

kortárs fejlődését a teljes audiovizuális mediatizáció és megfigyelés disztópiájává extrapolálja. A média - a technológiai média - mindenütt jelenlétének benyomása...

A megfigyelést és a megfigyelést tovább erősíti a technikailag mediált kommunikációs helyzetek gyakori megrendezése telefonos és intercom beszélgetések, magnófelvételek, videóközvetítések és CCTV-felvételek formájában.

A *THX 1138* a saját kulturális pillanatára adott válaszként művészi reflexiót alkot az akusztikus média akkori átalakulására, amit Jonathan Sterne "hangszóróskultúrának" nevezett (Sterne 2015: 113). A film technikai kommunikációból és automatizált bemondásokból álló, mon- tázsokkal átszőtt hangzásvilágában felvetődik a kérdés, hogy a szereplők valóban emberi beszélgetőpartnerekkel vagy csupán szalagon tárolt, automatikusan kiváltott válaszokkal interakciónak-e valójában. A telefontechnológiára való utalások már a jelenetek kialakításában is megjelennek: az IMDB trivia szekciója szerint egy Pacific Bell kapcsolóterem szolgált a forgatás helyszínéül:

A végtelennek tűnő Vezérlőterem, ahol az android rendőrök megpróbálják sarokba szorítani THX-et és SRT-t, akik megtudják, hogy LUH-t szerv- újrahasonosítás céljából fogyasztották el, a Pacific Bell Telephone Com- pany San Francisco-i telephelyének kapcsolóterme volt. A Pacific Bell engedélyezte George Lucasnak, hogy ott forgassa a filmet, mert az egész helyiséget és az ott található hardvereket éppen le akarták szerelni, mivel a telefontársaság éppen a hangalapú telefontechnológiára állt át (IMDB 2019).

Lucas még a film címét is a San Franciscó-i telefonszámáról, a 849-1138-ról nevezte el, ahol a THX betűk a számjegyek gombjain található betűknek felelnek meg, és a 8,4,Söt9., a film hangzásvilágában felbukkanó elektro-akusztikus hangeffektek közül sok a telefon tárcsahangokból származik, amelyeket a szerkesztő és hangszerkesztő Walter Murch a musique concrete-ből származó kompozíciós módszerek alkalmazásával manipulált. Az automatikus beszédrendszerek embertelen és névtelen ábrázolása nagyrészt a korai IVR-rendszerek felhasználói élményének rekonstruálásával vagy utánzásával érhető el: a film során automatikusan induló üzenetek és közlemények ismétlődőek és monotonok, és nem hagynak kétséget afelől, hogy a jövő társadalmának polgárainak a rendszerhez kell alkalmazkodniuk, és nem fordítva. A filmet majdnem fél évszázaddal a bemutatása után újra megnézve nem lehet nem asszociálni a jelenlegi mesterséges intelligencia alapú közfelügyeleti rendszerekre, például a kínai Social Credit Systemre.

3. "A beszéd egy kiaknázatlan aranybánya": A mesterséges intelligencia alkalmazása a Contact Centerben és a virtuális hangalapú asszisztenseknél.

A még mindig nyilvánvaló korlátaik ellenére a legújabb beszédfelismerő és beszédszintetizáló rendszerek, mint például a hangalapú asszisztensek,

ismerősebben hangzanak, és kevésbé robotikusak és névtelenek, mint a mantraszerű emlékeztetők és bejelentések, amelyek

a *THX 1138* hangsávját. Az automatikus beszédfelismerés (ASR) korai példái közé tartoznak a beszélt hangok, például számjegyek és szavak korlátozott együttesének felismerésére szolgáló mintaalapú modellek, ahol a kimondott számjegy vagy szó felismerését a tárolt referencia-minták halmazával való korrelációja határozza meg (Davis et al. 1952: 194). Az ilyen alkalmazások ismert korai példái közül a Bell Laboratories "Audrey" (Pieraccini 2012: 55-59) és az IBM "Shoebox" (Dersch 1962) képes volt felismerni a beszélt számjegyeket, és a Shoebox esetében akár korlátozott számú parancsot is, ha azt egy ismerős hang mondta.⁷ A "HARPY", a Carnegie Mellon Egyetemen az 1970-es évek közepén a beszédmegértés kutatására irányuló első ARPA-projekt részeként kifejlesztett beszédfelismerő már képes volt szavak szókinccsének felismerésére 1,011 (Lowerre 1976). Az 1970-es évek végén az IBM Dragon rendszere a rejtett Markov-modelleken alapuló ASR-rendszerek új korszakát hirdette meg, amelynek leszármazottait az 1990-es évektől kezdve a legtöbb IVR-rendszerben használták (Pieraccini 2012).

Amint az előző szakaszban említettük, a beszédfelismerési kutatások az emberi hang lehetséges felhasználását eredményezték az automatizált rendszerek vezérlésére és az in- formáció továbbítására. A beszélt nyelv így az automatizált telefonrendszerekben a DTMF-hangok mellett a pro- duktív hangok alternatív típusát jelentette. A hangvezérlés integrálása a főbb számítógépes operációs rendszerekbe, mint a Windows vagy a MacOS, nem utolsósorban a látássérült felhasználók hozzáférhetőségének javítása érdekében, a beszélgető rendszerek felé mutat, amelyeket ma már az okostelefonok és az intelligens otthoni eszközök jelenlegi alkalmazásaiban és platformjaiban látunk. Napjainkban az automatikus beszédfelismerés, -megértés és -szinkronizálás kombinációját - amely ma már nagyrészt mesterséges intelligencia-megközelítéseken alapul - természetes nyelvi feldolgozásnak nevezzük. A kiterjesztett hangügynök-interakció ezen szakasza felé vezető döntő lépés a gépi tanulás és a mély tanulási technikák alkalmazása volt, amelyek többnyire mély neurális hálózatokon (DNN) alapuló tanulási algoritmusokra támaszkodnak. A korábbi módszerekkel összehasonlítva, a beszédfelismerés és beszédészítés fentebb ismertetett történelmi előzményei nyomán a DNN-ek lehetővé teszik a beszédhang elemzését és feldolgozását sokkal nagyobb pontossággal és természetességgel (vö. Mary 2018: 50). A javulás elsősorban a feldolgozási teljesítmény általános növekedésének, a felhőalapú számítástechnika használatának és a hatalmas mennyiségű képzési adathoz való hozzáférésnek köszönhető. Ez az oka annak is, hogy a nagy technológiai vállalatok az elmúlt években egyre inkább fejlesztették a természetes nyelvi feldolgozást, és AI-megoldásokat kínáltak call centerek és hangalapú asszisztensek számára okostelefonos vagy otthoni használatra.⁸ Speciális alkalmazások és platformok, mint például az Amazon Lex, a Google Dialog-

6 Az "Audrey" név az "automatikus számjegyfelismerés" laza rövidítése.

7 Ezek a rendszerek egyre inkább egy központosított internetes infrastruktúrán alapulnak,

amelyet néhány nagy piacvezető, az Amazon (AWS), a Google (Google Cloud) és a Microsoft (Azure) által nyújtott felhőalapú szolgáltatások uralnak. A beszéd felismerő modellek, az érzelem elemelési metrikák és a

flow, a Facebook Wit.ai, az IBM Watson és a Microsoft LUIS meglehetősen egyszerű megoldásokat kínálnak a társalgási botok létrehozására. Nem meglepő, hogy az AI alkalmazásának egyik fő szakmai területe a contact center-iparág. A Google például azzal büszkélkedik, hogy felhőszolgáltatásai "AI-alapú virtuális ügynököket biztosítanak a contact center számára, beleértve a telefonos társalgási ügynököket, az úgynevezett interaktív hangreakciót (IVR)" (Google 2019).

A call centerek ideális feltételeket kínálnak a hangközpontú mesterséges intelligencia-technológiák bevezetéséhez, mivel - amint azt a szakasz bemutatja - rendkívül2, kompartmentális, folyamatorientált és automatizált beszélgetési környezetet alkotnak, amely az ember-gép integráció hosszú múltra tekint vissza. Az IVR-ek fejlesztői és forgalmazói a kezdetektől fogva olyan rendszereket terveztek, amelyekhez mind az ügyfeleknek, mind az ügynököknek alkalmazkodniuk kell. Most a virtuális asszisztensek legújabb példáival megfigyelhetjük, hogy ez a logika fennmaradt és új, félautomata társalgási környezetekké alakult át: ahhoz, hogy a rendszerek "megértsék" őket, a felhasználóknak alkalmazkodniuk kell a beszédmódjukhoz és a használt szavakhoz. Ezért a telefonos munka IVR-ek és a társalgási AI révén történő automatizálásának folyamatait jobban le lehet írni azzal, amit Hamid R. Ekbia és Bonnie A. Nardi (2017) "heteromációnak" neveztek, azaz "a gazdasági érték kinyerésének a számítógépes hálózatokban alacsony költségű vagy ingyenes munkából" - ebben az esetben a call center ügynökök és az ügyfelek által egyaránt végzett munkáról van szó.

A kortárs kapcsolattartó központ alapfunkciója nem különbözik alapvetően az eredeti, történelmi feladattól, a megkeresések kezelésétől és az ügyfélelégedettség javításától. Az online vásárlás és az e-kereskedelem más formáinak növekedése azonban hatalmas igényt támasztott a virtuális ügyfélszolgálatok iránt, ami egybeesik a természetes nyelvi feldolgozási képességek és a szintetikus beszédmodellek jelentős fejlődésével az elmúlt évtizedben (Kopparapu 2015: 5). Az ilyen optimalizált rendszerek használata nemcsak a híváselosztás és -irányítás automatizálását ígéri, hanem az egyedibb ügyfélinterakciókat is, mint például az összetett háromfaktoros számlahitelesítés, aminek hatására tovább csökken az emberi kiszolgálókkal való közvetlen kapcsolat szükségessége - esetleg addig, amíg végül az emberek közötti beszélgetések a normából a kivételek közé kerülnek.⁹ Az IVR-rendszerek, amelyek kiterjedt döntési fáikhoz hosszú, automatizált, beszélt menüket igényelnek, megjelenésük óta kritikát kaptak személytelenségük és idegesítő jellegük miatt (vö. Smith 2016). Egy másik ok az intelligens személyes as- integrálására.

a szintetikus hangok kialakítása, amelyek a kortárs és jövőbeli autonóm gazdasági ágensek középontjában állnak, így mind az e vállalatok által meghatározott protokolloktól és szabályozásuktól függenek.

8 Ezzel párhuzamos fejlődés történt a szövegalapú chatbotok területén, amelyek teljesítménye

ma már meggyőző a standard felhasználási esetekben (Sheth et al. 2019), bár jelenleg a legtöbb ügyfélinterakció még mindig telefonon történik.

A növérek célja tehát az, hogy az ügyfél számára a "személyes," látszólag egyéni beszélgetés élményét nyújtsák, azzal a céllal, hogy az IVR-rendszerek észlelt hiányosságait kiküszöböljék azáltal, hogy a mögöttes hierarchikus struktúrát elrejtik az ügyfél érzékelése elől. Az Apple Siri például kezdettől fogva szellemes és szórakoztató, személyiséggel rendelkező alkalmazásként lett megjelölve, hogy elhatárolódjon az olyan névtelen automatizált rendszerektől, mint az IVR-ek.¹⁰

Az automatikus beszédfelismerés használata a hagyományos IVR-rendszerek kiegészítésére és az emberi ügynökök helyettesítésére azonban nem a fő célja a mesterséges intelligencia technológia bevezetésének a kapcsolattartó központban. Inkább az AI jéghegy csúcsát jelenti, amely az ügyfél számára általában látható vagy érzékelhető. A mai kontaktközpontban nagy valószínűséggel nem csak egy, hanem egyre több különböző típusú mesterséges intelligencia-megoldással találkozhatunk egyszerre. Az iparág egyik vezető szaklapja, a *Call Center Helper* szerint a mesterséges intelligencia megoldásokat nemcsak a hívások kezelésére használják, hanem egyre inkább arra is, hogy az ügyfélinterakciókból származó adatok rögzítésével, nagy adatelemzés alkalmazásával, az ügyfelek viselkedésének előrejelzésével vagy a tanácsadói teljesítmény nyomon követésével új ismereteket állítsanak elő az ügyfelekről és a call center ügynökeiről (Call Center Helper 2018). Egy iparági képviselő ezért azt jósolja, hogy "a gépekkel való jövőnk a partnerség és a fejlesztés lesz (és kell, hogy legyen), nem pedig az átfogó helyettesítés" (Call Center Helper 2019). A call centerek általában hatalmas mennyiségű tárolt hangfelvétellel rendelkeznek, ami különösen alkalmassá teszi őket az analitikus AI-alkalmazások, különösen a prediktív analitika és a beszédanalitika számára. Ahogy egy iparági fehér könyv fogalmaz: "A beszéd egy kiaknázatlan aranybánya". (CallMiner 2019: 5)

A prediktív analitika lehetővé teszi a hívásközpontok számára, hogy értékes információkat generáljanak valós időben.

Idő, például az ügyfél hajlandósága a tartozás visszafizetésére, az ügyfélszolgálati ügynök hatékonyasága az egyes aggályok kezelésében, valamint a hívó általános hangulata és az előzmények alapján valószínűleg kielégítő intézkedések. A beszédanalitika viszont,

túlmutat a felismerésen, és nem csak a hívó fél által elmondott szavakat értelmezi, hanem azt is, ahogyan ezeket a szavakat kimondja. [Más néven hangelemzés, ez a technológia] olyan tényezőket érzékel, mint a hangszín, az érzelmek, a szókincs, a néma szünetek és még a hívó kora is, és ezeket a tényezőket elemezve a hívókat az ideális ügynökhöz irányítja az ügynökök sikerességi aránya, speciális tudása és erősségei, valamint az ügyfél személyisége és egyéb viselkedési jellemzői alapján. (Stanley 2018, n.p.)

Ez különösen az összes rendelkezésre álló hangfelvétel mindenféle elemzéshez történő visszatekintését érinti, valamint azt a törekvést, hogy az emberi hangnak nemcsak a szemantikai, hanem az érzelmi aspektusait is felismerjék és elemezzék az affektív számítástechnika módszereinek kihasználásával (Picard 1997; Jeon

9 A "Siri" a "beszéértelmező és -felismerő interfész" rövidítése.

A mesterséges intelligencia technológiának a contact centerbe való integrálása felé az a probléma, hogy az csak részben az ügyfélnek, és elsősorban az azt használó vállalkozásnak dolgozik. Amivel érvelni szeretnék, az az, hogy ugyanez vonatkozik az otthoni virtuális hangalapú asszisztensekre is, amelyek esetében a kontaktközpontok a korai bevezetés teszt pályájaként szolgáltak (Davis 2019). Ennek érdekében a mesterséges intelligencia különböző felhasználási módjai az ügyfélszolgálati iparágban rámutatnak a mesterséges intelligencia számos, potenciálisan láthatatlan vagy rejtett felhasználási módjára a hazai hangalapú asszisztensek számára. A hangalapú interfésszel rendelkező intelligens hangszórókat kényelmes interfészként márkázták meg mind a helyi, mind a felhőalapú digitális szolgáltatásokhoz. Ezért úgy tervezték őket, hogy személyiséget szimuláljanak, hogy szórakoztatóbb legyen a használatuk. Nem szabad azonban becsapnunk magunkat azzal, hogy csak egy és csakis egyetlen mesterséges intelligenciával van dolgunk - amelynek működését a mesterséges hang formájában reprezentálják és sűrítik össze. Inkább azt kellene felismernünk, hogy valószínűleg egy tucatnyi másik mesterséges intelligencia-rendszer is lehallgatja és elemzi a hangadatainkból a szolgáltatók felhőszerverére továbbított információkat. Végző soron az intelligens személyi asszisztensek nem pusztán a felhasználókért, hanem rajtuk dolgoznak. A háztartási felhasználók és az irodai dolgozók a hangalapú asszisztenseket a kényelem és az ismétlődő feladatok, például a webes keresés és a napi rutinok elvégzésének hatékonysága miatt fogadják el. A vállalkozások, a technológiai vállalatok, a felügyeleti államok és más szereplők azonban azért versengenek, hogy hozzáférjenek magához a felhasználók hangjához, amelyet rendkívül értékes adatforrásnak - "aranybányának" - tekintenek az AI-alapú elemzések számára.

4. Következtetés

A DTMF-jelzés bevezetésével az 1960-as években a speciális célú telefonvevőkészülékeket általános célú távvezérlőkké alakították át, ami alapvető első lépést jelentett a telefonrendszer hosszú távú átalakulása felé, amely a pusztán kommunikációs eszközből a társzűkös sokoldalú médiumává vált. Nagyjából két-három évtized alatt az IVR-rendszerekkel kombinált érintéshangos hívás lassan megtanította a felhasználókat arra, hogy hogyan lépjenek kapcsolatba a távoli automatikus és félautomata információs rendszerekkel a telefonhálózaton keresztül. Tekintettel arra, hogy ezek a technológiák szinte kizárólag a teljes egészében akusztikus inter- face-okra és egy kis billentyűzetre támaszkodtak, úgy tekinthetjük, hogy a produktív hangok mobilizálása végző soron előkészítette az utat ahhoz, amit visszamenőleg a mindennapi "online gyakorlatok" első generációjának nevezhetünk. A produktív hangok különböző iterációi, ahogyan azt már kifejtettem, ily módon képezték az alapját a távközlésből a távszűkösbe való lassú átmenetnek: először alapvető és operatív értelemben többfrekvenciás

jelek formájában; később a call center ügynökök által végzett hangmunka, előre felvett üzenetek, várakozási zene és más dizájnelemek formájában, amelyek a telefonos várakozási hurkok és akusztikus interfészek részét képezték az automata-

interaktív hangreakciós rendszerek; és végül a természetes nyelvi feldolgozáson alapuló társalgási mesterséges intelligencia rendszerek.

Amint azt a Touch-Tone telefonálás bevezetése megmutatta, a "régi" médiaiparágak, különösen a távközlési ágazat, már a hálózatba kapcsolt információs technológiákon és szolgáltatásokon alapuló jövő megvalósításán dolgoztak. Az ügyfélszolgálati hívások auto- mációja feltárta, hogy az infrastrukturális innovációk hogyan alapozták meg az elektronikus és a megtestesült "adatgyakorlatokon" alapuló új szolgáltatások megjelenését, és hogy ezek az átalakulások az áramköri kapcsolású telefonhálózatokban a személyi számítógépek és az internet elterjedése előtt és jóval a számítástechnikai iparon kívül zajlottak. Azáltal, hogy nyomon követtem a különböző technológiai ágensek és munkaformák közötti kapcsolatokat a call centerek kooperatív csoportjain belül, megmutattam, hogy a hanggal kapcsolatos mesterséges intelligencia rendszerek fejlődését az ember-gép interakció egy nagyobb történetének részeként kell felfogni, amelynek gyakorlatai továbbra is alakítják a felhasználók és a kortárs hangalapú asszisztensek közötti kapcsolatokat. Ez az átalakulás nem annyira egy bomlasztó forradalom, mint inkább a (félautomata ember-gép rendszerek egymást követő kombinációin és újrakombinációin alapuló történelmi folytonosságok, valamint a kooperatív gyakorlatok szorgalmas infrastrukturálásának, hálózatba szervezésének és delegálásának formájában következett be, ami végül a virtuális call center ügynökökhöz és a hazai hangalapú asszisztensekhez vezetett.

A hangalapú asszisztensek és az intelligens hangszórók használata emlékeztet az elvekre.

az önkiszolgáló telefonközpontok rendszerének használatára vonatkozó gyakorlatok. Ezért próbálom őket az otthoni hívásközpontoknak nevezni. Ráadásul az elkövetkező években a hangvezérlés, különösen a kihangosított környezetben, például mozgó járművekben, valószínűleg mindenütt jelenlévő és naturalizált interfész-gyakorlattá válik. A kortárs kapcsolattartó központokban a beszélgetések kezelése és automatizálása a munkaerőköltségek csökkentése és a hatékonyság növelése érdekében már nem az egyetlen motiváció a mesterséges intelligencia megoldások bevezetésére; ugyanilyen fontos a felhasználói adatok elemzése a predikciók készítéséhez és új, kereskedelmileg hasznosítható felismerések előállításához. A virtuális hangalapú asszisztensekben alkalmazott mesterséges intelligenciát ezért nemcsak a mindennapi feladataink kényelmes irányításának új módjainak megteremtésére használják, hanem a vezérlőjelek (azaz a hangbevitel) mint hasznosítható ügyfeladatok adatbányászatára is. A call centerek gyakorlatának tanulmányozása tehát egy módja lehet a hangalapú asszisztensek megértésének, és politikájukat így talán legjobban egy háttorzongató együttműködési paktummal lehet magyarázni: Egyrészt a hangalapú asszisztenseket arra találták ki, hogy segítsenek nekünk, és ezt jól teszik, sőt, a képességeik fejlődésével egyre jobbak lesznek. Másrészt, mivel a virtuális hangalapú asszisztensek digitalizált hangjelzéseinket távoli felhőszerverekre továbbítják feldolgozásra, a felhasználók - képletesen szólva -

minden egyes beszélgetéssel meghívják otthonukba és táplálják a névtelen háttérben működő mesterséges intelligencia rutinjait. A call center rendszerekben leggyakrabban előzetesen rögzített kijelentés a virtuális hangalapú asszisztensekre is érvényes: "A hívását figyelni fogjuk."

Köszönetnyilvánítás

Ezt a kutatást a Német Kutatási Alapítvány (DFG) finanszírozta az "1187Együtműködés médiumai" (Medien der Kooperation) Együtműködési Kutatóközpont A01 projektjének részeként. Szeretnék köszönetet mondani Kyle Stine-nak a kritikai megjegyzésekért és javaslatokért. Szeretnék köszönetet mondani Kyle Stine-nak a kritikai megjegyzésekért és javaslatokért, valamint Thomas Bjørnsten-nek a dolgozathoz adott értékes információkért. Szeretnék továbbá köszönetet mondani Sheldon H. Hochheisernek és Melissa Wassonnak az AT&T Archives and History Center (Warren, NJ) nagylelkű támogatásáért.

Bibliográfia

- Aharon, Dan/Laqab, Daryush (2018): "A Contact Center átalakítása mesterséges intelligenciával". Google Cloud Blog. <https://cloud.google.com/blog/products/gcp/transforming-the-contact-center-with-ai/> (2019. június 10.).
- Austin, John Langshaw (1975): *How to Do Things with Words*. Oxford: Oxford: Oxford University Press.
- Bodin, Madeline (2002): *A Call Center szótár: A Call Center és az ügyfélszolgálati technológiai megoldások teljes útmutatója*. Boca Raton: CRC Press.
- Call Center Helper (2018): "A mesterséges intelligencia 12 legfontosabb felhasználási módja a kapcsolattartásban". Központ." Call Center Helper. <https://www.callcentrehelper.com/12-top-uses-of-artificial-intelligence-in-the-contact-centre-123361.htm> (2019. június 10.).
- Call Center Helper (2019): "Mesterséges intelligencia a kapcsolattartó központban: Mit kell VALÓBAN tudnia." Call Center Helper. <https://www.callcentrehelper.com/artificial-intelligence-contact-centre-should-know-142841.htm> (június 10, 2019).
- CallMinerEureka(2019): "HowAIImprovstheCustomerExperience.RealUseCases of Engagement Analytics & Automation for Contact Center Success". CallMiner. <https://learn.callminer.com/whitepapers/how-ai-improves-the-customer-experience> (2019. június 10.).
- Canalys (2019): "Canalys: Global Smart Speaker Installed Base to Top Million200 by End of 2019." <https://www.canalys.com/newsroom/canalys-global-smart-speaker-installed-base-to-top-200-million-by-end-of-2019> (2019. június 10.).
- Dahlbom, C. A./Horton, Jr., A. W./Moody, D. L. (1949): "Applications of Multifrequency Pulsing in Switching" (A többfrekvenciás pulzálás alkalmazása a kapcsolásban). In: *Trans. AIEE* 68, pp. 392-96.
- Davis, Jessica (2019): "Hangalapú asszisztensek a vállalatoknál". *InformationWeek*. <https://www.informationweek.com/strategic-cio/it-strategy/voice-as>

[sistants-coming-to-the-enterprise/d/d-id/1333642](#) (2019. június 10.).

- Davis, K. H./Biddulph, R./Balashek, S. (1952): "Automatic Recognition of Spoken Digits." (Beszélt számjegyek automatikus felismerése). In: *Journal of the Acoustical Society of America* 24/6, pp. 637-642.
- Dersch, W. C. (1962): "Cipősdoboz: A Voice Responsive Machine." In: *Datamation* 8/6, pp. 47-50.
- Ekbia, Hamid R./Nardi, Bonnie A. (2017): *Heteromation, and Other Stories of Computing and Capitalism*. Cambridge, Mass.: MIT Press.
- Gillespie, Tarleton (2010): "A platformok politikája." In: *Új média és társadalom* 12/3, 347-364. o.: *New Media & Society* 12/3, pp. 347-364.
- Gopinath, Sumanth (2013): *Gopanth Gathanth: The Ringtone Dialectic: Economy and Cultural Form*. Cambridge, Mass.: The MIT Press.
- Green, Venus (1995): "Goodbye Central: A Bell rendszer automatizálása és a "személyes szolgáltatás" hanyatlása, 1878-1921". In: *Technology and Culture* 36/4, pp. 912-949.
- Haigh, Thomas/Russell, Andrew L./Dutton, William H. (2015): "Az internet történetei: Introducing a Special Issue of *Information & Culture*." (Az Információ és Kultúra különszámának bemutatása). In: *Information & Culture* 50/2, pp. 143-159.
- Harr, J. A., E. S. Hoover/Smith, R. B. (1964): "Az 1. számú Ess tárolt program szervezése". In: *Bell System Technical Journal* 43/5, pp. 1923-1959.
- Hopkins, Harris F. (1960): "Tárcsázás." In: *Bell Laboratories Record* 38/3, pp. 82-87.
- IMDb (2019) "THX 1138 (1971)-Trivia." IMDb.com. <https://www.imdb.com/title/tt0066434/trivia> (2019. június 10.).
- Jeon, Myoungsoon (szerk.) (2017): *Emotions and Affect in Human Factors and Human-Computer Interaction*. London; San Diego, Cal: Elsevier/Academic Press.
- Kopparapu, Sunil Kumar (2015): *Call Center Conversations nem nyelvi elemzése*. New York: Springer.
- Lipartito, Kenneth (2003): "A Picturephone és az információs korszak: A kudarc társadalmi jelentése". In: *Technológia és kultúra* 44/1, pp. 50-81.
- Lowerre, Bruce T. (1976): "A HARPYPY beszédfelismerő rendszer. Doktori disszertáció. Carnegie-Mellon University.
- Mary, Leena (2018): *Prosody extraction of Prosody for Automatic Speaker, Language, Emotion and Speech Recognition*. New York: Springer.
- Meacham, L. A./Power, J. R./West, F. (1958): "Tone Ringing and Pushbutton Calling: Two Integrated Exploratory Developments (Két integrált felfedező fejlesztés)." In: *Bell System Technical Journal* 37/2, pp. 339-360.
- Nexidia Interaction Analytics (2017): "www.nexidia.com/media/2522/whitepaper-using-ai-powered-analytics-jan-2017.pdf (június 20 1910.,.).
- Noweck, H. E. (1961): "Az érintéses hanghívás sokoldalúsága". In: *Bell Laboratories Record* 39/9, pp. 312-316.

- Perez, Sarah (2019): "Az amerikai felnőttek több mint egynegyede rendelkezik intelligens hangszóróval, jellemzően egy Amazon Echo készülékkel." TechCrunch. <http://social.techcrunch.com/2019/03/08/over-a-quarter-of-u-s-adults-now-own-a-smart-speaker-typically-an-ama-zon-echo/> (2019. június 13.).
- Picard, Rosalind W. (1997): *Affektív számítástechnika*. Cambridge, Mass.: The MIT Press.
- Pieraccini, Roberto (2012): *A hang a gépben: Building Computers That Understand Speech*. Cambridge, Mass.: MIT Press.
- Schüttpelz, Erhard (2017): "Infrastrukturális média és közmédia". In: *Media in Action* 1/1, pp. 13-61.
- Sheth, Amit/Yip, Hong Yung/Iyengar, Arun/Tepper, Paul (2019): "Cognitive Services and Intelligent Chatbots: Current Perspectives and Special Issue Introduction". In: *IEEE Internet Computing* 23/2, pp. 6-12.
- Smith, Ernie (2016): Ernie Smith: "The History of the Call Center Explains How Customer Service Got So Annoying". Vice. https://www.vice.com/en_us/article/xyg4mn/the-history-of-the-call-center-explains-how-customer-service-got-so-annoying (2019. június 10.).
- Soderberg, J. H. (1969): Soderberg: "Machines at Your Fingertips". In: *Bell Laboratories Record* A 47/7, pp. 199-203.
- Stanley, Robert (2018): "A mesterséges intelligencia átfogó története a call centerben: ACD-től a prediktív analitikáig és azon túl". CallMiner. <https://callminer.com/blog/comprehensive-history-ai-call-center-acds-predictive-analytics-beyond/> (2019. június 10.).
- Star, Susan Leigh/Bowker, Geoffrey C. (2002): "Hogyan kell az infrastruktúrát működtetni?" In: Leah A. Lievrouw/Sonia Livingstone (szerk.): *Az új média kézikönyve: Social Shaping and Social Consequences of ICTs*, London: Sage, pp. 151-162.
- Star, Susan Leigh/Griesemer, James R. (1989): "Az intézményi ökológia, a "fordítások" és a határtárgyak: Amatőrök és szakemberek a Berkeley's Museum of Vertebrate Zoology, 1907-39". *Social Studies of Science* 19/3, pp. 387-420.
- Sterne, Jonathan (2006): "Az Mp3 mint kulturális műtárgy". In: *Új média és társadalom* 8/5, pp. 825-842.
- Sterne, Jonathan (2012): *MP3: Egy formátum jelentése*. Durham: Duke University Press.
- Sterne, Jonathan (2015): "Tér a térben: Mesterséges visszhang és a leválasztható visszhang". In: *Szürke szoba* pp60, pp. 110-31.
- Volmar, Axel (2017): Volmar, Axel: "Formátumok mint az együttműködés médiumai." In: *Media in Action* 1/2, pp. 9-28.

Algoritmikus kereskedés, mesterséges intelligencia és a megismerés politikája

Armin Beverungen

Ebben a fejezetben a pénzügyi piacokon folytatott algoritmikus kereskedésben bekövetkezett változásokra összpontosítok, amelyeket a gépi tanulás és a mesterséges intelligencia (AI) fejlődése idézett elő. A pénzügyi kereskedés hosszú ideig az adatfeldolgozás és a számítás rendkívül kifinomult formái uralták a "kvantok" uralmában. Az elmúlt két évtizedben azonban a nagyfrekvenciás kereskedés (HFT), mint az automatizált, algoritmikus kereskedés egy olyan formája, amely inkább a sebességre és a volumenre, mint az intelligenciára összpontosít, uralta a pénzügyi piacok fegyverkezési versenyét. Szeretném azt sugallni, hogy a gépi tanulás és a mesterséges intelligencia ma megváltoztatja ennek a fegyverkezési versenynek a kognitív paramétereit, és elmozdítja a határokat a nagyfrekvenciás kereskedés (HFT) "buta" algoritmusai és az algoritmikus kereskedés más formáinak "intelligens" algoritmusai között. Míg a HFT nagyrészt a pénzügyi piacok belső adataira és dinamikájára összpontosít, az AI által lehetővé tett algoritmikus kereskedés új formái kibővítik a pénzügyi piacok ökológiáját, mivel az automatizált kereskedés szélesebb körű adatokra, például a társadalmi adatokra támaszkodik az olyan elemzésekhez, mint a hangulatelemzés. Azt szeretném javasolni, hogy e változások politikájának megértéséhez érdemes a megismerésre, mint a pénzügyi piacok csataterére összpontosítani, ahol az AI és a gépi tanulás a megismerés további újraelosztásához és új időbeliségéhez vezet. A megismerés politikájának meg kell küzdenie a pénzügyi piacokon zajló algoritmikus kereskedés homályosságával és időbeliségével, amelyek korlátokat jelentenek a pénzügyek demokratizálásának és társadalmi szabályozásának.

Tudatosság és kapitalizmus

A pénzügyi piacok vitathatatlanul a kortárs kapitalizmusban a megismerés körüli küzdelem élvonalában állnak. Ha a mai kapitalizmust az jellemzi, hogy a pénzügyek elsődleges eszközként szolgálnak az életre gyakorolt erőszakra és az életből való értékkivonásra, valamint arra, ahogyan a tőke felhalmozza és kisajátítja a kognitív képességeket, hogy fenntartsa ezt a kivonást

(Fumagalli/Mezzadra 2010), akkor a pénzügyi piacok

bizonyára kulcsszerepet játszanak ebben a pénzügyi és kognitív kapitalizmusban (Beverungen 2018). A pénzügyi piacok ekkor "kollektív kapitalista agynak" tűnhetnek, amelyen keresztül a tőke kognitív módon szervezi az értékkivonást, amelyet csak a nagyfrekvenciás kereskedéshez (HFT) kapcsolódó "alkalmi és véletlenszerű katasztrófák" akadályoznak (Terranova 2013: 66). A HFT-t pedig úgy is felfoghatjuk, mint a "kibernetikai innováció magas határát" a tőkének az ellenségei és a munkásosztály ellen vívott háborújában, amelyben a számítógépek "a fejlett tőke által használt fegyverek" (Dyer-Witheford 2016: 3551.). Bár a mesterséges intelligencia (AI) és az al-algoritmikus kereskedés közelebbi vizsgálata egy összetett képet fog mutatni, amelyben a kollektív tőkés agy távolról sem érzékelhető, és az osztályharc talán kevésbé látható, mint az egyes tőkések közötti verseny, ez egy fontos elemzési keret, amelyet az elemzésem során szem előtt kell tartani.

Azt azonban, hogy a pénzügyek az érték kinyerésével foglalkoznak, természetesnek vehetjük. Ez az előfeltevés a pénzügyi kereskedő szemszögéből is nyilvánvaló, ahol a pénzügyi piacokon az érték kinyerésének kérdése a helyes üzletkötés kérdése. Ahogy Beunza és Stark érvel: "Mi számít?" - ez az a kérdés, amely "a legtömörebben kifejezi azt a kihívást, amellyel az értékpapír-kereskedők a kvantitatív pénzügyek korában szembesülnek" (2008: 253), és feltehetően minden más pénzügyi kereskedő is, beleértve a nagyfrekvenciás kereskedőket is. A feladat elsősorban az információval kapcsolatos, a kereskedők "az információk virtuális áradatában merülnek el", ahol "a kereskedők számára nem a gyorsabb, magasabb, erősebb - mintha az adatmennyiség problémáját még több adat összegyűjtésével lehetne megoldani -, hanem annak kiválasztása, hogy mi számít, és a kiválasztás értelmezése" (Beunza/Stark 2008: 253). A "számítási gyakorlatok", amelyeket a kereskedők a "mi számít?" kérdésre adott válaszként alkalmaznak, "személyek és eszközök között oszlanak meg" (Beunza/Stark 2008: 254). Beunza és Stark itt egy bizonyos információs problémát feltételez, ahol a feladat a releváns információk kiválasztása, amelyek a pénzügyi kereskedésben számításba vehetők, ami többletet eredményez. Már most láthatjuk, hogy a mesterséges intelligencia és napjaink legfontosabb vívmánya - a mesterséges neurális hálózatok - milyen nagy segítséget nyújthat.¹ Az alábbiakban megvizsgálom, hogy a mesterséges intelligencia különböző típusait hogyan alkalmazták a pénzügyi kereskedésben, és megjegyzem, hogy ezek hogyan változtatták meg a Beunza és Stark által leírt kihívás paramétereit.

Mielőtt azonban továbbmennék, szeretném historizálni Beunza és Stark prem-jét, és kibővítem elemzésüket, ami lehetővé teszi számomra, hogy később visszatérjek a pénzügyi és kognitív kapitalizmus elvontabb politikai elemzéséhez. Mindenekelőtt fontos megjegyezni, hogy Beunza és Stark piacát az információáramlás és a szűrés kognitív kihívása jellemzi.

¹ Kissé mulatságos módon Ray Kurzweil, a futurista és "a mintafelismerési technikák világsője" 1999 óta működtet egy "FatKat" nevű céget, amely "iparágvezető eszközöket épít a kvantitatív

alapú befektetéshez". Erről a cégről keveset tudni, és a honlapja (www.fatkat.com) még mindig 2012-es dátummal rendelkezik Lásd 2001.Patterson 2012: 306.

az információ, hogy a sikeres kereskedés szempontjából hasznos információkhoz jusson, nem történelmi adottság. Eltartott egy ideig, amíg a piacot információfeldolgozóként értelmezték és erre a célra tervezték. Mirowski és Nik-Khah (2017) részletesen feltárta Hayek fejlődő piacelfogásának hatását a közgazdaságtan tudományára és a piactervezés gyakorlatára. Hayek munkásságában legalább három olyan szakaszt azonosítanak, amelyben a piacokat, a tudást és az információt másként értelmezik (Nik-Khah/Mirowski 2019: 38-44). Először a tudást úgy értelmezik, mint valami nehezen felhalmozható, olyan feladatot, amelyet csak a piacok képesek teljesíteni. Ezután a tudás valami hallgatóságos dologgá válik, és ezért olyasmivé, amit csak a piacok tudnak előtérbe helyezni. Végül a tudás információvá válik, valami szuprapersonális, a piacon belül tartózkodó dologgá: "egy újfajta virtuális információ" (Nik-Khah/ Mirowski 2019: 43; kiemelés az eredetiben).

Nik-Khah és Mirowski bemutatják, hogy ezek a piacelfogások befolyásolták a piactervezés különböző iskoláit, amelyekben a közgazdászok a piacok, például a pénzügyi piacok tervezőiként működnek. Az alábbiakban megemlítem, hogy az olyan piactervezők, mint Alvin Roth és mások hogyan vesznek részt azoknak a pénzügyi piacoknak a tervezésében, amelyeken az algoritmikus kereskedés zajlik, és ahol a mesterséges intelligenciát alkalmazzák. A fontos szempont, amelyet meg kell jegyeznünk, és amely egybeesik azzal, ahogyan a piactervezés "a neolib-eralizmus logikusan végigvitt parancsait jelenti" (Nik-Khah/Mirowski 2019: 63), az, ahogyan az emberi tudatosság és megismerés egyre kevésbé válik relevánssá a piacok számára, és végső soron diszkontálódik, a piacot pedig "ember-ma- gyar rendszerként", "hibrid számítási eszközként" képzelik el, "a gondolkodást a dolgokra hártva" (Nik-Khah/Mirowski 2019: 6153.). Ahogy Mirowski és Nik-Khah fogalmaz: "Az ágensek be lennének hajtogatva a személy-gép rendszerbe, és többé nem tartanák ca- pabilisnak, hogy megértsék, miért hozzák meg azokat a döntéseket, amelyeket meghoznak. Gondoljunk a kényszerhelyzetükre úgy, mint mesterséges tudatlanságra". (2017: 238-239). Ironikusnak tűnhet, hogy ezt a pénzügyi kereskedőkre is jellemző "mesterséges tudatlanságot" a gépek mesterséges intelligenciájával kívánják kiegészíteni. De ahogyan alább bemutatom, a mesterséges intelligencia bevetése az algoritmikus kereskedésben pontosan követi a közgazdászok és a piactervezők premisszáit: az információ a piacon található, és a mesterséges intelligencia feladata, hogy azt - az "alfát" - kinyerje, hogy ezzel növelje a kereskedést.

Annak érdekében, hogy megértsük, hogyan alakulnak a piacok mind emberi, mind gépi formában.

ic és számítási, valamint arról, hogy a gondolkodás hogyan "terhelődik át a dolgokra", szeretnék támaszkodni Hayles nemrégiben végzett munkájára a nem tudatos megismerésről (Amoore 2019), mivel ez hasznos módot kínál arra, hogy értelmet adjon annak, hogyan oszlik meg a megismerés a pénzügyi piacokon (lásd még Beverungen/Lange 2018). Hayles különbséget tesz a "gondolkodás" és a "megismerés" között, azt sugallva, hogy a gondolkodás az emberi, tudatos megismerés, míg a megismerés "egy sokkal szélesebb körű képesség, amely

bizonyos fokig minden biológiai életformában és számos technikai rendszerben jelen van" (2017: 14). A megismerést "olyan folyamatként definiálja, amely az információt olyan kontextusokban értelmezi, amelyek jelentéssel kapcsolják össze" (Hayles 2017: 22), és amely nem-tudatosan is végbemehet. Felajánlja, hogy helyettesítse

az emberi és nem emberi közötti különbséget a "megismerők" és a "nem megismerők" megkülönböztetésével, ahol az első kategóriába tartoznak az emberek, a biológiai élet, de az olyan technikai rendszerek is, mint például a mesterséges intelligenciát alkalmazó rendszerek (Hayles 2017: 30). Fontos, hogy ez lehetővé teszi számunkra, hogy megértsük a pénzügyi piacok összetételét, amelyet számos megismerő (emberi és gépi) alkot, hogy megvizsgáljuk, hogyan oszlik meg a megismerés e megismerők között (tudatos és nem tudatos kognícióként egyaránt), és hogy feltárjuk, milyen autonómiát kapnak a gépek az algoritmikus kereskedésben "olyan zsebekben, amelyeken belül a technikai rendszerek autonóm módon működnek" egy "punctuated agency" (Hayles 2017: 32) keretében.

Hayles (2017: 142-177) saját elemzést kínál a pénzügyekről és a HFT-ről, és azt javasolja, hogy a HFT-t "olyan evolúciós miliőnek tekinthetjük, amelyben a tudatosság helyett a sebesség vált fegyverré a nem tudatos kognitív fegyverkezési versenyben - egy olyan fegyver, amely azzal fenyeget, hogy autonóm pályán halad egy olyan időbeli rendszerben, amely a közvetlen tudatos beavatkozás számára elérhetetlen" (2017: 165). A következő fejezetekben Haylesre és a Lange-val közösen végzett korábbi munkámra (Beverungen/ Lange 2017; 2018) építve szeretném megvizsgálni, hogyan alakítja ezt a "nem-tudatos kognitív fegyverkezési versenyt" a mesterséges intelligencia. Azt fogom javasolni, hogy az AI egy másik fegyvert - az okosságot - kínál ebben a versenyben a gyorsasággal szemben, amely eltolja a pénzügyi piacok időbeli és kognitív paramétereit, és amelynek további értelmet adhat, ha a pénzügyi és kognitív kapitalizmus körüli vitákat szem előtt tartjuk.

Nagyfrekvenciás és kvantitatív kereskedés

A pénzügyi piacok kereskedési platformjainak automatizálása előtt "a piac kakofóniáját és a kereskedelem látszólagos véletlenszerűségét" főként az emberi szocialitás koordinálta; ma ez "az anonim befektetők tömegeitől érkező megbízásokat kódoló, pontozott elektronikus jelek kezelésének kérdése", amelyet "a mai tőzsdék anyagi alapjait képező fürge algoritmusokkal, kifinomult számítógépes processzorokkal, feltört routerekkel és speciális távközlési rendszerekkel való játszadozással" (Pardo-Guerra 2019: 23) értek el. A manuális kereskedés még mindig létezik, bár minden megbízást automatizált platformokon keresztül kell végrehajtani, és az algoritmikus kereskedés teszi ki a finanszírozási piacokon folytatott kereskedés nagy többségét. Kirilenko és Lo az algoritmikus kereskedést "matematikai modellek, számítógépek és telekommunikációs hálózatok felhasználásaként határozza meg, hogy automatizálják a pénzügyi értékpapírok vételét és eladását" (2013: 52). Az elmúlt két és fél évtizedben a felemelkedését elősegítette, hogy a pénzügyi rendszer egyre összetettebbé vált, "egy sor áttörés történt a pénzügyi piacok kvantitatív

modellezésében", valamint "a számítástechnika terén elért áttörések szinte párhuzamos sora" (Kirilenko/Lo 2013: 53). A piacok automatizálódtak, a kereskedési stratégiák számítógépvezéreltek, és a kereskedést nagyrészt algoritmusok hajtják végre.

A pénzügyi piacok már az automatizált kereskedési platformok bevezetése előtt is lehetőséget kínáltak a gyorsaságon és az okosságon alapuló kereskedési stratégiákra, és a megismerés bizonyos formáit feltételezték. A tick-er tape bevezetése, ahogyan azt például Preda (2006) tárgyalja, megváltoztatta a tőzsde időbeli rendszerét, mivel az árfolyamváltozások folyamatos adatáramlását kínálta, minden eszközzel és céllal valós időben: "a papírcetlik rongyos időszerkezetét felváltotta a ticker tape sima, megszakítás nélküli, egyedi ideje" (Preda 2006: 767). A futószalaghoz diagramok és a vizualizáció más formái, valamint "diszkurzív módok" is társultak, amelyek "a chartot mint kognitív instrukciót támogatták, ami viszont tekintélyt kölcsönzött a részvényelemzőnek, mint az egyetlen elég képzettnak ahhoz, hogy a szaggatott vonalakban felfedezze a piac igazságát" (Preda 2006: 770). A futószalag gyorsasága önmagában nem vezetett versenyelőnyhöz; a részvényelemző okosságára volt szükség ahhoz, hogy hozzáférjen a piac igazságához és annak alapján cselekedjen. A sebesség és az okosság gazdaságossága tovább fejlődik, például a Reuters Stockmaster árfolyam-lekérdező szolgáltatásának bevezetésével (Reuters Stockmaster) vagy 1964 az első automatikus jegyzési rendszer, a NASDAQ elindításával (1972 lásd Mirowski 2007: 216), és a stratégiák differenciálódását eredményezi az algoritmikus kereskedésben.

A futószalag, és a piaci infrastruktúrák, mint például a távközlési a világot behálózó grafikonok már előrevetítik, hogy milyen infrastrukturális beruházásokra van szükség a HFT-hez, mint az algoritmikus kereskedés egy olyan formájához, amelyet nagy sebességű és nagy volumenű kereskedés jellemez. A HFT az 1980-as évek vége óta kulcsszerepet játszik a pénzügyi piacok automatizálásában. Mackenzie és Pardo-Guerra (2014) például felidézi az 1995-ben indított új elektronikus kereskedési platform, az Island szerepét, azt, hogy hogyan hívta ki a meglévő, nem teljesen automatizált kereskedési platformokat, és hogyan vezette be már akkor az automatizált kereskedési platformok olyan kulcsfontosságú aspektusait, mint az ultragyors megfeleltetési motorok, a finomszemcsés árazás vagy a ko-lokáció. Azt is elmesélik, hogy milyen szimbiózis volt a kapcsolat az Island és az Automated Trading Desk között, az egyik első HFT vállalat között, amely 1989-ben kezdte meg a kereskedést, és hamarosan a legnagyobb ügyfelévé vált. MacKenzie részletezi, hogy az Automated Trading Desknek hogyan sikerült bricolage révén HFT-vé válnia, többek között "ok-okozati szerepet" játszott az all-to-all piacok bevezetésében, előmozdította a kereskedés számítógépesítését, és kidolgozta a HFT nagy volumenű és különleges piaci árfolyamokon alapuló üzleti modelljét (MacKenzie 2016: 175, 180). MacKenzie összefoglalja: "Az algoritmusok használata hozzájárult az algoritmusokhoz anyagilag jobban illeszkedő piacok létrehozásához" (2016: 190). Az ezt követő HFT "fegyverkezési verseny" mára a pénzügyi piacok "piackialakításának állandó elemévé" vált (Budish et al. 2015: 1553).

Az olyan infrastrukturális beruházások révén, mint az optikai szálak vagy a mikro-és

a kereskedési helyszínek közötti hullámkapcsolatok, a ko-lokációs központok és még a HFT-re optimalizált számítógépes architektúra is (Zook/Grote 2017; MacKenzie et al. 2012), a piacok kialakítása és időbeli rendszere információszimmetriát eredményez.

metrikák, amelyek lehetővé teszik a nagy sebességen alapuló, milli-, mikro- vagy akár nanoszekundumokban működő kereskedési stratégiákat (Markoff 2018) és a pénzügyi piacok "vízvezeték-rendszerének kijátszását" (Toscano 2013). A HFT-ben a sebesség végső soron felülmúlja az okosságot. Ennek következtében a kereskedési algoritmusok meglehetősen "buták": a sebességhez alacsony késleltetési időre van szükség, és minden információfeldolgozás időbe telik. A HFT-algoritmusokat ezért a lehető legegyszerűbbnek kell tartani ahhoz, hogy gyorsan reagáljanak az információváltozásokra és automatikusan végrehajtsák a kereskedést, ezért állandó emberi felügyeletet igényelnek (Beverungen/Lange 2018: 86-91). Ahogy Arnoldi (2016: 46) fogalmaz, "a kereskedés 'naiv' algókra bízása [...] gazdasági szükségszerűségből adódó választás lehet a nagyfrekvenciás kereskedők számára [...]. Durván fogalmazva, az algók gyorsabbak lesznek, de nem okosabbak". A HFT a pénzügyi piacok merőlegességét kihasználva a belső piaci dinamikára és az információk aszimmetriáira összpontosít, és mikro- vagy mostanra már nanoszekundumos időbeli előnyökkel dolgozik, ezért nem engedheti meg magának, hogy időt szánjon az olyan komplex számításokra, mint amilyenek a mesterséges intelligenciához szükségesek. Az algoritmusok "punctuated agency", azaz az a tér, amelyben "következtetéseket vonnak le, kontextusokat elemeznek és döntéseket hoznak milliszekundumokban" (Hayles 2017:1 42), egyszerűen nem hagy időt az AI számára.

Ez nem jelenti azt, hogy a mesterséges intelligencia nem tájékoztathatja a HFT-stratégiákat. Például az Au-

tomated Trading Desk, alapvető AI, mint például a lineáris regressziós egyenleteket használták az árak előrejelzésére: a gép kiszámítja a kérdéses részvény "kiigazított elméleti értékét", az árfolyamának előrejelzését 30 másodpercre a jövőben", olyan piaci adatok alapján, mint "a [legjobb] ajánlat mérete a [legjobb] offer méretéhez képest, valamint 'a részvény tranzakciós árának rövid távú trendváltozója". (MacKenzie 2017: 182-186). Ez ma már aligha számítana mesterséges intelligenciának, de korai példát szolgált arra, hogy milyen modellek és számítások kerültek a HFT-algoritmusok tervezésébe. Az Ann-Christina Lange által végzett terepmunka során a nagyfrekvenciás kereskedők arról számoltak be, hogy évekbe telik, mire a mesterséges intelligencia relevánssá válik a HFT számára, és alkalmazása csak kísérleti stádiumban van (Beverungen/Lange 2018: 89). A megerősítéses tanuláson, mély neurális hálózatokon vagy tekervényes neurális hálózatokon alapuló HFT-megközelítéseket kidolgozó legújabb tudományos munkák (pl. Kearns/Nevmyvaka 2013; Arévalo et al. 2016; Ganesh/Rakheja 2018) hasonlóan azt sugallják, hogy sok a kísérletezés, de kevés a megvalósítás. Egy nemrégiben készült iparági jelentés szerint míg a HFT "a sebességről szól, a gépi tanulás a mélységről és a betekintés szélességéről", és bár a sebesség még mindig számít, "ez egy másfajta sebesség", mint a HFT (McCauley 2016: 4, 7).

Bár Scott Patterson *Sötét medencék* című könyvének címe: *A hagyományos mesterséges intelligencia felemelkedése...*

ing Machines and the Looming Threat to Wall Street (2012) azt sugallná, hogy a HFT

nagyrészt mesterséges intelligencián alapul, nem mindig világos, hogy mi számít mesterséges intelligenciának, és példái vagy a kvantitatív pénzügyekhez inkább kapcsolódó kereskedési stratégiákkal foglalkoznak, vagy olyan példákkal, mint a Trading Machines, amely a 2000-es évek végén szakértői rendszerekre épülő automatizált kereskedési stratégiát működtetett, de amely "egy lumber- ing technósbéka volt a tőzsdén felemelkedő újfajta sebességu botokhoz képest".

(Patterson 2012: 38). A kvantitatív pénzügyekben tágabb értelemben az olyan fejlesztések, mint a portfólióoptimalizálás elmélete, a tőkeeszköz-árazási modell és - ami talán a legfontosabb - a Black-Scholes-féle opcióárazási formula (Kirilenko/Lo 2013: 53-55), számítási eszközöket kínáltak annak eldöntésére, hogy milyen pénzügyi eszközökbe érdemes befektetni, hogyan kell kockázati stratégiákat kidolgozni és hogyan kell árazni az olyan pénzügyi eszközöket, mint az opciók. Ez tette lehetővé, hogy a "kvantok" az 1980-as évektől kezdve meghódítsák a Wall Streetet (Patterson 2010), többnyire a fedezeti alapok részeként, és pénzügyi modelljeik képére formálják a pénzügyi piacokat (MacKenzie 2006). A kvantitatív kereskedés ma már algoritmikusan is vásárol, azaz a megbízások végrehajtása automatikus, és a kereskedési döntések nagy részét is algoritmusok hozzák. Bár sok fedezeti alap is a mikroszekundumokban kifejezett gyors kereskedelemre specializálódott, a HFT-vel ellentétben itt a hangsúly nem pusztán a sebességre, hanem az okosságra helyeződik, és nem annyira a nagy volumenű, nagy sebességű kereskedés során a pénzügyi piacok vízvezetékrendszerének kihasználására, mint inkább az információs aszimmetriák kihasználására a másodpercek helyett órák, napok vagy hetek tartási idejével működő kereskedésben.

Bár a fedezeti alapok és kvantitatív kereskedők rendkívül titokzatosak, a mesterséges intelligencia alkalmazásának néhány példája ismert, és a közelmúltban széles körben elterjedt használatra utal. Például a Renaissance Technologies, az egyik legnagyobb és "sokak által a világ legsikeresebb hedge fundjának tartott" (Patterson 2012: 107), "a pénzügy legfeketébb dobozának" is nevezett (Burton 2016) Renaissance Technologies nagymértékben az amerikai kormányzat kriptográfusaiból és az IBM beszédfelismerő programjából toborozta munkatársait (Patterson 2012: 107-117). Egyik szakértőjük Robert Mercer volt, aki az 1970-es években Frederick Jelinek beszédfelismerő csapatának tagjaként dolgozott a Brown-féle klaszterezésen.² Vagy vegyük Haim Bodeket, aki a Hull Tradingnél, egy kvantitatív algoritmikus kereskedéssel foglalkozó cégnél dolgozott egészen addig 1997, amíg a Goldman Sachs meg nem vásárolta (1999Patterson 2012: 28-30). Bodek korábban a csalások felderítésével foglalkozott, és gépi tanulási ismereteit a Hullnál kamatoztatta (Patterson 2012: 28), majd később megalapította a Trading Machines-t, amely az 2011 egyike első teljesen automatizált és magasabb frekvenciájú kereskedési 2007vállalkozásként működött (Patterson 2012: 32-60).³ Vannak újabb példák is *Patterson Dark Pooljai* között, például az Apama, egy "komplex eseményfeldolgozó" motor, amelyet a Software AG alapított és 1999 vett át (2013Patterson 2012: 62), amely már rámutat arra, hogy a kvantitatív kereskedés a piaci és a piaci adatokon túl az "alternatív" adatok szélesebb körét is felöleli.

2 Robert Mercer ma már hírhedt a jobboldali politikában való szerepvállalásáról, például Donald Trump és a Brexit támogatásáról, valamint a Cambridge Analytica-botrányban való részvételéről. 2017Politikai nyomásra lemondott a Renaissance-tól. Lásd Cadwalladr 2017.

- 3 Haim Bodek a Wall Street talán leghíresebb feljelentője, mert felfedett egy titkos megbízástípust, amelyet a nagyfrekvenciás kereskedők használtak, és amely tönkretette Bodek saját kereskedési stratégiáit a Trading Machines-nál. Bodek a *The Wall Street Code* (2013) című dokumentumfilm főszereplője.

kereskedési adatok - különösen a hírek és a társadalmi adatok - elemzése és a kereskedési stratégiákba való visszacsatolás céljából.

Az algoritmikus kereskedésben a mesterséges intelligenciával kapcsolatos három nagy, a tárgyalásom szempontjából releváns jelenlegi fejlemény van. Először is, a kvantitatív kereskedés automatizálása felé haladunk, azaz a számítás felhasználásával mind a megbízások leadására, mind a stratégiák kiszámítására, hasonlóan a HFT automatizálásához. Úgy tűnik, hogy egyes vállalatok ezt a stratégiát támogatják a HFT-egységek felvásárlásával, például a Citadel 2016-ban megvásárolta az Automated Trading Desket. A Rebellion Research volt talán az első teljesen automatizált hedge fund, amelynek Bayes-hálózatokon alapuló "Star" algoritmus 2005 óta önállóan kereskedik, és a decemberben bevezetett frissített "Star 2.0" (2016Patterson 2012: 323-335; Metz 2016). Egy másik friss példa az Aidyia, egy másik teljesen automatizált AI fedezeti alap, amely "a mesterséges intelligencia több formájára támaszkodik, köztük a genetikai evolúció által inspirált és egy másik valószínűségi logikán alapuló" (Metz 2016). Az azonban továbbra is kérdéses, hogy a kereskedés itt milyen mértékben automatizált valóban, és úgy tűnik, hogy az iparág felismerte az AI-ra való túlzott támaszkodás és az AI-ba vetett "helytelen bizalom" veszélyét, valamint az emberekre való igényt (McCauley 2016: 14, 16). Ahogy a HFT esetében, ahol a kereskedők valószínűleg nem hagyják felügyelet nélkül algoritmusait (Beverungen/Lange 2018), az esetek itt is hasonlóak lehetnek a Trading Machines esetéhez, ahol Bodek is folyamatosan szuper- vizálta a volatilis piacon működő algoritmusait: "Bodek inkább a saját agyában bízott. Miközben olyan mesterséges intelligencia-módszereket használt, mint a szakértői rendszerek az algói megalkotásához, ő inkább a kereskedési nap során végig fenntartotta az irányítást. Ezért soha nem hagyta el a helyét, még mosdószünetre sem". (Patterson 2012: 38; lásd még Satariano/ Kumar 2017). Mindazonáltal ez az automatizálás további elmozdulást jelez a pénzügyi piacokon a ma- chine-machine ökológia felé.

Másodsor, míg az Aidyia és a Rebellion Research viszonylag kicsi, addig a nagy fedezeti alapok nagy többsége is azt állítja, hogy ma már mesterséges intelligenciával dolgozik (lásd pl. Satariano/Kumar 2017 a Man Groupról), és a jelenleg AI-t fejlesztő vállalatok és kutatóintézetek, valamint a fedezeti alapok között jelentős az ex-váltás. David Ferruci, az IBM Watson fejlesztője 2012-ben az IBM-től a Bridgewater Associates vezető technológusa lett (Vardi 2016). Li Deng a Microsoft mesterséges intelligenciáért felelős vezető kutatói posztjáról a Citadelhez költözött in, hogy a mesterséges intelligencia vezetője 2017 legyen. Pedro Domingos, a *The Master Algorithm* (2015) című könyv szerzője és a markov logikai hálózatok szakértője csatlakozott a D.E. Shaw-t a gépi tanulási kutatócsoport 2018 vezetésére. Ezek a magas szintű mozgások azt sugallják, hogy a fedezeti alapok kulcsszerepet fognak játszani a mesterséges intelligencia fejlesztésében és politikájában az elkövetkező évtizedekben, olyan intézményeken keresztül is, mint az Ox- ford-Man Institute of Quantitative Finance, és ez azt sugallja, hogy a különböző típusú mesterséges intelligenciákat, amelyekben ezek a kutatók szakértelemmel rendelkeznek, széles

körben fogják alkalmazni az algoritmikus kereskedésben. Ez azonban nem jelenti azt, hogy a mesterséges intelligencia alkalmazása az algoritmikus kereskedésben egyszerű vagy egyszerű lesz. Li Deng például a következőket javasolja

hogy legalább három kihívás van: alacsony jel-zaj arány az elemzett információkban a minták felismeréséhez; erős nem-stacionárius, sok hamis adattal, amelyeket ki kell küszöbölni; és az adatok sokfélesége, a beszédttől a szövegen át a képekig, amelyeket össze kell vonni és elemezni kell (Deng 2018; lásd még Frontiers A.I. 2018). Mégis, ez jelentős változást jelent a pénzügyi piacok kognitív ökológiájában, a mesterséges intelligenciát a kereskedés gyorsabbá és okosabbá tételére használják.

Harmadszor, jelentősen bővülnek azok az adatforrások, amelyekkel az algoritmikus kereskedés működik, és amelyekből a kereskedési lehetőségeket kínáló mintákat igyekszik kinyerni, ami a kereskedési stratégiák differenciálódásához vezet (McCauley 2016: 4). A HFT-ben az adatforrások a piaci adatok egyértelmű készletére korlátozódnak, amelyek többnyire azon kereskedési platformok megbízási könyveihez kapcsolódnak, amelyekben a nagyfrekvenciás kereskedők tevékenykednek, más algoritmikus kereskedés pedig az olyan vállalatok, mint a Reuters vagy a Bloomberg által szolgáltatott piaci és gazdasági adatok viszonylag korlátozott készletére támaszkodik. Napjainkban azonban az adatforrások megsokszorozódnak, ahogyan azok a vállalatok is, amelyek adatfolyam- és elemzési szolgáltatásokat kínálnak az algoritmikus kereskedéshez, különösen a közösségi médiával kapcsolatban. Az olyan fedezeti alapok, mint a BlackRock, átnézik a közösségi médiát és figyelik a keresőmotorokat, hogy segítsék befektetési döntéseiket (De Aenlle 2018), és vannak olyan vállalatok, mint az EquBot, amelyek saját AI és az IBM Watson segítségével "cikket és hírforrásokat millióit elemzik, hogy katalizátorokat és eseményeket tárjanak fel a piaci felértékelődés valószínűségének maximalizálása érdekében" ("Artificial Intelligence (AI) and the Technology behind EquBot"), beleértve a piaci hangulelemzést is (De Aenlle 2018; McCauley 2016: 8). Vannak olyan vállalatok is, mint a Quandl, a RavenPack, az Eagle Alpha vagy a DataMinr, amelyek adatelemzési szolgáltatásokat kínálnak az algoritmikus kereskedéshez. A DataMinr például "alternatív adatokra" specializálódott, mint például "közösségi média, műholdas képek, időjárás adatok és egyéb" ("AI Data Tips for Traders | Dataminr"), és azt sugallja, hogy a kereskedők közel 80%-a már ilyen "alternatív" adatokat használ ("Report: Investors Embrace Alternative Data | Dataminr").⁴ Az algoritmikus kereskedés adatökológiájának ilyen mértékű bővülése a mintafelismeréshez mesterséges intelligenciát igényel, és egy emberi felismerő számára lehetetlen lenne mindezen információk figyelembe vétele.⁵

4 A Palantir Technologies, a Cambridge Analytica botrányban való részvételéről elhíresült Palantir Technologies nagy adatelemző vállalat a Palantir Foundryn keresztül szolgáltatásokat kínál a pénzügyek számára is, amelyeket azonban nagyrészt csalásfelderítésre használnak. A

Palantir működéséről lásd: Munn 2018: 27-56.

- 5 Az olyan kritikai művészeti projektek, mint a Rybn ADM Trading Bot (lásd <http://www.rybn.org/ANTI/ADM8/>) és Derek Curry hacktivistája, taktikai médiaprojektje, a *Public Dissentiment* (lásd <http://www.publicdissentiment.org/>) a pénzügyi piacok megzavarására és a "tudatosság növelésére törekszik, hogy a közösségi média mostanra hogyan kapcsolódik össze a részvénykereskedelemmel" (Curry 2018: 108).

Változó kognitív ökológiák

Talán nem meglepő, hogy a mesterséges intelligencia az algoritmikus kereskedés központi eleme, és hogy a mesterséges intelligencia újabb fejlesztéseit, például a mélytanulás fajtáit az algoritmikus kereskedésben adaptálják. Ha a piacokat úgy alakították ki, hogy ne helyezzenek hangsúlyt az emberi megismerésre, és feltételezzék, hogy az igazság a piacot jelentő információfeldolgozóban rejlik, akkor nem meglepő, hogy az emberi megismerést a mesterségesen intelligens gépek által gyakorolt nem tudatos megismerés még inkább megkerüli. Ami talán még meglepőbb, hogy a tudatos, emberi megismerés még mindig központi szerepet játszik az összes algoritmikus kereskedési forma esetében, kivéve annak legautomatizáltabb változatait. A HFT-ben a "nem-tudatos kognitív fegyverkezési verseny" (Hayles 2017: 165) azt jelentette, hogy az emberi tudatos megismerés helyébe a gépi nem-tudatos megismerés *sebessége lépett*, ugyanakkor a "tudatosság költségeit" (Hayles 2017: 41-45) - a lassú reakcióidőt, az emberek korlátozott racionalitását stb. - a "nem-tudatos megismerés költségeivel" (Beverungen/Lange 2018: 80) szemben kellett egyensúlyozni, ami pénzügyileg katasztrofálisnak bizonyulhatott. A mesterséges intelligenciába és a ma- chine learningbe történő beruházások határozottan eltolták a pénzügyi piacok kognitív ökológiáját afelé, hogy ne csak a *sebességre* helyezzenek prémiumot - a HFT továbbra is kihasználja a pénzügyi piacok vízvezetékrendszerét -, hanem az *okosságra* is - *egy olyan* mesterséges okosságra, amely még inkább kihívást jelent az emberi megismerés számára. Az emberi tudat most már nem tud lépést tartani sem azzal a sebességgel, amellyel a nagyfrekvenciás algoritmusok kereskednek, sem azzal az intelligenciával, amellyel a mesterségesen intelligens gépek értelmezik az adatokat és találják meg a kereskedési stratégiák számára előnyös mintákat. A nem tudatos megismerés e másfajta fajtájának - az algoritmikus kereskedelemben szerepet játszó különböző mesterséges intelligenciáknak - a költségeit még fel kell sorolni.

Biztonságosnak tűnik feltételezni, hogy a kognitív ökológia egyik költségét az algoritmikus kereskedés által termelt piaci volatilitás. Már most is rengeteg példa van arra, hogy mind a kvantitatív, mind a HFT milyen módon okozott összeomlásokat (lásd Kirilenko/Lo 2013: 60-67 a nem teljes listát). A kvantitatív kereskedés esetében a legsúlyosabb esemény a 2007. augusztusi "kvant-rengés" volt, az akkor kibontakozó pénzügyi válság kellős közepén. A látszólag csekély piaci nyomás ellenére a fedezeti alapok összehangolt kényszertörlesztésekbe és az azt követő tőkeáttétel-csökkentésbe kezdtek, ami a fedezeti alapok számára hatalmas veszteségeket eredményezett (Kirilenko/Lo 2013: 61-62). A HFT esetében a leghíresebb példa a május 6 f lash crash 2010., amelyben a Dow Jones ipari átlag "egész története során a legnagyobb egynapos pontcsökkenést szenvedte el intraday alapon, és a világ néhány legnagyobb vállalatának részvényárfolyamai érthetetlen árfolyamokon kereskedtek", mindez nagyrészt a nagyfrekvenciás algoritmusok negatív kölcsönhatásának köszönhetően (Kirilenko/Lo 2013: 62-63; Borch 2016). A f lash összeomlása azonban nem egyedi esemény volt: Johnson et

al. több mint 18.000 "ultragyors szélsőséges eseményt" azonosított egy ötéves időszakon belül, ami szerintük összhangban van a "versengő gépek kialakulóban lévő ökológiájának megfigyelésével, amelyben ragadozó algoritmusok "tömegei" vannak" (2013: 1). Továbbá, az egyik

A pénzügyi piacok kibővített adatökológiája által okozott volatilitás példáját Karppi és Crawford (2016) a "hack crash" -ként írta le, amelyben egy hamis hír bejelentése a Twitteren a pénzügyi piacokon a DataMinr által táplált automatizált kereskedési algoritmusok által okozott rázkódáshoz vezetett. Ezek a példák azt sugallják, hogy a "kivételes helyzetek, ritka események és fekete hattyúk rejtélye" - amely már a származtatott ügyletekhez és a pénzügyi piacok más aspektusaihoz kapcsolódik - megmarad, és hogy a "sötét és zavaros empirizmus terepét" (Vogl 2015: 15), amely a pénzügyi piacokat jellemzi, az AI csak tovább súlyosbítja.

Mindezek a példák azt is mutatják, hogy a volatilitás nagy része a pénzügyi piacokon működő (mind "buta", mind "intelligens") automatizált kereskedési algoritmusok kölcsönhatásából ered. Ezek az automatizált ügynökök jelentősen hozzájárulnak ahhoz, hogy a pénzügyi piacokat olyan interaktív dinamikák jellemzik, mint az utánzás (pl. Borch 2016; Lange 2016). Mégis "még mindig messze vagyunk attól, hogy ro-bust megérthessük, hogy a kereskedési algoritmusok hogyan lépnek interakcióba", noha azt, hogy egy algoritmus "anyagilag hogyan cselekszik, az interakció alakítja", így az algoritmusokat "relációsan kell megérteni" (MacKenzie 2019a: 55). Az "automatizált kereskedés gép-gép ökológiája" (Hayles 2017: 175) kikerül az emberek megértése és ellenőrzése alól, ahogy végső soron a mesterségesen intelligens ágensok megértése és ellenőrzése alól is. Elképzelhető talán egy gyümölcsöző, szimbiózisra épülő kölcsönhatás az "intelligens" kereskedési algoritmusok között, és a piactervezés területén belül minden bizonnyal még mindig él az a törekvés és remény, hogy a több ágensből álló mesterséges intelligencia-rendszereket, beleértve azok interakciós szabályait is, a semmiből tervezzék meg, és egyfajta *machina economicust* hozzanak létre (Parkes/ Wellmann 2015: 272). A piactervezés ellenére azonban az AI kereskedési algoritmusok nagyrészt önállóan működnek, és ebből a szempontból a pénzügyi piacok sem alkotnak "kollektív kapitalista agyat" (Terranova 2013: 66), hanem az okos ágensok egymással versengő kis kapitalista agyak összességét alkotják.

Ezt a helyzetet súlyosbítja a többszörös homály, amely a pervasive a pénzügyi piacokon. Burrell szerint a gépi tanulási algoritmusok egyes homályosságai áthidalhatatlanok, és a gépi tanulás működésének alapvető részét képezik az architektúrák és a skálák tekintetében (Burrell 2016: 4-5). Az olyan stratégiák, mint a megmagyarázható mesterséges intelligencia, jelenleg szintén nem eredményezik az opacitások csökkentését (Sudmann 2018: 187-191). A mesterséges intelligencia ezen opacitásai azonban csak a legújabb adalékok a pénzügyi piacok egyéb opacitásaihoz, és ezeket súlyosbítják az algoritmikus kereskedők fentebb már említett titokzatos stratégiái. Már említettem, hogy a nagyfrekvenciás kereskedők hogyan használják ki a pénzügyi piacok vízvezetékrendszerét és információs aszimmetriáit. Mivel ezek versenyelőnyt jelentenek, a lehető legteljesebb mértékben titokban tartják őket; csak az olyan felfedezések, mint a fent említett Bodek vagy Michael Lewis *Flash Boys* (2014) című könyvében tett felfedezések vezettek ahhoz, hogy a pénzügyi piacok mikrostruktúrája nyilvánosan is ismertté vált. Léteznek továbbá a dark

poolok (MacKenzie 2019b), amelyek - ahogy nevük is sugallja - nagyrészt a sötétben működnek, a megbízási könyvek és platformjaik számos más jellemzője nagyrészt hozzáférhetetlen.

a nyilvánosság számára. Lange (2016) arról is beszámol, hogy a HFT propshopok felállítása egyfajta szervezeti tudatlanságot eredményez, ahol a kereskedők és a kódolók között olyan korlátokat állítanak fel, amelyek az utánpótlás elkerülését szolgálják, de káros mellékhatásokhoz is vezethetnek.

Az algoritmikus kereskedelem homályosságának és fekete dobozainak politikai kikezdése ezért komoly felfordulást igényelne a pénzügyi piacokon. A szabályozási kísérletek csak korlátozottan foglalkoztak ezekkel az átláthatatlanságokkal, például a HFT-algoritmusok azonosíthatóságának előírásával (pl. Coombs 2016). Más kísérletek is léteznek a piacok kialakításának megváltoztatására az átláthatatlanságok csökkentése érdekében. Az Investors Exchange (IEX) például egy olyan kereskedési platform, amelyet Lewis (2014) a HFT elleni küzdelemként ünnepel: egy 61 km hosszú kábel tekeresése az adatközpont körül mintegy 7 milliszekundumot ad az algoritmusok "oda-vissza útjához", és gyakorlatilag kizárja a HFT működőképességét a platformon. Az IEX emellett sokkal átfogóbb díjstruktúrával rendelkezik, és "tisztességesebb" kereskedési feltételeket kínál. Egy másik javaslat Budish et al. (2015) tollából származik, akik azt javasolják, hogy a folyamatos limitáras megbízási könyveket - a kereskedési platformok jelenlegi módját a megbízások egyeztetésének megszervezésére - váltsák fel kötegelte aukciókkal, amelyek másodpercenként zajlanának, és ezáltal a nagyfrekvenciás kereskedők időbeli előnyeit is nagymértékben megtagadnák (lásd még Hayles 2017: 165-169). Roth, a piactervezés egyik fő támogatója és Budish tanára támogatja ezeket a javaslatokat (2015: 81-100). Mirowski és Nik-Khah (2008) más kontextusban óva int attól, hogy a piactervezésnek ezt a neoliberális árnyalatú konstruktivista perspektíváját átvegyük. Bár az algoritmikus kereskedés politikájáról más árnyalt megfontolások is léteznek (lásd pl. Lange et al. 2016), e javaslatok egyike sem foglalkozik az algoritmikus kereskedés mesterséges intelligenciájának homályosságával.

Az sem lenne világos, hogy ezek a változások milyen mértékben vezetnének demok-

az algoritmikus kereskedés és a mesterséges intelligencia ratifikációja. Ahogy MacKenzie és Pardo-Guerra ref előadást tart az Island kapcsán, amelynek megbízási könyve nyitott volt, "lehetővé téve bárki számára a valós idejű betekintést a megbízási könyvébe", ellentétben az összes jelenlegi kereskedési helyszínnel: "az információ talán szabad akart lenni, de a kapitalizmusnak más prioritásai voltak" (2014: 171). Különösen a fentebb tárgyalt, a pénzügyi piacok bővülő adatökológiai körüli fejlemények azt sugallják, hogy az algoritmikus kereskedés és a mesterséges intelligencia fejlődése demokratizálódás helyett inkább a mindennapi élet további finanszírozásához vezet (Martin 2002). A közösségi médiában és máshol rögzített társadalmi élet ma már "pénzügyi társadalmi gépekbe táplálkozhat, amelyek integrálják az innovatív nagysebességű hálózatot, a közösségi médiából származó információkat és az egyének kereskedési döntéseit, hogy pontosabb árelőrejelzéseket adjanak, ami a pénzügyi piacok jobb integrációjához vezet" (Ma/ McGroarty 2017: 245). Itt a mélytanulás

"nagy ígérete", amely "nem csupán arra irányul, hogy a gépek megértsék a világot, hanem arra is, hogy azt egyre többféleképpen kiszámíthatóvá tegyék: hogyan alakul a tőzsde, mit akarnak vásárolni az emberek, meghal-e egy ember vagy sem, és így tovább" (Sudmann 2018: 193), beíródik abba, amit Hayles úgy hív, hogy

"vámpiros kapitalizmus" (2017: 159), illetve amit fentebb a pénzügyi és kognitív kapitalizmus kapcsán tárgyaltam.

A "modern AI-technológia infra-mediális feltételeire és azok politikai dimenziójára" (Sudmann 2018: 185) való összpontosítás, ahogyan azok a pénzügyi piacokkal kapcsolatban mutatkoznak, valamint a pénzügyi piacok "elemzési fókuszunknak a pénzügyi piacok infrastruktúrái felé történő áthelyezése" (Pardo-Guerra 2019: 31), ahogyan azt ez a hozzájárulás megkísérli, feltárja, hogy az algoritmikus kereskedés és az AI újabb alkalmazása ennek részeként milyen alaposan magával ragadja a pénzügyi és kognitív kapitalizmust. Ahhoz, hogy az algoritmikus kereskedésben a mesterséges intelligencia politikáját megragadjuk, elemezni kell, hogy a mesterséges intelligenciát hogyan vetik be az érték kivonás szolgálatában, legújabbban a társadalmi életből, ahogy azt a közösségi médiában és máshol rögzítik. A fenti vázlat azt mutatja, hogy a mesterséges intelligencia politikája egyre szorosabban összefonódik a pénzügyekkel és azokkal a kognitív ökológiákkal, amelyekben működik. A műveletek politikájának kibővített megértésének részeként (Mezzadra és Neilson 2019) a pénzügyek részeként alkalmazott mesterséges intelligencia megmutatja, hogy a finanszírozás révén hogyan vesz részt az értékek kinyerésében, amelyből való kitéréshez többre lenne szükség, mint a piaci tervezés néhány apró módosítására. A legközvetlenebb módon az AI politikája a pénzügyi piacokon a megismerés politikájaként jelenik meg, amelyben jelenleg a "nem tudatos kognitív fegyverkezési verseny" (Hayles 2017: 165) határozottan eltolódik egy olyan terep felé, ahol az AI a neoliberais pénzügyi tőkével cinkos. Ez a megismerés olyan politikáját igényli, amely végiggondolja, hogyan lehetne a mesterséges intelligenciát kivonni ebből a cinkosságból, és más célokra felhasználni, amelyek nem feltétlenül egyeznek meg a pénzügyi és kognitív kapitalizmussal.

Köszönetnyilvánítás

Köszönöm Ann-Christina Lange-nak, hogy olyan alaposan bevezetett a nagyfrekvenciás kereskedés jelenségébe, és az ezt követő közös írói projektekért. Köszönet a bochumi Center for Advanced Internet Studies (CAIS) kutatóközpontnak, ahol 2019 nyarán ösztöndíjas voltam, amikor ezt a szöveget elkészítettem.

Hivatkozások

- "Alt Data Tips for Traders | Dataminr". n.d. (<https://www.dataminr.com/resources/tips-for-traders-to-take-advantage-of-alt-data>). Hozzáférés: 2019. július 11.
- Amoore, Louise (2019): "Bevezetés: Algoritmussal gondolkodva: Thinking with Algorithms: Katherine Hayles munkásságában". In: Theory, Culture &

Societaty 36/2, pp. 3-16.

Arévalo, Andrés/Niño, Jaime/Hernández, German/Sandoval, Javier (2016): "Mély neurális hálózatokon alapuló nagyfrekvenciás kereskedési stratégia". In: De .

- Shuang Huang/Kyungsook Han/Abir Hussain (szerk.), *Intelligent Computing Methodologies*, Cham: Springer International Publishing, pp. 424-36.
- Arnoldi, Jakob (2016): "Computer Algorithms, Market Manipulation and the Institutionalization of High Frequency Trading". In: *Theory, Culture & Society* 33/1, pp. 29-52.
- "Artificial Intelligence(AI) and the Technology behind EquBot". n.d. (<https://equbot.com/technology/>). Hozzáférés július 11. 2019.
- Beunza, Daniel/Stark, David (2008): "A szakma eszközei: Az arbitrázs társadalmi-technológiai háttere egy Wall Street-i kereskedőteremben". In: Trevor J. Pinch/Richard Swedberg (szerk.), *Living in a Material World: Economic Sociology Meets Science and Technology Studies*, Cambridge, MA: MIT Press, pp. 253-90.
- Beverungen, Armin (2018): "Kognitiver Kapitalismus? Nichtbewusste Kognition Und Massenintellektualität.". In: *Zeitschrift Für Medienwissenschaft* 18, pp. 37-49.
- Beverungen, Armin (2017): "Zeitlichkeit und Kognition im Hochfrequenzhandel". In: *Archiv für Mediengeschichte* 17, pp. 9-20.
- Beverungen, Armin/Lange, Ann-Christina (2018): "Kogníció a nagyfrekvenciás kereskedésben: The Costs of Consciousness and the Limits of Automation". In: *Theory, Culture & Society* 35/6, pp. 75-95.
- Borch, Christian (2016): "High-Frequency Trading, Algorithmic Finance and the Flash Crash: Reflections on Eventalization." In: *Gazdaság és társadalom* 45/3-4, pp. 350-78.
- Budish, Eric/Cramton, Peter/Shim, John (2015): "The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response". In: *The Quarterly Journal of Economics* 130/4, pp. 1547-1621.
- Burrell, Jenna (2016): "Hogyan "gondolkodik" a gép: Understanding Opacity in Machine Learning Algorithms." In: *Big Data & Society* 3/1, pp. 1-12.
- Burton, Catherine (2016): Burton: "Inside a Moneymaking Machine Like No Other". In: *Bloomberg Markets*, 2016. november 21. (<https://www.bloomberg.com/news/articles/2016-11-21/how-renaissance-s-medallion-fund-became-finance-s-blackst-box>). Hozzáférés: 2019. július 11.
- Coombs, Nathan (2016): "Mi az algoritmus? Pénzügyi szabályozás a nagyfrekvenciás kereskedés korában.". In: *Gazdaság és társadalom* 45/2, pp. 278-302.
- Curry, Derek (2018): Curry Curry: "Public Dissentiment: Hacktivism in the Age of High-Frequency Traders". In: *Visual Resources* 34/1-2, pp. 93-115.
- De Aenlle, Conrad (2018): "A.I. megérkezett a befektetésbe. Az emberek még mindig uralkodnak". In: *A.I.A: The New York Times*, június 8., sz. Business (<https://www.nytimes.com/2018/01/12/business/ai-investing-humans-dominating.html>). Hozzáférés: 2019. július 11.

- Deng, Li (2018): "AI in Finance." Előadás az EmTech Digitalon, San Francisco, március (2018) <https://events.technologyreview.com/video/watch/li-deng-ai-finance/>). Hozzáférés július 11 2019.
- Dyer-Witheford, Nick (2016): "Cybernetics and the Making of a Global Proletariat". In: *The Political Economy of Communication* 4/1, pp. 35-65.
- Frontiers, A.I. (2018): "Utazás a mesterséges intelligenciával: a beszéd-től a pénzügyekig". 201824. augusztus (<https://medium.com/aifrontiers/a-journey-with-ai-from-speech-to-finance-2a60cb3422a>). Hozzáférés: 2019. július 11.
- Fumagalli, Andrea/Mezzadra, Sandro (szerk.) (2010): *Válság a globális gazdaságban*: Los Angeles/ Cambridge, MA: Semiotext(e)/Distributed by the MIT Press.
- Ganesh, Prakar/Rakheja, Puneet (2018): "Deep Neural Networks in High Frequency Trading". In: ArXiv:1809.01506 [Cs, q-Fin, Stat], szeptember (<http://arxiv.org/abs/1809.01506>).
- Hayles, N. Katherine (2017): *Gondolatlan: The Power of the Cognitive Nonconscious*, Chicago/London: The University of Chicago Press.
- Johnson, Neil/Zhao, Guannan/Hunsader, Eric/Qi, Hong/Johnson, Nicholas/ Meng, Jing/Tivnan, Brian (2013): "Abrupt Rise of New Machine Ecology Beyond Human Response Time" (Az új gépi ökológia hirtelen felemelkedése az emberi reakcióidőn túl). In: *Scientific Reports* 3/September (<https://doi.org/10.1038/srep02627>).
- Karppi, Tero/Crawford, Kate (2016): In: "Social Media, Financial Algorithms and the Hack Crash": *Theory, Culture & Society* 33/1, pp. 73-92.
- Kearns, Michael/Nevmyvaka, Yuriy (2013): "Machine Learning for Market Micro-structure and High Frequency Trading". In: David Easley/Marcos Lopez de Prado/ Maureen O'Hara (szerk.), *High Frequency Trading-New Realities for Traders, Markets and Regulators*, London: Risk Books.
- Kirilenko, Andrei A./Lo, Andrew W. (2013): "Moore törvénye kontra Murphy-törvénye: Algorithmic Trading and Its Discontents (Algoritmikus kereskedés és annak ellenérzései)". In: *Journal of Economic Perspectives* 27/2, pp. 51-72.
- Lange, Ann-Christina (2016): "Szervezeti tudatlanság: An Ethnographic Study of High-Frequency Trading". In: *Gazdaság és Társadalom* 45/2, pp. 230-50.
- Lange, Ann-Christina/Lenglet, Marc/Seyfert, Robert (2016): "Cultures of High-Frequency Trading: Mapping the Landscape of Algorithmic Developments in Contemporary Financial Markets (Az algoritmikus fejlesztések tájképének feltérképezése a kortárs pénzügyi piacokon)". In: *Economy and Society* 45/2, pp. 149-65.
- Lewis, Michael (2014): *Flash Boys: A Wall Street Revolt*, New York: W.W. Norton & Company.
- MacKenzie, Donald (2006): *An Engine, Not a Camera: How Financial Models Shape Markets*, Cambridge, MA: MIT Press.
- MacKenzie, Donald (2017): "A Material Political Economy: Automated Trading

Desk and Price Prediction in High-Frequency Trading". In: *Social Studies of Science* 47/2, pp. 172-94.

- MacKenzie, Donald (2019a): "How Algorithms Interact: Goffman's Interaction Order' in Automated Trading". In: *Theory, Culture & Society* 36/2, pp. 39-59.
- MacKenzie, Donald (2019b): "Market Devices and Structural Dependency: The Origins and Development of 'Dark Pools' ". In: *Finance and Society* 5/1, pp. 1-19.
- MacKenzie, Donald/Beunza, Daniel/Millo, Yuval/Pardo-Guerra, Juan Pablo (2012): "Fúrás az Allegheny-hegységen keresztül: Liquidity, Materiality and High-Frequency Trading (Likviditás, lényegesség és nagyfrekvenciás kereskedés)". In: *Journal of Cultural Economy* 5/3, pp. 279-96.
- MacKenzie, Donald/Pardo-Guerra, Juan Pablo (2014): "Insurgent Capitalism: Island, Bricolage and the Re-Making of Finance". In: *Gazdaság és társadalom* 43/2, pp. 153-82.
- Markoff, John (2018): "A nanoszekundumra osztott idő pontosan az, amit a Wall Street akar". *The New York Times*, 2018. június 30., sec. Technology (<https://www.nytimes.com/2018/06/29/technology/computer-networks-speed-nasdaq.html>). Hozzáférés: 2019. július 11.
- Martin, Randy (2002): *Philadelphia: A mindennapi élet finanszírozása*: Temple University Press: Philadelphia: Temple University Press.
- McCauley, Dennis (2016): *McCauley McCauley: Ghosts in the machine: London: Artificial intelligence, risks and regulation in financial markets, London: Thought Leadership Consulting (www.euromoneythoughtleadership.com/ghostsinthemachine)*.
- Metz, Cade (2016): "A mesterségesen intelligens fedezeti alapok felemelkedése". *Wired*, January (25,2016<https://www.wired.com/2016/01/the-rise-of-the-artificially-intel-ligent-hedge-fund/>). Hozzáférés július 11 2019.
- Mezzadra, Sandro/Neilson, Brett (2019): *The Politics of Operations: Excavating Contemporary Capitalism*, Durham: Duke University Press.
- Mirowski, Philip (2007): "Markets Come to Bits: Evolution, Computation and Markomata in Economic Science". In: *Journal of Economic Behavior & Organization* 63/2, pp. 209-42.
- Mirowski, Philip/Nik-Khah, Edward (2008): "Parancsnoki teljesítmény: Exploring What STS Thinks It Takes to Build a Market." In: Trebor Pinch/Richard Swed-berg (szerk.), *Living in a Material World: Economic Sociology Meets Science and Technology Studies*, Cambridge, MA: The MIT Press, pp. 89-128.
- Mirowski, Philip/Nik-Khah, Edward (2017): *The Knowledge We Have Lost information: The History of Information in Modern Economics*, New York: Oxford University Press.
- Munn, Luke (2018): *Lüneburg*: Meson press.
- Nik-Khah, Edward/Mirowski, Philip (2019): "The Ghosts of Hayek in Orthodox Microeconomics: Marekts as Information Processors." In: Armin Beverungen/Jens Schröter/Philip Mirowski/Edward Nik-Khah, *Markets*, Minneapolis/Lüneburg: University of Minnesota Press/meson press.

- Parkes, David C./Wellman, Michael P. (2015): "Gazdasági gondolkodás és mesterséges intelligencia". In: *Science* 349/6245, pp. 267-272.
- Patterson, Scott (2010): *The Quants: New York, New York: How a New Breed of Math Whizzes Conquered Wall Street and Nearly Destroyed It*, New York, NY: Crown Business.
- Patterson, Scott (2013): *Sötét medencék. The Rise of A.I. Trading Machines and the Looming Threat to Wall Street*, London: Random House Business.
- "Jelentés: Dataminr", n.d. (<https://www.data-minr.com/resources/report-investors-embrace-alternative-data>). Hozzáférés: 11. július 2019.
- Roth, Alvin E. (2015): *Ki mit kap - és miért: The New Economics of Match-making and Market Design*, Boston, MA: Houghton Mifflin Harcourt.
- Satariano, Adam/Kumar, Nishant (2017): "Ez a hatalmas fedezeti alap a mesterséges intelligenciára fogad". In: *Bloomberg Markets*, 2017. szeptember 27. (<https://www.bloomberg.com/news/features/2017-09-27/the-massive-hedge-fund-betting-on-ai>). Hozzáférés: 2019. július 11.
- Sudmann, Andreas (2018): Sudmann: "A mesterséges intelligencia médiapolitikai dimenziójáról". In: *Digital Culture & Society* 4/1, pp. 181-200.
- Terranova, Tiziana (2013): Tiziana Tiziano: Tiziano Tiziano: "A kognitív kapitalizmus hétköznapi pszichopatológiái". In: Arne De Boever/Warren Neidich (szerk.), *A kognitív kapitalizmus pszichopatológiái: Első rész*, Berlin: Archive Books, pp. 45-68.
- Toscano, Alberto (2013): "Gaming the Plumbing: High-Frequency Trading and the Spaces of Capital". In: *Mute Magazine* 3/4, pp. 74-85.
- Vardi, Nathan (2016): "Wall Street Keeps Raiding Silicon Valley for Tech Talent", *Forbes* március (11,2016) (<https://www.forbes.com/sites/nathanvardi/2016/03/11/wall-street-keeps-raiding-silicon-valley-for-tech-talent/>). Hozzáférés: 2019. július 11.
- Vogl, Joseph (2015): *The Specter of Capital*, Stanford, CA: Stanford University Press.
- Zook, Matthew/Grote, Michael H. (2017): "The Microgeographies of Global Finance: High-Frequency Trading and the Construction of Information Inequality." In: *Környezet és tervezés* A 49/1, pp. 121-40.

A használható adatok keresése Gépi tanulási algoritmusok építése a közzsféra archívumaiból

Lisa Reutter/Hendrik Storstein Spilker

Ez a fejezet a norvég kormány egyik korai törekvését elemzi, amely a közszolgáltatások javítására irányult a közzsféra archívumaiból származó adatok felhasználásával. A munkaügyi és jóléti adminisztráció (NAV) mesterséges intelligencia-alapú szolgáltatások fejlesztésére irányuló kezdeményezését vizsgálja. A norvég közzsféra úttörő hangulatban van. A digitalizáció új hulláma a platformok, a felhők és az algoritmusok felé irányítja a figyelmet. A mesterséges intelligencia magában hordozza a közzsféra forradalmasításának lehetőségét és ígéretét. Különösen a felügyelt gépi tanulás vált a választott módszerre az adatvezérelt válás végső és valahogyan diffúz céljának eléréséhez.¹ Nagy az izgalom azzal kapcsolatban, hogy a gépi tanulási algoritmusokat hogyan lehet felhasználni a jobb és személyre szabottabb szolgáltatások nyújtására, a bürokrácia megváltoztatására és a polgárok felhatalmázására. Az állampolgárokra vonatkozó információk rögzítése, tárolása és feldolgozása már régóta a modern állam kulcsfontosságú eleme; az ehhez szükséges számítási rendszerek és technikák azonban egyre gyorsabbá, átfogóbbá és önállóbbá váltak (Beer 2017).

A technológiai magánvállalkozásokkal szemben a közzsféra szervezetei egy nyilvánvaló előnnyel rendelkeznek - legalábbis "papíron". A polgárokról hatalmas, személyes jellegű, gyakran hosszú időn keresztül rögzített és folyamatosan frissített adathalmazokkal rendelkeznek. Ahogy Redden megjegyzi, "ez adatelemzési szempontból hihetetlenül értékessé teszi őket" (2018:1). Informátoraink nagyon is tisztában vannak ezzel a potenciális előnnyel - egyesek a nagy kormányzati adatokra úgy hivatkoznak, mint "a mi aranyunkra". Az aranyat gazdagnak, átfogónak, izgalmasnak és egyedülállónak írják le a bányászok. A gépi tanulás lehetőségként mutatkozik be az archívumokban rejlő arany kibányászására, új és meglepő betekintést nyújtva az ügyintézőknek saját munkájukról és az általuk irányított polgárokról.

Azonban, mint a valós bányászatban, az arany kitermelése az ércekből nem feltétlenül egyszerű. Valakinek ki kell ásni, meg kell különböztetni a többi aranytól.

¹ A mesterséges neurális hálózatokkal (ANN) kapcsolatos technikák alkalmazása nem megengedett a következőkben közszolgálat, mivel nem átlátható, és a döntések nem magyarázhatók meg.

tételek, mossuk és tisztítsuk meg, hogy alkalmassá tegyük a közjavak előállítására. A NAV-ban erre a kérdésre egy új adattudományi környezet kialakítása volt a válasz. Ez a fejezet a váratlan kihívásokról szól, amelyekkel az AI-osztálynak szembe kellett néznie - és arról a hétköznapi munkáról, amely a gépi tanulás gyakorlatának háttérében áll. Kutatási kérdésünk tehát kettős: *Milyen kihívások kapcsolódnak a közzsféra archívumaiból származó AI-alapú szolgáltatások fejlesztéséhez? Hogyan tükrözik ezek a korai kihívások azokat a bizonytalanságokat, amelyek az AI-nek a közzszolgáltatásban való megjelenésével kapcsolatos hype mögött rejlenek?*

Ezek fontos szempontok, mivel a felelősség, hogy a kormány és a polgárok közötti kapcsolatokban a mesterséges intelligencia állítólagos felhatalmazó és demokratizáló potenciálját megvalósítsuk, végső soron azokon múlik, akik az adatokat előkészítik és az algoritmusokkal bütykölnek. A közzszférán belül eddig figyelemre méltó mennyiségű optimizmus és hype volt tapasztalható az AI-alapú szolgáltatások fejlesztésével kapcsolatban (Vivento AS/Kaupán AS 2015; Technologirádet 2017). Ugyanakkor egyre inkább tudatosulnak azok az aggályok, amelyek a mesterséges intelligenciáról szóló társadalomtudományi diskurzus nagy részét uralják. Mindennapi életünk egyre több aspektusát érinti az adatifikáció, amikor az emberi tevékenységet és viselkedést digitális adatok elemezhető formájává alakítják át és többféleképpen hasznosítják (Mayer-Schönberg/Cukier 2013). A big data hasznosítása komoly adatvédelmi, adatbiztonsági és etikai kérdéseket vet fel. Ezek a kérdések természetesen még kritikusabbak, ha a mesterséges intelligenciát a magánszektorhoz képest a közzszférában alkalmazzák (vö. Sudmann 2018). Jelentős a megfigyelés lehetősége, valamint az igazságtalan gyakorlatok automatizálásának kockázata (vö. Pasquale 2015; Cheney-Lippold 2017; Crawford/boyd 2012).

Természetesen ezek az aggályok arra is ösztönzőleg hatnak, hogy a kutatás során vizsgálják és mélyebben megértsék azokat a folyamatokat, amelyek során a (hagyományos) közzszféra archívumai (modern) gépi tanulási algoritmusokká alakulnak át. A demokratikus befolyás és ellenőrzés lehetővé tétele és biztosítása érdekében nemcsak a kész algoritmusok hatását kell tanulmányozni, hanem a megalkotott algoritmusokat is (Latour 1987-et idézve). Elméleti szempontból a "kritikai algoritmustanulmányok" új területén végzett munkákból tájékozódunk (Beer 2017; Kitchin 2017; Gillespie 2014). Az algoritmustanulmányok a digitális tartalmak és interakciók tanulmányozásán túlmutató elmozdulást jelentenek a digitális tartalmak láthatóságát és az interakciók mintáit kondicionáló infrastruktúrák vizsgálata felé. A kritikai algoritmustanulmányok központi feladata az algoritmusalapú infrastruktúrák struktúráinak, dinamikájának és következményeinek feltárása, mivel ezek az infrastruktúrák gyakran technikai és semleges, átláthatatlan és áthatolhatatlan infrastruktúrákként jelennek meg (Burrell 2016).

Mivel azonban az algoritmusalapú infrastruktúrák egyre több döntés és ajánlás alapját képezik a társadalmi, politikai és gazdasági területeken, sürgősen foglalkozni kell szerepükkel és működésükkel. Pasquale (2015)

gyakran használta a "fekete doboz" metaforáját annak megjelölésére, hogy a létfontosságú társadalmi döntések hogyan alakulnak ki a láthatóságon és ellenőrzésen kívül. Pasquale felvázol egy forgatókönyvet

egy *belső*, technológiai cégekből, adattudósokból és titkos és átláthatatlan algoritmusaikból álló, hatalmat és ellenőrzést gyakorló részleggel, és egy jogfosztott külsővel, ahol mi többiek, polgárok, vásárlók, az egész régi társadalom lakunk.

A kritikus algoritmus-tanulmányok értékes betekintést nyújtottak az adatszerkesztési gyakorlatok mögött vagy alatt álló szereplők és szervezetek működésébe, valamint abba, hogy hogyan járulnak hozzá a társadalmi rendezéshez. Flyverbom és Murray szerint azonban eddig "kevés mondanivalójuk volt azokról a tényleges, *belső* folyamatokról, amelyek révén az adatok megszerveződnek és strukturálódnak" (2018: 5-6). Boyd és Elish is kiemeli az adatok gyűjtésének, tisztításának és kuratírozásának hétköznapi munkájának importanciáját, mert "ezen a hétköznapi munkán keresztül [az] kulturális értékek beágyazódnak a rendszerekbe" (2018: 69). Annak ellenére, hogy többször felszólítottak több etnográfiai vizsgálatra, eddig kevés készült (Kitchin 2017). Így a NAV adattudományi környezetének "laboratóriumi vizsgálatára" vonatkozó döntésünk egyik fontos motivációja az ilyen tanulmányok hiányának felismerésén és azon a vágyon alapult, hogy megvizsgáljuk az algoritmus-konstrukció folyamatainak apró részleteit. A végső cél az volt, hogy *ex amine* megvizsgáljuk a gépi tanulás tényleges gyakorlatát, valamint azokat a bizonytalanságokat és módszertani kihívásokat, amelyek a mesterséges intelligencia közszolgálati hype mögött húzódnak (boyd/Elish 2018).

Esettanulmány: Munkaügyi és Jóléti Igazgatóság

A NAV, az egyik legnagyobb norvég közhivatal, a folyamatban lévő országos digitális átalakulás élén áll. A NAV egy közjóléti ügynökség, amely több mint 60 különböző ellátást és szolgáltatást nyújt, például munkanélküli segélyt és nyugdíjat. Az állami ügynökség a teljes norvég állami költségvetés mintegy harmadát kezeli, és a munkaügyi minisztérium alá tartozik. A NAV körülbelül 19 000 alkalmazott, akik közül körülbelül 14 000 a központi kormányzat alkalmazásában áll, további 5000 pedig helyi szinten.

A NAV adattudományi környezete az informatikai osztályon belül egy újonnan létrehozott részleg része. Ez a részleg a Munkaügyi és Jóléti Igazgatóságon belül az adattermékeket fejlesztő és kezelő összes környezettel kíván foglalkozni. Feladata tehát az állampolgárok adatosításának megszervezése. Az adattudományi környezetet a következő évben hozták létre, és 2017 akkoriban néhány adattudós és egy csoportvezető megfigyeléséből állt. E csapat tagjai az elképzelt adatvezérelt közhivatal kulcselemei.

Az adatvezérelt válás iránti késztetés a szervezeten belül és kívül egyaránt megjelenik. A szervezeten belül az egyének már egy ideje kísérleteznek a nagy adatokkal. A szervezeten kívül a társadalmi és gazdasági

tendenciák, mint például az olajszektor visszaesése, a magasabb bevándorlási arányok és az iparágak automatizálódása új kihívások elé állítják az adminisztrációt és a jólétet.

állam általánosságban. A javasolt megoldás? Egy adatvezérelt jóléti állam. A politikai vezetők ezért a gépi tanulás és a nagyméretű adatok vizsgálatát kérték:

Természetes feltételezés, hogy a big data az olyan technológiákkal együtt, mint az automatizálás és a mesterséges intelligencia, képes lesz megváltoztatni a kormányzat szolgáltatástermelésének jövőbeli működését (Kommunal- og moderniseringsdepartementet 2016: 109).

Az adatvezérelt digitális átalakulás ezen első szakaszában a gépi tanulás algoritmusait elsősorban döntéstámogató eszközként fejlesztik ki. Ezt például egy olyan projekten keresztül lehet szemléltetni, amely az önkormányzati és kormányzati adatokat kívánja egyesíteni a felhasználók nyomon követésének javítása érdekében. A projekt egyik ambíciója az új munkanélküliségi esetek sebezhetőségének azonosítása. A tervezett végtermék egy osztályozó eszköz, amely az újonnan munkanélküli polgárokat két csoportba sorolja: azokba, akiknek valószínűleg intenzív nyomon követésre van szükségük a NAV-tól, és azokba, akik valószínűleg rövid időn belül, kevés beavatkozással elhelyezkednek. Ezt az értékelést korábban a humán felhasználói támogatás végezte el.

A felhasználó szükségleteinek első értékelését a lehető legnagyobb mértékben automatizáltan kell elvégezni, és annak ismeretén kell alapulnia, hogy mely tényezők befolyásolják a felhasználó munkaerőpiacra való belépési lehetőségeit. (NAV-ekspertgruppen 2015: 13)

A terepmunka januárban zajlott, és 2018magában foglalta az adattudományi környezet háromhetes megfigyelését, 11 mélyinterjút a csapaton belüli és kívüli kulcsfontosságú alkalmazottakkal, valamint a belső dokumentumok elemzését, amelyek a nagy adatok gépi tanulással történő felhasználásával kapcsolatos munkát vitatják meg és mutatják be.

A nyilvános archívumok aranybányájának bányászata: A használható adatok keresése

A modern állam és az adatok elválaszthatatlanul összefonódnak, mivel a statisztikai információk nyilvánosság számára való hozzáférhetősége minden demokrácia feltétele és szükségszerűsége (Desrosières 1998: 324). A kormányozandó alanyokra vonatkozó digitális adatok mennyisége, szemcsézettsége, közvetlensége és sokfélesége egyedülálló a kortárs kormányok számára (Ruppert/Isin/Bigo 2017). A NAV a második legnagyobb adattermelő a norvég közszférában. Az adatok mindig is fontos szerepet játszottak a közigazgatásban, mivel hivatalos statisztikákat és jelentéseket készít a politikai döntéshozatalhoz, például a betegszabadságról és a munkanélküliségről.

A Munkaügyi és Jóléti Igazgatóság a polgárokról és saját munkájukról szóló hatalmas mennyiségű információ archiválásának, összegyűjtésének és tárolásának kultúráját gyakorolja.

Meglepő módon a kormányzati szervek hajlamosak elfelejteni a birtokukban lévő adatokat, hacsak egy válság vagy vizsgálat nem vezet arra, hogy foglalkozzanak az elfelejtett vagy rosszul tárolt adatokkal, vagy a pontokkal, amelyeket nem sikerült összekapcsolniuk (Prince 2017: 236). Az adatokat eddig a statisztikák előállításához használták fel, majd egy nyilvános archívumba vagy adatbázisba vitték át. Az archívum a gépi tanulás megjelenésével megváltozik a szervezeten belüli szerepe - az adatok passzív befogadójából és gyűjtőjéből aktív adatszolgáltatóvá válik. Ahelyett, hogy porosodnak, az adatok a közigazgatás mindennapi munkájának irányítását szolgálják. A közigazgatás egy eddig fel nem fedezett értéket feltételez a közlevéltárakban, amely kulcsfontosságú lehet a közigazgatás túlélése szempontjából. Az archívum így az érték és a hatalom forrásává válik. A benne tárolt információ aktív célpontjává válik a feltárásnak.

Az aranybányászat azonban piszkos üzlet. Az olyan vállalatok, mint például a Google/Alphabet, a Facebook és az Amazon, úgy tűnik, könnyedén visszatáplálják az adatokat a gyakorlatba, és bányásznak aranyat, miközben létrehozzák. Ezzel szemben a közszférában a ma- chine learning algoritmusok létrehozása a már meglévő adatokra és infrastruktúrákra támaszkodhat és támaszkodnia is kell. Ezen túlmenően a régóta létező gyakorlatokhoz és értékrendekhez kell igazodnia. A hatalmas állami archívumok felbecsülhetetlen és korlátlan adatforrást ígérnek a gépi tanulási algoritmusok létrehozásához. A gyakorlatban azonban a felhasználásukkal kapcsolatos kihívások széles skálája létezik.

A Munkaügyi és Jóléti Igazgatóságnak a már meglévő digitális infrastruktúrára adatfelhasználási infrastruktúrát kell építenie, ami egyszerre korlátozza és teszi lehetővé a gépi tanúlással kapcsolatos munkát. Az, hogy milyen adatokat és hogyan használnak fel, befolyásolja az algoritmusok által készített előrejelzéseket. A gépi tanulási algoritmusok előállításához nagy mennyiségű adatra van szükség, amelyek alapján az algoritmusokat finomítani és tesztelni lehet. Egyik informátorunk úgy foglalja össze az adatmunka általános fontosságát, hogy úgy jellemzi azt, mint a közigazgatás adatvezérelt jövőjének alapját, amelyen a kezdeményezések sikere vagy kudarca múlik:

Tehát, ha tudjuk, milyen adatokkal rendelkezünk, milyen minőségűek az adatok, mire használhatjuk fel őket, akkor szerintem számolnunk kell azzal, hogy sok időt kell erre fordítanunk. Szerintem ez lesz az alap. És amit erre építesz, az sem lesz jobb, mint az alap.²

Az adattudományi környezetben végzett munka nagy része messze nem korlátozódik az adatelemzés és az informatika gyakorlatára. Mielőtt bármilyen algo-

2 Az adminisztrációval kötött közzétételi megállapodás miatt az informátorok egyikét sem azonosítják semmilyen metainformációval vagy álnévvel. Minden jelöletlen idézet tehát a 11 interjú bármelyikéből származik. Bár ez veszélyezteti az elemzés átláthatóságát, a megfigyelés során a csapat mérete miatt erre szükség volt.

rithmus felépíthető, az adattudósoknak maguknak kell összeállítaniuk az adatokat, amelyeket algoritmusokhoz lehet táplálni. A csapatnak tárgyalnia kell a képzési és tesztadatokhoz való hozzáférésről, meg kell értenie az adatok etikus felhasználását támogató jogi kereteket, és értékelnie kell az adatok minőségét. A mesterséges intelligenciával foglalkozó munkatársaknak gépi tanulhatóvá kell tenniük az adatokat. Ez bizonyos fokú frusztrációhoz és bizonytalansághoz vezet az adattudósok körében, amit azonban szükségesnek tartanak a gépi tanulás megfelelő használatának és előállításának biztosításához. Nézzük tehát közelebbről az adatok összeállításának folyamatait, az adatok megszerezését és strukturálását a gyakorlatban.

Hozzáférés

A nyilvános archívum szerepének általános változása megköveteli, hogy az adattudósok aktívan részt vegyenek a leporolt archívumban, így hozzáférve annak belső működéséhez. A közhivatal szabványosított és jó rutinokkal rendelkezik a nyilvános statisztikákban használt felhalmozott adatokra vonatkozóan. Az adatok elérhetők és megtalálhatók egy adattárházban. Ezeket az adatokat megtisztítják és kiigazítják a hagyományos elemzéshez.

De ami minket most érdekel, az az adatok százalékos 95 aránya, amelyek nem szerepelnek a adattárházban, hanem a nyers adatbázisokban.

A szükséges adatok eltérnek a hagyományos statisztikában használt és nyersként leírt adatoktól. Az utóbbiak egyfajta természetes, feldolgozatlan és korlátlan újraforrás. Hogyan lehet tehát hozzáférni ehhez a forráshoz, és milyen adatokkal rendelkezik valójában a szervezet? Az állítólagos nyers adatokhoz korántsem könnyű hozzáférni. A korábbi re- szervezetek egy elosztott adattárolási rendszerhez vezettek az adminisztrációban. Az adatoknak ezért rengetegféle tulajdonosa van, és a szervezetben mindenhol elhelyezve. Adatosított énjeink messze nem központosított, egységes entitások. A polgárokról szóló információdarabkák mennyisége, tartalma és holléte of- ten bizonytalan.

Az adatokhoz való gyakorlati, technikai hozzáférés pedig finoman szólva is késedelmesnek tűnik. Sokkal többre lett volna időnk, ha az adattudósoknak nem kellett volna ennyi idő ahhoz, hogy kitalálják, milyen adataink vannak, hol vannak, és a szervezet melyik egységével kell konzultálni a hozzáféréshez.

Az önkormányzati hivatalok és a központi kormányzat közötti szervezeti és adminisztratív szakadék ráadásul megnehezíti az adatok továbbítását. A különböző szervezeti egységekben tárolt adatokat még nem engedték összeállítani, illetve nem állították össze. Az adatokhoz való hozzáférést például a

speciálisan formázott számítógépeken biztosítják, de nem feltétlenül a megfelelő eszközökkel rendelkező számítógépeken.

az adatok elemzéséhez. Ezenkívül az önkormányzati és a kormányzati adatok összekapcsolása még nem volt lehetséges.

Ismét a régi és az új egyértelmű érzékelése. Ez nem csak az adatok történelmi perspektívájáról szól, hanem arról is, hogy az adatoknak milyen szerepet kell játszaniuk. Az új adatok az előrejelzések szerint agilisek és dinamikusak lesznek, f törvénytelenül vándorolnak a szervezet egészen keresztül. Az aranybányához való hozzáférés a különböző forrásokból származó adatok összehozását és formázását jelenti, vagy - metaforikusan szólva - alagutak és aknák építését az arany eléréséhez és szállításához, hogy az feldolgozható legyen. A pontok összekapcsolásáról van szó, egy már meglévő infrastruktúrára építve, hogy az adatáramlást a gépi tanulási algoritmusok felé irányítsuk. Mivel az adatok összegyűjtésére tett korábbi kísérletek gyakran kudarcot vallottak, és úgy tűnik, kevesen érzik magukat felelősnek az adatokhoz való hozzáférés általános kezeléséért és azért, hogy milyen adatok milyen formátumban állnak rendelkezésre, az adattudósok jelentős időt fordítanak arra, hogy szövetségeket keressenek az elosztott nyilvános archívumokban. Ezek az archívumok azonban egyértelmű jeleit mutatják annak, hogy sosem szánták őket bányászatra, és olyan kapuőrökkel rendelkeznek, akik még nincsenek tudatában kapuőri szerepüknek.

Minőség

Az adatokhoz való hozzáférés megszerzése után az adatokat gyakran vizualizálják és megvizsgálják a minőségük megállapítása érdekében. A minőséget itt az adatok mennyiségében és teljességében, valamint az adatokban tárolt információk pontosságában mérik. Az adatminőséghez jelentős bizonytalanság kapcsolódik, mivel az adatok tulajdonosai keveset tudnak az adatkészleteikről. A gépi tanulási algoritmusok nemcsak hatalmas mennyiségű adatra, hanem bizonyos fokú minőségű adatokra is támaszkodnak ahhoz, hogy bármilyen osztályozást vagy előrejelzést tudjanak készíteni.

Fontos azonban, hogy megértsük, hogyan hatnak az adatokra, és hogy ezek az adatok mit mondhatnak el nekünk, és hogyan fogják befolyásolni az általunk épített modelleket is. Mert a modelljeink mindennek ellenére nem többek annál, mint amit beljük táplálunk és amire betanítjuk őket.

Ez az aranybányászati folyamatnak ebben a szakaszában van, hogy a teljes arany metafora megreped. Az adatok, az arannyal ellentétben, nem természetes módon jelennek meg a természetben (Cheney-Lippold 2017). Hetven informátor hangsúlyozza annak megértésének fontosságát, hogy az adminisztrációban tárolt adatok nagy részét gépekkel együttműködő emberek állították elő. Nincs olyan, hogy nyers adat. A nyers adatok fogalma, ahogy Bowker (2005) rámutat, oximoron.

Mielőtt egy adatbázisban vagy archívumban tárolnák, számos szelekciós és

ma- nipulációs lehetőség van. Még ha az adathalmazok többé-kevésbé teljesnek is tűnnek, az adatbejegyzések értelmezésével kapcsolatban további komplexitás merül fel: mit

pontosan mérik, és hogyan végezték a méréseket? Az adatok szituált tudás, társadalmilag konstruáltak, történelmileg kontingensek és kontextusfüggők. Az adatok nyilvántartásának és tárolásának megfelelő megértése kulcsfontosságú az adatközpontúvá válás átfogó céljának eléréséhez. A nyilvános archívumban található adatok a közigazgatásban alkalmazott munkamódszerek eredményei.

Kontextus nélkül az adatok értelmetlennek tűnnek a felhasználók számára. Amikor például a polgárok foglalkoztatási/munkanélküliségi státuszára vonatkozó egyszerű nyilvántartási adatokat vizualizálták, a csapat hamarosan üres helyeket fedezett fel. Mik ezek az üres helyek? Egy munkáltató, aki elfelejtett bejelenteni egy alkalmazottat, vagy egy munkanélküli, aki nem jelentette be a munkanélküliségét? Esetleg elírás történt az út során, vagy esetleg hiba történt valamelyik regisztrációs infrastruktúrában? Egyszerűen nem könnyű megmondani, hogy mi történt, és ezért kihívást jelent a kezelésük. Ezért egy felhasználó-támogató munkatársat kértek fel, hogy kontextusba helyezze a regisztrált adatokat, megvitatta a munkamódszereket az adattudományi környezettel. A hiba kívánt számszerűsítése azonban nem történt meg a megfigyeléskor. A gépi tanulást gyakran vádolják azzal, hogy legitimálja társadalmi hatalmát, mivel matematikusnak, logikusnak, pártatlannak, következetesnek és ezáltal objektívnek tűnik (Gillespie 2014). Meglepő módon az objektivitás nem eleme a csapat artikulációs munkájának. Itt a résztvevők hangsúlyozzák, hogy maga a prototípusuk, a közügynökség felhasználói támogatás, nem objektív. Módszereiknek ezért nem kell kemény tények. A pontosság fontosabb a csapat számára, mint az objektivitás. Nincsenek tökéletes adatok vagy nyers adatok. Az informátorok mégis úgy gondolják, hogy képesek lesznek valamilyen alkalmazható jelentést kihozni az adatkészletekből, amelyek kibővítik a hagyományos statisztikákból származó ismereteket.

Adatvédelem

Az adatok előkészítésében az adatok tudósai számára egy harmadik összetettség az adatok biztonságával kapcsolatos kérdésekhez kapcsolódik. Ki használhatja az adatokat? Milyen adatok használhatók fel? Milyen adatok nem elemezhetők együtt? Hogyan lehet biztonságosan elszállítani az aranyat a bányából az algoritmusba? Ez egy interdiszciplináris és széleskörű kihívás. Több informátor az adatvédelemmel és az információbiztonsággal kapcsolatos munkát tartja munkája legfontosabb és egyben legmegerőltetőbb részének. Mivel nincs konkrét keret arra vonatkozóan, hogy az adatokat hogyan lehet és kell felhasználni, és milyen adatokat lehet felhasználni, a résztvevőknek új kereteket kell kialakítani az adatok etikai és jogi felhasználására a közhivatalokban. A nagyméretű adatok felhasználása újdonság a szervezet, valamint a norvég közszféra számára. Több hivatalos jelentés is rámutat a jogi iránymutatások hiányára a gépi tanuláson keresztül történő nagy adatfelhasználáson belül (Technologiråd 2018). Bár a politikai beszédekben gyakran említik, a da- ta-vezérelt jóléti állam egy jövőbeli

képzelet, gyakorlati jelenbeli iránymutatások nélkül.

A nem létező jogi keret bizonytalansághoz vezet az adattudósok körében. Csak azért, mert az adatok hozzáférhetőek, még nem etikus automatikusan feldolgozni ezeket az adatokat. A gépi tanulás nemcsak a magánélet, hanem az igazságszolgáltatás területét is érinti. Bár a közigazgatás már régóta felelős hatalmas mennyiségű, rendkívül érzékeny adat kezeléséért, az adatok újbóli felhasználását a saját gyakorlatában még nem vizsgálták. A korábban figyelmen kívül hagyott adatok 95%-át nem kezelik kellőképpen. A rendkívül érzékeny adatokkal való munka során a józan ész és a megérzések függvénye igényesnek és nemkívánatosnak tekinthető. A hibák következményeit jelentősnek képzelik.

Ezt nem hagyhatjuk. Minden leállna. Nagyrészt hihetetlen mennyiségű információval rendelkezünk Norvégia teljes lakosságáról. És rengeteg információval rendelkezünk az emberek életének legkiszolgáltatottabb és legnehezebb helyzetéről.

Az adatvédelem az adatok etikus és biztonságos felhasználásának értékeléséről szól. A helyes HSE megvalósításáról szól az Ön aranybányászati projektjében. Több informátor összehasonlítja a közigazgatásban végzett munkát a magánszektorban végzett gépi tanulási algoritmusokkal kapcsolatos munkával. Bár a magánszektor nagy utat tett meg a gépi tanulás területén, a résztvevők nem feltétlenül akarják átvenni a magánszektor szereplői által készített gyakorlatokat és modelleket. Fontos számukra, hogy megfelelő módon teremtsenek és elősegítsék a felhasználók közötti bizalmat. A polgárok elvárják tőlük az adatok biztonságos kezelését. A jogi irányelvek hiánya itt a kísérletezés szabad terét jelenti. Több informátor is kiemeli, hogy nemcsak a jogi, hanem az erkölcsi és etikai szempontok figyelembevételével is fontos eljárni. Az erkölcsi és etikai magatartás számszerűsítése és alkalmazása az adatokkal való munka során azonban korántsem egyszerű. Eddig a közigazgatás inkább tartózkodik az adatok felhasználásától, minthogy olyan hibákat kövessen el, amelyek hatással lehetnek az állampolgárok bizalmára.

Megbeszélés és következtetés

Ezt a vitát és a következtetéseket azzal kezdjük, hogy visszatérünk az algoritmikus infrastruktúrák "fekete dobozok" metaforájához. Ez a metafora a hatalmat és az ellenőrzést gyakorló belső vállalat és az erőtlen és tudatlan külső hatalom képzetét idézi fel. Természetesen, mivel egyre több döntést gépi tanulási modellek alapján hoznak meg, az átláthatóság és a befolyás hiánya komoly demokratikus fenyegetést jelent. Ezért a kritikai algoritmus-tanulmányok egyik központi feladata az ilyen "fekete dobozok" konstitutív elemeinek kibontása és vizsgálata.

Itt az átláthatóságot nem lehet egyszerűen egy közzétételi kódexszel elérni,

amit a közsférában egyesek javasoltak. Úgy véljük, hogy az algoritmuskonstrukció apró részleteinek etnográfiai tanulmányozása fontos hozzájárulás az algoritmuskonstrukcióhoz.

A belső ellenőrzés mértékének árnyaltabb felfogása a belső ellenőrzés mértékének árnyaltabb felfogása. Seaver (2017) terepmunkája bemutatja a programozás összetettségét és rendetlenségét, valamint az adattudósok bizonytalanságát az algoritmikus feldolgozás beléptetése és eredménye közötti kapcsolatról. Tanulmányunk a kontroll képzetének egy másik részét bontja le, az algoritmusok bizonytalan alapjának bemutatásával. Az algoritmusokba táplálható adatokra vonatkozó döntések ritkán egyértelműek és egyértelműek, hanem a hiányzó értékek, szöveges esetlegességek, kontextusfüggőségek és értelmezési hiányosságok kezelését foglalják magukban. Az adatok gépi tanulásra való alkalmassá tételének folyamata gyakran láthatatlan.

E kihívások némelyike általánosítható az adatelőkészítés minden típusára, a magánvállalkozásokon belül is, valamint a mélytanulás és a neurális hálózatok alkalmazásaira. Végül is nincs mesterséges intelligencia adatok nélkül. Mások inkább a közzféra archívumainak felhasználására jellemzőek. A közintézményekben található hatalmas adathalmazokat - informátoraink is - "aranybányaként" jellemezték a gépi tanulási algoritmusok fejlesztéséhez, amelyek felhasználhatók a polgárok jobb és személyre szabottabb szolgáltatások nyújtására. A politikusok és döntéshozók sok reményt és izgalmat fűztek az adat-aranybányához. Esettanulmányunk azonban azt mutatja, hogy az ilyen archívumok hasznosításával kapcsolatos kihívások, ha nem is leküzdhetetlenek, de legalábbis sokkal nagyobbak és igényesebbek a vártnál. A gépi tanuláshoz olyan varázslatos érzés kötődik, amely minimálisra csökkenti a figyelmet az eredmények előállításához szükséges módszerekre és erőforrásokra (boyd/Elish 2018).

Az első kutatási kérdésünk a közzféra archívumaiból származó, mesterséges intelligencia alapú közszolgáltatások fejlesztésével kapcsolatos kihívásokra vonatkozott. Ebben a fejezetben három olyan típusú kihívást választottunk, amelyekkel az adattudósok munkájuk korai szakaszában szembesültek. Először is, az adatokhoz való hozzáféréssel kapcsolatos jelentős akadályok, mind szervezeti, mind technikai szempontból. Ezek az akadályok abból adódnak, hogy a kormányzati adatoknak rengeteg különböző tulajdonosa van, és a szervezet egész területén vannak elhelyezve, mivel a korábbi átszervezések elosztott adattárolási rendszerhez vezettek. Ezen túlmenően a közigazgatáson belül az egyes adatkészletek kapuőreit gyakran nem könnyű megtalálni, vagy nincsenek tudatában kapuőri szerepüknek. Az információbiztonsági kockázatok miatt az adatok nehezen vándorolnak a szervezeten belül. Egy másik kihívás az adatkészletek adatainak minőségével és jelentésük értelmezésével kapcsolatos. Az adattudósok hamar felfedezték, hogy az adatkészletek közül sok hiányzó értékekkel és közelítésekkel van tele, és hogy a számok nehezen értelmezhetők a cél és a nyilvántartásuk kontextusának ismerete nélkül. Mit is mértek pontosan? Hogyan készültek a mérések? Végül az adattudományi környezetnek számos összetett jogi és biztonsági kérdéssel kell foglalkoznia, ami halmozottan megnehezíti a munka előrehaladását. Ki használhatja fel az adatokat? Milyen

adatok használhatók fel? Mely adatkészletek kapcsolhatók össze? Mivel nincs létező hivatalos jogi keret arra vonatkozóan, hogy hogyan kell dolgozni

az adatokkal a gépi tanulással együtt, az adattudósoknak iránymutatásokat kell kidolgozniuk az út mentén - extra biztonsági tartalékokkal kiegészítve.

Érdekes módon sok hasonlóságot találhatunk az adattudományi környezet tárgyalta kihívásai és a társadalomtudósok által felvetett kritikus kérdések között (Crawford / Boyd 2012). A gépi tanulási algo- rithmusokon dolgozó adattudósok jól ismerik a terület komplexitását és f törvényeit, amelyben tevékenykednek. Emellett a módszertani kihívásokban is találhatunk hasonlóságokat a társadalomtudományok és a gépi tanulás művelése között. Boydhoz és Elishhez (2018) hasonlóan ezért mi is az adattudósok és a társadalomtudósok közötti szakértelemcserére kívánjuk felhívni a gépi tanulóval foglalkozók figyelmét. A szakudás szélesebb körének bevonása az egyik út a digitális infrastruktúrák alakítására gyakorolt társadalmi befolyás növeléséhez (Ananny/Crawford 2018).

Az előkészítő munka mennyisége és összetettsége bizonyos mértékig meglepetésként érte az adminisztrációt - az adattudósok számára napot kellett eltöltenie azzal, hogy fel és alá bolyongott a folyosókon, be és ki az irodákból, poros archívumokban kutatott, régi adathalmazokat vizsgált és értelmezett, és megismerkedett a homályos jogi keretekkel és a zavaros szervezeti biztonsági iránnyelvekkel. Így a napjai olyan feladatokkal teltek ki, amelyek elvileg nem tartoztak a szakterületéhez, miközben alig kezdett hozzá azokhoz a feladatokhoz, amelyekre alkalmazták - gépi tanulási algoritmusok létrehozásához és bővíthetőségéhez. A terepmunka a felfedezés és a bizonytalanság fázisában folyt. Az újonnan kialakult adattudományi környezet még nem érte el az úgynevezett zúzásponot. Az adattudományi környezet még mindig azon dolgozott, hogy kikövesse az utat a gépi tanulási algoritmusok felé, hogy az adatokat gépi tanulhatóvá tegye. A jövőbeli adatvezérelt képalkotás diffúz volt, és nem rendelkezett jelenlegi iránnyelvekkel. A felmerült kihívások így szakítást jelentettek a zökkenőmentes és lenyűgöző funkcionalitás adatvezérelt mítoszával, és komoly kérdéseket vetettek fel azzal kapcsolatban, hogy mi lehetséges és mi a ténylegesen reális (boyd/Elish 2018). Az adattudósok ahelyett, hogy munkájukat úgy jellemezték volna, mint az adatvezérelt válás felé irányuló munkát, inkább úgy érezték, hogy tudatosabb kapcsolatot kezdeményeztek az adatokkal.

Végző soron úgy tűnik, hogy az adattudományi környezetet arra a küldetésre állították be, hogy

a szervezet általános adatkezelési gyakorlatának átalakítása. Ez azonban nem korlátozódott az adatgyakorlat pusztá automatizálására. A csapat célja az volt, hogy megváltoztassa az adminisztrációban tárolt adatok és maga az adminisztráció közötti kapcsolatot. Az adatokat itt úgy képzelték el, hogy nagyobb hatalmat és bizalmat kapjanak a személyre szabás, a hatékonyság fokozása és a felhatalmazás általános céljának elérése érdekében. Azonban azok, akik a legnagyobb hatalmat tulajdonították a nyilvános archívumnak, nem azok voltak, akik közvetlenül a gépi tanulási algoritmusokon dolgoztak. Az adattudósok számára állandó küzdelem folyt az adatvezérelt jóléti állam nagy mítosza és a gépi tanulóval kapcsolatos valós tapasztalatok között. Ezt erősíti a saját küzdelemünk is, hogy összehangoljuk az aranybánya metaforát,

formants a mi empirikus bizonyítékunkban szereplő megállapításokkal. A feltételezett aranybánya talán nem is tartalmaz aranyat. Az adatvezérelt képképzés alapja maga is bizonytalannak tűnt.

Még nincs egységes megoldás arra, hogy hogyan lehet és kell megközelíteni a data-vezérelt képeket. Ez azt is jelenti, hogy a mesterséges intelligencia-alapú közszolgáltatások fejlesztésével kapcsolatos adatgyakorlatok rekonstrukcióinak és konfigurációinak még van helye. Ahogy Cheney-Lippold (2017: 13) érvel: "Aki az adatok nevében beszél, [...] rendkívüli hatalommal rendelkezik ahhoz, hogy keretbe foglalja, hogyan jutunk el egy jelenség magyarázatához". Maga a gépi tanulás demokratizálására irányuló felhívás diffúz és fluid, ahogyan a közszférán belüli adatvezérelt válasz átfogó célja is (vö. Sudmann 2018). A mesterséges intelligencia felhatalmazó és demokratizáló potenciáljának reanalízisa a kormányzat és a polgárok közötti kapcsolatokban végső soron azokon múlik, akik az adatokat előkészítik és az algoritmusokat konstruálják. Attól függ, hogy az adattudósok és a szervezetek hogyan tesznek eleget a felmerülő bizonytalanságoknak és módszertani kihívásoknak. Ahhoz, hogy ne ragadjanak magukkal a mesterséges intelligenciát övező mítoszok és hipterek, kutatnunk kell a hétköznapi tárgyalásokat és döntéseket, és figyelmünket a gépi tanulás előállításához szükséges módszerek és erőforrások felé kell fordítanunk. Csak az ilyen jellegű munkával kapcsolatos valós tapasztalatokba való betekintéssel leszünk képesek elkezdni feltenni a helyes kérdéseket, és képesek leszünk az adatvezérelt jövőnk irányítására.

Hivatkozások

- Ananny, Mike/Crawford, Kate (2018): "Látni tudás nélkül: Az átláthatóság eszményének korlátai és alkalmazása az algoritmikus elszámoltathatóságra." In: *New Media & Society* 20/1, pp. 973-989.
- Beer, David (2017): Beer: "Az algoritmusok társadalmi hatalma." In: *Information, Communication & Society* 20/1, pp. 1-13.
- Bowker, Geoffrey C. (2005): Cambridge, Massachusetts: *Memory Practices in the Sciences*: MIT Press.
- boyd, danah/Elish, Madeleine Clare (2018): "A módszerek elhelyezése a Big Data és a mesterséges intelligencia varázslatában". In: *Communication Monographs* 85/1, pp. 57-80.
- Burrell, Jenna (2016): "Hogyan "gondolkodik" a gép: Machine learning algoritmusok átláthatatlanságának megértése." In: *Big Data & Society* 3/1, pp. 1-12.
- Cheney-Lippold, John (2017): *We Are Data: Algorithms and the Making of Our Digital Selves*, New York: NYU Press.
- Crawford, Kate/boyd, danah (2012) "Kritikus kérdések a nagy adatokhoz - felhívások egy kulturális, technológiai és tudományos jelenséghez". In:

- Desrosières, Alain (1998): *La politique des grands nombres*. Cambridge, Massachusetts: Harvard University Press.
- Flyverbom, Mikkel/Murray, John (2018) "Datastructuring-Organizing and curating digital traces into action". In: *Big Data and Society* 5/2, pp. 1-12.
- Gillespie, Tarleton (2014): "Az algoritmusok jelentősége". In: Tarleton Gillespie/Pablo J. Boczkowski/Kirsten A. Foot (szerk.), *Media Technologies: Essays on Communication, Materiality, and Society*, Cambridge, Massachusetts: The MIT Press.
- Kitchin, Rob (2017) "Az algoritmusokról való kritikus gondolkodás és kutatás". In: *Information, Communication & Society* 20/1, pp. 14-29.
- Kommunal- og moderniseringsdepartementet (2016): *Digital agenda for Norge: IKT for en enklere hverdag og økt produktivitet* (Meld. St. nr. (272015-2016)), Oslo: Kommunal- og moderniseringsdepartementet.
- Latour, Bruno (1987) *Science in action: how to follow scientists and engineers through society*, Milton Keynes: Open University Press.
- Mayer-Schönberg, Viktor/Cukier, Kenneth (2013) *Big Data: London: A Revolution That Will Change How We Live, Work and Think*: John Murray.
- NAV-ekspertgruppen (2015): *Et NAV med muligheter*. Sluttrapport fra ekspertgruppen, Oslo: Arbeids- og sosialdepartementet.
- Pasquale, Frank (2015) *The black box society*: Cambridge, Massachusetts: Harvard University Press.
- Prince, Christopher (2017): Prince, Prince: "Big Data and Privacy: Adatvédelem: Miért fogadják el a közzféra szervezetei a nagy adatokat?". In: *Canadian Journal of Information & Library Sciences* 41/4, pp. 233-244.
- Redden, Joanna (2018) "Demokratikus kormányzás az adatosítás korában: A kormányzati diskurzusok és gyakorlatok feltérképezésének tanulságai". In: *Big Data and Society*, pp. 1-13.
- Ruppert, Evelyn/Isin, Engin/Bigo, Didier (2017) "Data politics". In: *Big Data & Society* 4/2., pp. 1-7.
- Sudmann, Andreas (2018) "A mesterséges intelligencia médiapolitikai dimenziójáról: Deep Learning as a Black Box and OpenAI". In: *Digital Culture & Society*, Volume Issue 4, Pages 1, 181-200.
- Teknologirådet (2017): *Denne gangen er det personlig: Det digitale skiftet i offentlig sektor*, Oslo: Teknologirådet.
- Teknologirådet (2018): *Kunstig intelligens-muligheter, utfordringer og en plan for Norge*, Oslo: Teknologirådet.
- Vivento AS/Agenda Kaupan AS. (2015): *Kartlegging og vurdering av stordata i offentlig sektor*. Oslo: Kommunal- og moderniseringsdepartementet.

Plurális, helyhez kötött szubjektumok a mesterséges intelligencia kritikájában

Tobias Matzner

1. Bevezetés

A mesterséges intelligencia (AI) területéről származó információs technológiákkal kapcsolatos számos jelenlegi kritikai álláspont az emberi alanyok és a technológia közötti különbségre összpontosít. Ezeknek az álláspontoknak két változata létezik. Az első változat a technikai semlegesség gondolata. Szerencsére az a régi érv, miszerint a technológia semleges, hogy társadalmi hatása "csak attól függ, hogy mit teszünk vele", egyre inkább veszít a hatásából.

Ezt az érvet azonban gyakran azzal cáfolják, hogy az algoritmusok nem újszerűek, mert emberek készítik őket. Hasonlóképpen, absztraktabb szinten gyakran állítják, hogy a gépi tanulási algoritmusok képzéséhez használt adathalmazok az emberi társadalmat tükrözik, és így annak igazságtalanságait és előítéleteit importálják (Campolo et al. 2017; O'Neil 2017). Ez azt jelenti, hogy az algoritmusok semlegesek lehetnének, ha az emberek nem rontanák el őket folyamatosan az elfogultságaikkal. Ez egy nagyon determinista, platói történet, ahol az emberi ötletek és cselekedetek a meghatározóak, amelyeket aztán kódba foglalnak és a gépek végrehajtanak (Chun 2008).

Ezért fontos, hogy a kritika második változatához forduljunk. Ez azokat az álláspontokat foglalja magában, amelyek azt mutatják, hogy az emberi szubjektivitás nem az információs technológián kívüli dolog, amelyet azután ez a technológia elfogult vagy elfogulatlan módon reprezentál.¹ Inkább azzal érvelnek, hogy a digitális technológia magával az emberi szubjektivitással tesz valamit.

E megközelítések többsége azonban általános ítéletet alkot az adatalapú vagy algoritmikus szubjektivitásról, amelyet általában a támogatott jellemzők egyfajta veszteségként írnak le. A következőkben az ilyen elméletekkel fogok foglalkozni, és néhány eset segítségével bemutatom, hogy az ilyen általános ítéletek miért hordozzák magukban annak a veszélyét, hogy nem veszik figyelembe azt a fontos tényezőt, hogy a mesterséges intelligencia konkrét alkalmazásai egészen más módon kapcsolódnak a már létező társadalmi-technikai helyzetekhez és a szubjektivitás megfelelő formáihoz. Linda Martin Alcoff posztkoloniális teoretikus munkásságát fogom felhasználni, hogy egy

I Az ilyen, a reprezentáció episztemikus kritikájára épülő megközelítéseket részletesen tárgyalja a következő művekben
(Matzner 2016).

a szubjektivitás olyan fogalma, amely képes megragadni a közelmúlt technológiai változásainak hatását, ugyanakkor a különösen szituált szubjektumok közötti különbségeket emeli ki normativitásának forrásaként - a szubjektivitás algoritmikus formáinak általános jellemzője vagy hiánya helyett.

2. A mesterséges intelligencia alkalmazásai és a kritika két formája

A mesterséges intelligencia technológiai egyre inkább strukturálják a digitális kommunikációt és interakciót, de az "offline" térnek tekintett tereket is. Különösen a gépi tanulásból származó prediktív technológiák központi szerepet játszanak a digitális platformok jelenlegi szolgáltatásaiban. Ezeket használják a keresési eredmények személyre szabására, a közösségi médiában megjelenő posztok szűrésére, arra, hogy javaslatot tegyenek arra, hogy milyen tartalmakat nézzünk, és kivel lépünk kapcsolatba. Az ilyen prediktív technológiák különböző intézményi és kereskedelmi folyamatokat is áthatottak. Közismert, hogy a hitelekkel, biztosításokkal és munkaerő-felvételekkel kapcsolatos döntéseket gépi tanulási algoritmusok által szolgáltatott pontszámok befolyásolják. A biztonsági ügynökségek és a rendőrség világszerte alkalmaz mesterséges intelligenciával feljavított megfigyelési technológiákat a határellenőrzésben, a vízum- és menedékkérelmek feldolgozásában, a CCTV-felvételek automatikus értékelésében vagy - az algoritmikus elfogultság posztergyermeké - a visszaesés előrejelzésében (Angwin és Larson 2016).² A gépi tanulás prediktív felhasználása a célzott hirdetéseket és más "predikciós termékek" létrehozását is elősegíti, ahogy Shoshanna Zuboff nevezi őket (Zuboff 2019). Az algo- ritmikus technológiák, a munka és az értékteremtés pontos kapcsolata a digitális gazdaságban azonban még tisztázásra vár (Heilmann 2015; Srnicek 2016).

Sok kritikai munka folyt a de- az ilyen algoritmikus előrejelzésekből, azok episztemikus státuszából és abból a tendenciából, hogy a gépi objektivitás aurájába burkolják az elfogultságokat (Aradau and Blanke 2015; Kitchin 2014, 2017; Pasquale 2015). Máshol amellett érveltem, hogy ezeket a fontos vizsgálatokat kritikai vizsgálattal kell kiegészíteni azzal kapcsolatban, hogy mit tesznek ezek az algoritmikus gyakorlatok az alanyokkal (Matzner 2016). Például a közösségi médián belüli mindennapi interakciók megfigyelési célú felhasználása a gyanakvás és a bizalmatlanság jelenségeit és gyakorlatát importálja ezekbe az interakciókba.

Ezt az intuíciót követve fontos megkérdezni, hogy a mesterséges intelligencia-alapú technológiák növekvő hatása milyen új szubjektivitási formákat, illetve a szubjektivitás formáinak változásait eredményezi. Számos kritikai beszámoló, beleértve az aktivista pozíciókat is, implicit módon feltételezi a szubjektivitásnak a liberális politikában uralkodó modelljét.

² Az itt említett mesterséges intelligencia-alkalmazások mindegyikének kutatása szinte naponta

bővül. Cathy O'Neil (2017) könyve jó kiindulópont az általam itt említett alkalmazásokról szóló hivatkozásokhoz - még akkor is, ha kritikája az általam kritizált technológiával szemben az autonóm alanyok védelmének vonalába esik.

gondolkodás: racionális, önreflexív és autonóm szubjektum. A gépi tanulási technológiákat alkalmazó algoritmikus folyamatokat úgy tekintik, mint amelyek mindezen szempontok érvényesülését akadályozzák. Az úgynevezett "szűrőbuborékról" szóló diskurzusok például arra összpontosítanak, hogy az algoritmikus szűrés révén inkább az érzelmi, mint a racionális diskurzusok érvényesülnek, az algoritmusok átláthatatlansága miatt az autonóm ítéletekhez szükséges önreflexív gondolkodás sérül, és így a technológia autonóm használata már nem lehetséges (Pariser 2011; Zuiderveen et al. 2016). A liberális szubjektumok és a technológiai kényszerek közötti egyértelmű ellentét azonban túl egyszerű. A kibernetika egész történetét, amely a jelenlegi konnektionista mesterséges intelligenciához vezetett (Sudmann 2018), a liberális eszmékhez való mélyen ambivalens viszony strukturálta. Egyfelől a kibernetikát az az elképzelés vezérelte, hogy új és erőteljes eszközöket fejlesszenek ki a szabad és hatékonyabb emberi cselekvésekhez. Másrészt az ebből következő elképzelés, miszerint az ember, az állat és a gép alapvetően az irányítás és a kommunikáció ügye, mélyen fenyegeti az autonómia és az önállóság eszméit (Hayles 1999: 87). A mesterséges intelligencia legújabb alkalmazásai is sok tekintetben liberális projektnek tekinthetők (Matzner 2019). Ráadásul a konkrét kihívások, amelyeket a mesterséges intelligencia jelenlegi alkalmazásai jelentenek, nem oldhatók meg könnyen egyéni szinten. Például a magánélet és az adatvédelem kérdései, ha a liberális paradigmán belül oldódnak meg, az adatok személyes adatokra való felosztását feltételezik, amelyeket minden egyes érintett egyén ellenőrizhet (Matzner 2014). A jelenlegi mesterséges intelligencia által vezérelt adatelemzés vonzereje azonban abban rejlik, hogy az adatokat aggregált szinten használja, ami olyan összefüggéseket és asszociációkat talál, amelyeket nem lehet az egyes felhasználók hozzájárulására redukálni. Még az olyan személyre szabott rendszerek, mint az ajánlórendszerek vagy az idővonal-szűrő algoritmusok sem tárolják általában a felhasználó digitális modelljét, ahogy azt az "adatkezelők" és más fogalmak használata sugallhatja (Lyon 2014). Inkább minden egyes elemre vonatkozóan hoznak döntést, amellyel kapcsolatban a felhasználó érdeklődésének egy közelítése a felhasználó aktuális adatfolyamából és kapcsolatainak állapotából de- rivedálódik.³ Így az ilyen problémákat az adatfelhasználás összesített szintjén kell kezelni, nem pedig csak az egyes részek esetében. Végezetül a liberális elméletet a feminista és más kritikai elméletek vizsgálat alá vonták, mert részt vesz abban, amit Hayles "olyan gyakorlatoknak nevez, amelyek rossz hírét keltették a liberalizmusnak" (Hayles 1999: 87).

Ezen okokból kifolyólag a mesterséges intelligencia alkalmazásainak kritikai elméletei, amelyek a normativitás más forrásai előnyösebbek. Ennek egyik kiemelkedő és kidolgozott példája Antoinette Rouvroy algoritmikus kormányzásról alkotott koncepciója. Ő a nor-

3 Mint általában, nehéz pontosan tudni, hogy az olyan kiemelkedő alkalmazások, mint a Twitter

idővonala vagy a Face- book hírfolyamának szűrése hogyan történik. Így megfigyelésemet a közzétett kutatásokból merítem. Már a Yahoo-nál végzett korai kutatások (De Francisci, Morales et al. 2012), amelyek a személyre szabott tartalmakkal kapcsolatos újrakeresések élére álltak, nem használtak tartós modelleket a felhasználóról. A megközelítés támogató vektorgépeket használ az osztályozáshoz. Időközben a személyre szabás, mint a legtöbb más gépi tanulási feladat, áttért a neurális hálózatokra, és így a még inkább adatvezérelt és dinamikus megközelítésekre. Lásd például a Microsoft Research egy nemrégiben megjelent tanulmányát (Zheng et al. 2018).

matív álláspontot az emberiség olyan eszméjéből, amely éppen a teljes autonómia és racionalitás hiányán alapul. Rouvroy Judith Butler és Louis Althusser elméleteit követi (Rouvroy 2013: 158). Mindketten úgy írják le a szubjektumokat, mint akik soha nem rendelkeznek önmaguk felett, mert alapvetően másoktól függenek. Ezek a mások azonban nem egyszerűen meghatározóak. A mások ránk gyakorolt hatása olyan társadalmi inter- akciókban történik, amelyeket sem mi, sem a mások nem ellenőriznek teljes mértékben. Az emberi cselekvésnek éppen ez a többlete és nyitottsága teszi lehetővé a kritikát és az értelmes interakciót. Bár éppen ezt a többletet fenyegetik az algoritmusok:

[W]it kell megőrizni, mint a "szubjektum" és a so- cialitás, mint a világ többlete az algoritmikus valóság felett, a "közös", ez a "köztes", a közös megjelenés (*komparáció*) tere, amelyen belül kölcsönösen megszólítjuk egymást. (Ibid.: 159-60)

Így Rouvroy úgy látja, hogy az emberi interakció az újdonság és a szponzivitás olyan lehetőségét rejti magában, amelyet a számítástechnika soha nem tud megragadni. Az algoritmikus kormányzás, a fenti leírással összhangban, nem az egyes alanyokra összpontosít. Az algoritmikus kormányzás inkább "[e]gyüttműködik az in- formációs és fizikai architektúrák és/vagy környezetek újrakonfigurálásán keresztül, amelyeken belül bizonyos dolgok lehetetlenné vagy elképzelhetlenné válnak, és riasztásokat vagy ingereket vetnek be, amelyek inkább reflexszerű válaszokat produkálnak, mint értelmezést és reflexiót". (Ibid.: 155.) Ez a leírás egyértelműen visszahangozza a szubjektum elvesztésével kapcsolatos kibernetikai aggodalmakat. Az algoritmusok Rouvroy szavaival élve a virtuálisat a ténylegesre, a lehetségeset a statisztikailag valószínűre, az élőt a számításossá redukálják (Rouvroy 2017). Így a Rouvroy által az algoritmikus kormányzással szemben megfogalmazott kritika fő irányvonala ismét a szubjektivitás bizonyos mértékű elvesztése, ebben az esetben a relációs szubjektivitás egy olyan formája, amely hozzájárulhat a politika és az ellenállás megteremtéséhez.

3. Általános szintű kritika és a helyhez kötött témák fontossága

Az ilyen elemzések fontos betekintést nyújtanak a mesterséges intelligencia alkalmazásának következményeibe. Rouvroy beszámolója különösen az AI-alapú ítéletek és tevékenységek számos újkori formájának sajátosságainak tesz igazságot, amelyek az egyén feletti szinten működnek, és amelyek inkább cselekvésre ösztönöznek, mint információra. Fontos megjegyezni, hogy a mesterséges intelligenciának vannak olyan alkalmazásai, amelyek nagyon is összhangban vannak a fegyelmi hatalom foucault-iasabb formáival (Matzner 2017). Ezek különösen a nyugati, kapitalista társadalmak határainál találhatók, amelyeket Rouvroy és a mesterséges intelligencia legtöbb más kritikusa a

fókuszba vesz. Mégis, ezeken a társadalmakon belül is helytállóak az ilyen elemzések. Azonban abban a kísérletükben, hogy általános ítéletet találjanak a szubjektívitás konkrét elvesztéséről a mesterséges intelligencia alkalmazásai révén, fontos minősítéseket hagynak ki.

Ez nemcsak leíró pontosság kérdése, hanem azt is jelenti, hogy a mesterséges intelligencia önmagában nem olyan antipolitikai technológia, mint amilyenek ezekben az elemzésekben tűnik. Az antipolitikai hatásai nem a szubjektumokra mint olyanokra, hanem egyes szubjektumokra esnek, és mindegyikre más-más módon.

Az ilyen általános ítéletek problémái például Wendy Hoi Kyong Chun szűrőbuborékokról szóló elemzésében láthatók. Bemutatja, hogy a szűrőbuborék elmélete a homofília fogalmán alapul: Az az elképzelés, hogy az emberek hajlamosak arra, hogy olyan emberek felé orientálódjanak, akik hasonlóak vagy hasonlóan gondolkodnak. A szűrőbuborék kritikusai azzal érvelnek, hogy az algoritmikus tartalomkészítés hajlamos ezt az emberi tendenciát veszélyes módon érvényesíteni, ami mindenféle szélsőséges és rasszista közösségekhez vezethet. Az algoritmikus szelekció problémáját azonban nem magában a tartalomban, hanem a hasonlóság fogalmában látják, amely minden tartalomra egyformán vonatkozik. Így a homofília

alibiként szolgál az általa leképezett egyenlőtlenséghez, miközben a politikát is elhárítja: a homofília (gyakran állítólag a diszkrimináltaké), nem pedig a rasszizmus, a szexizmus és az egyenlőtlenségek forrása lesz az egyenlőtlenség, ami az igazságtalanságot "természetessé" és "ökológiai" teszi. (Chun 2018: 76)

Az algoritmikus szűrés, amely példaértékű esete annak, amit Rouvroy "az információs [...] architektúrák újrakonfigurálásának" (Rouvroy 2013: 155) nevez, az emberi szubjektumképződés egyetemes vonása tekintetében kritizálódik. Chun megmutatja, hogy szükséges figyelembe venni a szubjektumok társadalmi helyzetét, amely lehetővé teszi a rasszizmust, a szexizmust, az in- egyenlőséget. Egy másik ilyen eset a ProPublica hírhedt elemzése lenne, amely kimutatta, hogy egy visszaesést előjelező szoftver elfogult volt a feketékkel szemben (Angwin és Larson 2016). Ezt az esetet már szinte túl sokat tárgyalták, ezért csak azt szeretném kiemelni, hogy a szoftver nem használt semmilyen faji jellemzőt bemenetként. Így, még ha az algoritmusok hatékonysága nem is a faji szempontok szerint működik, akkor is foglalkozik és termel faji szempontokat.

Annak érdekében, hogy legyőzzük az elején említett kritikát, amely az elfogult adatok által elrontott semleges technológiát feltételezi, meg kell mutatni, hogy a média és különösen a mesterséges intelligencia hogyan kapcsolódik a társadalmilag és kulturálisan meghatározott témákhoz - beleértve a faji kérdéseket is.

4. Helyhez kötött tárgyak

Az általa "szokásszerű új médiának" nevezett könyvében Chun az adatelemzésről és az egyéntől való elfordulásukról ír, ami nagyon hasonló Rouvroy-hoz. Elemzésének középpontjában a szokás fogalma áll: ahelyett, hogy az egyéni

alanyra összpontosítana, az adatelemzés a szokásokat, a bevett cselekvési módokat próbálja megragadni, és következésképpen megpróbálja megragadni a szokásokat.

hogy kialakítsa és befolyásolja ezeket a szokásokat. Ennek érdekében inkább a szokások közötti összefüggésekre összpontosítanak, mint az egyes cselekedetekre vagy akár az egyes mintákra. "Ezáltal az egyéni cselekedetek a kollektív minták jelzéseiivé válnak, nem pedig kivételekké." (Chun 2016: 57) Ezek a minták az optimalizálás tárgyát képezik, egészen hasonlóan Rouvroy leírásához, aki az építmények és környezetek újrakonfigurálását írja le bizonyos viselkedések elérése érdekében.

Itt nem tudom követni azt a részletes konceptuális munkát, amelyben Chun a szokás fogalmával foglalkozik. Szeretném azonban követni azt a javaslatát, hogy a szokás médiateóriából származó felfogását összekapcsoljam a szokásról szóló gondolatokkal, amelyek az alteritással kapcsolatosak:

A megszokás a nyilvánosság: mások tapasztalata, a sebhely, ami az énben megmarad. A ha- bitek a múlt maradványai - múltbeli célok/önmagunk, múltbeli tapasztalatok -, amelyek tovább élnek a környezetünkre adott reakcióinkban ma, ahogyan a holnapot várjuk. A megszokás révén az alteritás lakik bennünk és lakik bennünk. (Ibid.: 95)

Chun arra ösztönöz, hogy tegyük fel a kérdést, hogyan változnak meg ezek a szokások a digitális média legújabb fejlődésével, és hogyan változhatnak meg újra a társadalom megváltoztatása érdekében (ibid.: 8). Ez azt jelenti, hogy nem a szokás önmagában a probléma, hanem a szokások közötti különbségek. Maga Chun azonban nem veszi elég komolyan ezeket a különbségeket. Őt elsősorban a liberális felszólítások foglalkoztatják, hogy megvédje a szubjektumot az alteritástól és a technológiai kényszertől. Mindkettő ötvözésével arra buzdít, hogy találjuk meg a módját annak, hogyan "lakjuk be" a másokhoz fűződő szokásszerű viszonyainkat, ami magában foglalja, hogy "óvatosan fogadjuk el" az adatok, kapcsolatok, szubjektivitás konfigurációinak számos új áramlását. Itt sokkal pozitívabban áll a technológiához, mint Rouvroy. Mégis alábecsüli, hogy a csere és az ebből következő szubjektivitás minden formáját a hatalom alakítja - nem csak a privát, liberális teret. Egyes szocio-technikai pozíciók meglehetősen bántóak az in- habitusra nézve. A következőkben tehát egy középutat szeretnék javasolni, amely azonban osztja azt a szemléletet, hogy az érzékelésünk és az adott helyzetben való cselekvésünk módjának változása nem csak az algoritmikus kormányzási eszközök célja (lehet). Ezek a szubjektivitás működésének alapvető módját jelentik. Ezt részletesen elemzi Linda Martín Alcoff látható identitások című könyvében.

Alcoff abból a foucault-i belátásból indul ki, hogy a hatalom nem csak egy ráerőltetés. Inkább azt jelenti, hogy gondolkodó és cselekvő szubjektumnak lenni azt is jelenti, hogy a hatalomban helyezkedünk el. A foucault-i elemzéssel ellentétben azonban, amely a normák szerinti fegyelmi alávetettségre összpontosít, Alcoff az alteritás és a habituálódás elméletén keresztül mutatja be, hogy észleléseinket és cselekedeteinket az általunk végzett gyakorlatok és a

helyzetek alakítják, amelyekben találtuk magunkat. Korábbi tapasztalataink nyomokat hagynak maguk után, amelyeket Alcoff a társadalmi gyakorlatok elméleteként (Reckwitz 2002) gyakran tárgyalt elmélet központi meglátásaival összhangban ír le: "[A]z identitásunkat alkotó értelmezési horizontot kétségtelenül a testben található hallgatólagos tudás gazdagsága konstituálja [...]". (Alcoff 2006: 106) Ilyen hallgatólagos tudás és

a szokássá válásoknak a gyakorlatokban van a helyük. Nem feltétlenül ránk kényszerülnek, inkább annak a módnak a növekvő maradványai, ahogyan cselekszünk - vagy ahogyan arra kényszerülünk, hogy cselekedjünk. Ez utóbbi természetesen továbbra is fontos, de nem az egyetlen módja annak, ahogyan a habituálódás létrejön. Ez szerves része annak, ahogyan értelmet adunk a helyzetünknek, és ahogyan strukturáljuk a cselekedeteinket. Az észlelés és a cselekvés ezen módjainak nagy része másoktól származik - az oktatáson keresztül, a különböző kontextusokban, ahol élünk, dolgozunk, játszunk, tanulunk stb. Mindezeket a kontextusokat vagy helyzeteket kollektív gyakorlatok strukturálják. Gyakorlatok, amelyekben teszünk valamit, de ugyanakkor szubjektivitásra teszünk szert. Mások mondják meg nekünk - többé-kevésbé implicit módon -, hogy kik vagyunk, mivé válunk vagy válnunk kellene bizonyos dolgok elvégzésével, mi illik vagy szokásos "egy magadfajta" számára stb. Ahogy Alcoff megállapítja:

A kollektív gyakorlat részben az én aspektusait hozza létre. A preferenciáinkat, a bizonyos típusú érzésekre való hajlamainkat bizonyos típusú helyzetekben, azt, hogy mi vált ki jellemzően félelmet, szorongást, nyugalmat, dühöt és így tovább, befolyásolja a kulturális és történelmi helyünk. Néha az emberek az ilyen belső érzéseket természetes eredetük bizonyítékának tekintik, mint például amikor egy homoszexuális csók undor érzését váltja ki. Az érzések nagyon is valóságosak lehetnek, de ez nem bizonyíték arra, hogy a homoszexualitás nem természetes; a fizikai reakciókat ismeretség és ismerkedés megváltoztathatja. Ez a példa arra utal, hogy a másik milyen erőteljes szerepet játszik az én-formálásban: a másik jelleme nem kis részben meghatározza az ént. (Ibid.: 115)

A szubjektumról szóló elméleteket illetően gyakran fontos kiemelni ezt a hatást a veleszületett tulajdonságokról szóló elképzelésekkel vagy a minél önreflexívebbé válás igényével szemben. Ilyenkor elegendő, hogy a szubjektum "helyzete" fontos - de nem annyira az, hogy ez a helyzet valójában mi is. Alcoff kiemeli, hogy a gyakorlatokat, amelyekben megszokottá válunk, mindenféle társadalmi különbség strukturálja. Elsősorban a faji és a nemi hovatartozást elemzi, de rámutat a társadalmi rétegekre, az iskolázottságra és a pénzügyi re- forrásokra is. Így a látszólag hétköznapi gyakorlatok különbözőek a különböző társadalmi pozíciókban élő alanyok számára. Simone de Beauvoir és Iris Marion Young munkásságára hivatkozva például mindenféle olyan dolgot felsorol, ami a nőkre jellemző:

Nemcsak dobálás és ülés, hanem állás, járás, futás, megszakításokkal járó és a témákat uraló beszélgetésminták, perceptuális orientációk, amelyek felöllelhetik a mellékes dolgokat, hogy észrevegyük a háztartási piszkot, a zaklatott gyerekeket, az unatkozó beszélgetőpartnereket és így tovább, valamint a saját érzelmi szubjektivitás belső élményét is. (Ibid.: 106)

A faji, valamint a kultúrák közötti és interszekcionális indexek esetében is hasonló listákat vezet (ibid.: 106 és köv.). Alcoff leírja, hogy a helyzeteket, tereket és személyeket dif-

különbözőképpen, attól függően, hogy milyen korábbi tapasztalataink vannak, milyen kultúrákban és jelentésekben mozogtunk. Másként lépünk be egy metróba férfiként vagy nőként, fehér bőrű vagy színes bőrű emberként. A kultúrákat és jelentéseket itt nem külsődleges struktúrákként kell felfogni. Ezek csak a kollektív gyakorlatokban maradnak fenn, és különösen abban, amit Alcoff "perceptuális gyakorlatnak" nevez (ibid.: 115).

Természetesen lehetséges reflektálni és foglalkozni a saját habituálissá válásunkkal és azokkal a gyakorlatokkal, amelyekben ez történik, de nem úgy, hogy teljesen átláthatóvá tesszük őket saját magunk számára. Nagyon is tudatosan cselekedhetünk annak tudatában, hogy észlelésünk és a cselekvés lehetséges formái mélyen összefonódnak a kontingens gyakorlatokkal. Mindazonáltal éppen ezek a gyakorlatok jelentik azt a kontextust, amelyben a jelentés és az észlelés létrejön. Ráadásul valamit megtapasztalni azt jelenti, hogy valahol lenni, és így nemcsak a megismerést teszi lehetővé, hanem a szubjektivitás megváltoztatásának lehetőségét is: "A megismerés egyfajta immanens elköteleződés, amelyben az ember saját énjét a világ [...] inkább bevonja, minthogy elkülönülten és fölötte állna." (Ibid.: 111.) Amikor tehát megpróbálunk elköteleződni saját helyzetünkkel, a gyakorlatok egyszerre alkotják ennek az elköteleződésnek a kontextusát és helyszínét. Ennek következtében a habituálódás csak a habituálódás egy másik formájának adhat teret:

A fenomenális világ folyamatosan visszahajlik önmagába, hozzáadva azt, ami korábban volt, és azt, ami még mindig a jelen pillanat háttérében marad; a múlt az, ami már meghaladta, mégis benne marad. Nincsenek teljes törések vagy teljes elválasztások, csak egy folyamatos szöveten belüli hajtások, amelyek terhesek a látens jelentéssel. (Ibid.: 110)

Ez azt is jelenti, hogy a szubjektivitások tartós megváltoztatása nem alapulhat egyéni próbálkozásokon. Sokkal inkább a gyakorlatokat, az ebből következő társadalmi viszonyokat kell megváltoztatni, hogy a szokássá válás és a szubjektivitások más formái jöjjenek létre:

A tapasztalatok számítanak, de jelentésük számunkra kétértelmű és dinamikus. Megtestesülünk, de nem redukálódunk fizikai meghatározottságokra, amelyeket a kultúrában és a történelemben elfoglalt helyünkön kívül létezőnek képzelünk el. Ez a beszámoló segít megragadni a társadalmi identitások dialektikáját, amelyben egyszerre vagyunk beillesztve a létező kategóriákba, és tesszük őket a sajátunkká. (Ibid.: 111)

A helyhez kötöttségnek és a társadalmi változások helyhez kötöttségének ez az elemzése következményekkel jár a Rouvroy által szorgalmazott politika szempontjából. Alcoff tagadja annak szükségességét, hogy az (emberi) lényt mindig többletként, vagy "a negáció vagy a f fény tiszta képességeként" (ibid.:

112) számon tartsuk. Még ha a politikáról szóló ilyen elképzeléseket mélyen inspirálja is a szubjektum kritikája, Alcoff azt állítja, hogy még mindig tartalmazzák a liberális beszámolókat inspiráló "dualizmus" maradványait, amelyek megpróbálják valahogyan elválasztani az egyént másoktól vagy a társadalomtól. A gyakorlatokban való habituális elhelyezkedés azonban

nem csupán az elnyomó és meghatározó identitások része - bár Alcoff vitájában ezek állnak az előtérben. Minden szubjektivitás részei, beleértve azokat is, amelyekkel azonosulunk, amelyekben örömet, barátságot, szolidaritást, szerencsét találunk. Következésképpen ezek eléréséhez nem kell túllépnünk a szituáltságon, csak a helyzetet kell megváltoztatnunk. Alcoff szavaival élve: "Az erkölcsi cselekvőképesség, a szubjektivitás és az érvelési képességek bizonyos típusú társadalmi hálózatokon belül válnak lehetővé. Nincs amorf szubsztancia vagy tiszta képesség, amely érintetlenül heverne az alatt a réteg alatt, ahol az identitás társadalmi konstrukciói érvényesülnek". (Ibid.)

5. Helyhez kötött szubjektumok, mesterséges intelligencia és politika

Alcoff maga nem tárgyalja a médiát és a technológiákat. Gondolkodását azonban mélyen inspirálja Merleau-Ponty fenomenológiája, amelynek középpontjában a tapasztalat mediatizált struktúrája áll - amit a vak emberbotjának hírhedt példája mutat be (Merleau-Ponty 1962: 152, lásd még Alcoff 2006: 188). Alcoff gondolata tehát könnyen kiegészíthető a médiatechnológiára vonatkozó szükséges reflexiókkal.

Christoph Ernst az interfészekről szóló értekezésében rámutat, hogy a digitális technológiával való interakció az interfészekeken keresztül egy szituált szubjektumot feltételez, beleértve a testet is (Ernst 2017: 100). Az interfészek csak azért működnek, mert képesek megszólítani az implicit tudást, amely gyakorlatokban gyökerezik, és így társadalmi szabályok által strukturált (ibid.: 102). Az interfész kutatás és -tervezés még ezt is tudatosan igyekszik kezelni azzal, amit Ernst a kognitív tudományra utalva "fogalmi modellnek" nevez (ibid.). Bár ez magában hordozza a manipulatív próbálkozások lehetőségét, ez önmagában nem manipuláció, hanem szükségszerűség ahhoz, hogy egy interfész működjön, vagyis hogy igazságot tegyen annak a ténynek, hogy az interfészek nem egyszerűen általános emberi lényekkel, hanem konkrétan szituált szubjektumokkal lépnek kölcsönhatásba.

Ernst az interfészeket tárgyalja, nem pedig az architektúrák vagy környezetek absztraktabb beállítását, amelyet Rouvroy hangsúlyoz, és amely "reflexszerű válaszokat kiváltó ingerek és jelek" révén működik. Ha azonban éppen ez a hatékonyság a technológiák de- fináló tényezője az algoritmikus kormányzásban, akkor az interfészekről nem eltérő módon kell kapcsolódnunk a habitualizált szubjektivitásokhoz (lásd még Distelmeyer 2017). Így, még ha ezek a technológiák nem is céloznak meg egy normakészletet és az ebből következő szubjektivitást, akkor is összefonódnak a szituált szubjektumokkal.

Ezt erősíti meg Chun megfigyelése is, miszerint a prediktív analitika a megszokott gyakorlatokhoz kötődik, amit fentebb már idéztem. Alcoff elméletét felhasználva most visszatérünk arra a pontra, hogy nem maga a szokássá válás a probléma. Az, hogy sok minden történik tudatos és szokásszintű szinten, nem

jelenti azt, hogy az AI alkalmazásai determinisztikusan hatnak ránk. Inkább az érzékelés és a cselekvés struktúráival lépnek kölcsönhatásba, amelyeket az algoritmusok bizonyára befolyásolhatnak, de amelyeket egy előre kialakított mélység is jellemez, amely a korábbi tapasztalatokból ered. Ez változhat

az algoritmikus kormányzás sokféleképpen előre látható eredménye, ami súrlódásmentes, szinte észrevétlen befolyást eredményez, valamint a Chun által javasolt módon a saját helyzet megélésének, elfogadásának és potenciális fejlődésének lehetőségeit - de sokféle feszültséget, sérülést és ellenállást is. Ez itt a fő szempont. A szubjektumok konkrétan elhelyezkedő szubjektumok, és az algoritmikus kormányzás, különösen azért, mert egy su- pra-szubjektív szinten működik, egészen másképp kapcsolódik a szubjektivitás különböző formáihoz.

Ez már az érzékelés szintjén elkezdődik: Például azok az uniós polgárok, akiket nem ismernek fel a mesterséges intelligencia alapú arcfelismerést alkalmazó automatizált határelenőrző terminálok, ezt azonnal a technológia hibás működésének fogják tekinteni. A migránsok ezt fenyegető döntésként érzékelhetik. A környezetek még kevésbé kézzelfogható adaptációi is a szituált szubjektumokhoz kapcsolódnak. Pontosan ez az oka annak, hogy a mesterséges intelligencia alkalmazásai nem semlegesek. Nem csak azért, mert elfogult képzési adatokon alapulnak, hanem azért is, mert különböző módon kapcsolódnak a szubjektivitás különböző formáihoz. A hírek algoritmikus szűrése azért problematikus, mert jobban kapcsolódik bizonyos, gyűlölet és másság által strukturált szubjektumokhoz és közösségekhez, mint más formákhoz. A visszaesés előrejelzése egy olyan biztonsági rendszert bővít ki és folytat tovább, amely faji megkülönböztetésen alapul. John Cheney-Lippold kimutatta, hogy az árucikkek algoritmikus szelekciója, amely egy olyan gépi tanulási rendszeren alapul, amely megpróbálja megjósolni a felhasználó nemét, jobban kapcsolódik a heteroszexuális, árucikké tett nemek formáihoz, mint másokhoz (Cheney-Lippold 2011).

A mesterséges intelligencia számos olyan technológiát eredményezett, amelyek fokozták a technológiák hatékonyságát abban az értelemben, hogy közvetlenül befolyásolják azt, ahogyan érzékeljük és cselekszünk a világban. Ez a hatás azonban nem jelenti a szubjektivitás általános elvesztését. Inkább a szubjektivitás különböző formáit alakítja át különböző módon. A kritika normális forrása tehát nem az algoritmikus irányítás alatt álló új szubjektivitás és az ezen valahogy túlmutató szubjektivitás közötti különbségben rejlik. A kritika forrása sokkal inkább a szubjektum-pozíciók között már meglévő különbségekben és azokban a sokféle módokban rejlik, ahogyan ezek a technológia révén eltolódnak. Chun javaslata, hogy "óvatosan fogadjuk el" ezt a helyzetet, lehet az egyik módja annak, hogy ilyen szituált módon próbáljunk meg jobbra forduló változást elérni. Más lehetőségek azonban a szituált alanyi pozíció elutasításában rejlenek, ami magában foglalhatja a magánélet védelmére vonatkozó követeléseket, valamint bizonyos technológiák használatának teljes beszüntetését. Ezekhez a követelésekhez társadalmi-technikai indexre lesz szükség. Vagyis nem a független szubjektum-pozícióhoz való visszatérés követelését jelentik, ahogyan azt a Chun és Rouvroy által egyaránt kritizált liberális irányzatok szeretnék. Mégis, a magánélet védelme, az adatok elrejtése, az automatizált elemzésbe való bevonás visszautasítása szükséges erőforrás lehet ahhoz, hogy jobb és

életképesebb helyzeteket találjunk azoknak a személyeknek, akiknek szubjektum-helyzete a mesterséges intelligencia bántó, visszaélő, jogfosztó alkalmazásaiba keveredik.

Ismétlem, a kritika kihívása nem a helyzettől való menekülés, hanem a változás. a helyzetet. Alcoff reflexiója azt mutatja, hogy ez mindig egy másokkal való szituáltság lesz; és ahogy elméleti szemléletének módosításával kapcsolatban kifejtettem, ez mindig is

a technológiával való szituáltság. Ez tehát végső soron politikai és demokratikus kihívást jelent. A mi helyzetünk mindig már eleve másokhoz kapcsolódik. A mesterséges intelligencia alkalmazásai ezt nagyon világossá teszik: a relációs adatokra összpontosítanak, és mint adatvezérelt technológiáknak csak aggregált szinten van értelmük.

Ugyanakkor, ahogy Alcoff mutatja, lehetetlen teljes mértékben tükrözni ezt a szituativitást és relationalitást. Ez nem egy rendszer vagy környezet, hanem sok különböző "múltbeli cél/elmúltbeli én/elmúltbeli tapasztalat" behatolása, ahogy Chun írja. Ez az egyes alanyok számára sok különbséget teremt az észlelésben és a cselekvési lehetőségekben. Így az ismeretelméleti és hatalmi kérdések összeolvadnak. Ebben az értelemben a politikai kihívások először is abban állnak, hogy megismerjük mások helyzetét, azt, hogy a technológiák hogyan kapcsolódnak a szubjektum pozíciójához. Második lépésben ezeket a pozíciókat össze kell egyeztetni, hogy a technológia új és jobb konfigurációját érnék el. Ennek demokratikus megoldásnak kell lennie, nem abban az értelemben, hogy kompromisszumot kell találni a már létező érdekek között, hanem abban az értelemben, hogy a szubjektumok mindig egy kapcsolódó, társadalmi-technikai helyzetű pluralitást alkotnak.

Hivatkozások

- Ahmed, Sara (2014): *Az érzelmek kulturális politikája*. Edinburgh University Press.
- Alcoff, Linda Martín (2006): *Látható identitások: faj, nem és én*. New York: Oxford University Press.
- Angwin, Julia; Larson, Jeff (2016): "MachineBias". *ProPublica*. Retrieved on 13.06.2018 (<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>).
- Aradau, Claudia; Blanke, Tobias (2015): "A (Big) Data-biztonság összeszerelése: Tudás és kritika". In: *Big Data & Society* 2 (2), pp. 1-12.
- Campolo, Alex; Sanfilippo, Madelyn; Whittaker, Meredith; et al. (2017): *AINow 2017 Jelentés*. New York: AI Now Institute.
- Cheney-Lippold, John (2011): "Egy új algoritmikus identitás: Lány biopolitika és az ellenőrzés modulációja". In: *Theory, Culture & Society*. 28 (6), pp. 164-181.
- Chun, Wendy Hui Kyong (2008): "A forráskeresésről", avagy a kód mint fétis. In: *Configurations*. 16 (3), pp. 299-324.
- Chun, Wendy Hui Kyong (2018): "Queering Homophily". In: Apprigh, Clemens; Cramer, Florian; Chun, Wendy Hui Kyong; et al. (eds.) *Pattern Discrimination*. Lüneburg: Meson pp. 59-98.
- Chun, Wendy Hui Kyong (2016): *A szokások új médiája*. Cambridge, MA: The MIT Press.
- De Francisci Morales, Gianmarco; Gionis, Aristides; Lucchese, Claudio (2012): "A csevegéstől a szalagcímekig: a valós idejű internet használása a személyre szabott

- hírajánlás". In: *ACM nemzetközi konferencián. a webes keresésről és adatbányászatról*, pp. 153-162.
- Distelmeyer, Jan (2017): "AN/LEITEN - Implikationen und Zwecke der Computerisierung". In: *Navigationen* 12 (7), pp. 99-116.
- Ernst, Christoph (2017): "Implizites Wissen, Kognition und die Praxistheorie des Interfaces". In: *Navigationen* (127), pp. 99-116.
- Hayles, N. Katherine (1999): *Hogyan lettünk poszthumánok*. Chicago: Chicago: University of Chicago Press.
- Heilmann, Till A. (2015): "Till Tilmann: Datenarbeit im 'Capture'-Kapitalismus: Datenarbeit im 'Capture'-Kapitalismus. Zur Ausweitung der Verwertungszone im Zeitalter informatischer Überwachung". In: *Zeitschrift für Medienwissenschaft*. (2), pp. 35-48.
- Kitchin, Rob (2014): "Big Data, új episztemológiák és paradigmaváltások". In: *Big Data & Society* (11), pp. 1-12.
- Kitchin, Rob (2017): "Az algoritmusokról való kritikus gondolkodás és az algoritmusok kutatása". In: *Information, Communication & Society*. 20 (1), pp. 14-29.
- Lyon, David (2014): "Snowden és a Big Data: A "Big Data: Lehetőségek, következmények, kritika". In: *Big Data & Society*. 1 (2), pp. 1-13.
- Matzner, Tobias (2016): "A Big Data performativitása a megfigyelésben". In: *Surveillance & Society*. 14 (2), pp. 197-210.
- Matzner, Tobias (2017): "A fegyelmezés és büntetés és napjainkban". In: *Foucault Studies* 23 , pp. 27-. 45.
- Matzner, Tobias (2019): "Az ember halott - éljen az algoritmus! Human-Algorithmic Ensembles and Liberal Subjectivity". In: *Theory, Culture & Society*. Online First.
- Matzner, Tobias (2014): "Miért nem elég a magánélet védelme az "ubiquitous computing" és a "big data" kontextusában". In: *Journal of Information, Communication and Ethics in Society* 12 (2), pp. 93-106.
- Merleau-Ponty, Maurice (1962[1945]): Maurice-Pauley-Pauley: *Az érzékelés fenomenológiája*. London: London: Routledge.
- O'Neil, Cathy (2017): *Weapons of math destruction: how big data increases inequality and threatens democracy*. New York: B/D/W/Y Broadway Books.
- Pariser, Eli (2011): *The filter bubble: what the Internet is hiding from you*. London: London: Viking.
- Pasquale, Frank (2015): *A fekete doboz társadalom: a pénzt irányító titkos algoritmusok és információ*. Cambridge: Harvard University Press.
- Reckwitz, Andreas (2002): "A társadalmi gyakorlatok elmélete felé: a kulturalista elméletalkotás fejlődése". In: *European journal of social theory* 5 (2), pp. 243-263.
- Rouvroy, Antoinette (2017): "Gouverner hors les normes: la gouvernementalité algorithmique". In: *Lacan Quotidien*. 733.

- Rouvroy, Antoinette (2013): "A kritika vége(i): adat-viselkedés vs. adekvátság." In: Mireille Hildebrandt, Katja de Vries (szerk.) *Privacy, Due Process and the Computational Turn-The Philosophy of Law Meets the Philosophy of Technology*. London: Routledge pp. 143-167.
- Srnicek, Nick (2016): *Platform capitalism*. Cambridge: Polity.
- Sudmann, Andreas (2018): "Szenarien des Postdigitalen. In: Engemann, Christoph; Sudmann, Andreas (szerk.) *Machine Learning-Medien, Infrastrukturen und Technologien der Künstlichen Intelligenz*. Bielefeld: transcript, pp. 55-74.
- Zheng, Guanjie; Zhang, Fuzheng; Zheng, Zihan; et al. (2018): "DRN: A Deep Reinforcement Learning Framework for News Recommendation". In: *Proceedings of the 2018 World Wide Web Conference (WWW '18)*, pp. 167-176.
- Zuboff, Shoshana (2019): *A felügyeleti kapitalizmus kora: harc a jövőért a hatalom új határai*. London: Profile Books.
- Zuiderveen Borgesius, Frederik; Trilling, Damian; Moeller, Judith; et al. (2016): "Kell-e aggódnunk a szűrőbuborékok miatt?". In: *Internet Policy Review* 5 (1).

A mélytanulás kormányzati jellege

A másik fekete doboz

Jonathan Roberge/Kevin Morin/Marius Senneville

Bevezetés

Frank Pasquale 2016-ban megjelent *The Black Box Society* című könyve mára mérföldkőnek számít a joggal kapcsolatos tudományágakban, a társadalomtudományokban és azon túl is. Maga a téma, a digitalizáció és az internet invazivitása rendkívül fontos. A könyv feltárja az üzemi titoktartás digitális platformok általi felforgatását és a felhasználók magánadataihoz való széles körű hozzáférést. A Facebook, a Google és társai bitek és bájtoknyi információt gyűjtenek és aggregálnak, hogy hatalmas profilalkotási rendszereket hozzanak létre, amelyeknek a *módjáról* keveset tudunk. Problémás, hogy magánszereplőként olyan masszív adatokra és "tudásra" tesznek szert a társadalomról és az egyéni viselkedésünkről, amire mi nem. Ehhez az összetett és minden bizonnyal kellemetlen valósághoz Pasquale elemzői szigora és kifinomultsága értékes meglátásokkal járul hozzá a lehetséges szabályozásokkal és az intelligensebb polgárság kialakításának lehetőségével kapcsolatban. Mi azonban azt szeretnénk állítani, hogy van egy másik, szélesebb körű és talán kulturális ok is, amiért *A fekete doboz társadalma* olyan nagy figyelmet keltett - ezzel is magyarázva a könyv sikerét. A rejtett, hálózatba kapcsolt entitás képe felidéz néhány általános félelmet. A "kontroll elvesztésének" érzését érzékelteti a legújabb au- tomációs folyamatokkal szemben. Egy ilyen algoritmikus fekete doboz, más szóval, a "tudott-ismeretlen", azaz az általunk felismert, többnyire rejtett tudástermelési formával kapcsolatos szorongást érinti. A fekete doboz képe kiábrándító; ebben rejlik az erőssége *és egyben* a gyengesége is.

Bár többnyire összhangban van Pasquale azon törekvésével, hogy megfejtse az átláthatatlanságot a mi adatvezérelt világban, az is jelentősnek tűnik, hogy megkérdőjelezzük a fekete doboz, mint heurisztikus, ha nem is holisztikus kép határait. Olyan tudósok, mint Geiger (2017), Sudmann (2018), Burrell (2016) vagy Bucher (2016) már feltárták ezt a területet. Utóbbi például azzal érvelt, hogy "az algoritmusok fekete dobozként való széles körben elterjedt felfogása inkább akadályozza, mint ösztönzi a kutatást", megjegyezve, hogy a fogalmat "túláságosan is könnyen használják" (84). Ezt követően az algoritmusok és algoritmikus rendszerek kritikus vizsgálatára szólít fel egy háromlépéses módszer segítségével: i) "ne féljünk a fekete doboztól!"; ii) "ne várjuk, hogy a megoldás benne van" és iii) "vegyük figyelembe a doboz dobozolást". Míg a

az első lépés úgy is felfogható, hogy az egész folyamatot magában foglalja, a többi lépés pedig úgy is felfogható, hogy egymást kiegészítő párost alkotnak, együtt vizsgálva a dobozon belüli és a dobozon kívüli oldalt. Ráadásul egy ilyen megközelítés különösen alkalmas a közelmúltban a mélytanulási algoritmusok felé történő elmozdulás elemzésére, de a különböző gépi tanulási technikák és mindazok elemzésére is, amelyeket manapság mesterséges intelligenciának neveznek. Hogyan és mennyiben különböznek ezek az algoritmusok gyakorlatilag és szimbolikusan a korábbi ún. Mit vonnának maguk után az átláthatatlanság, a többértelműség és a homályosság szempontjából? Hol, mit és kit kellene megnéznünk, hogy kritikai éleslátást és robusztus értelmezést fejlesszünk ki? Ezek a kérdések, ha egyszer dokkolunk Bucher lépéseihez, útmutatóként szolgálhatnak e fejezetben.

Közelebbről megvizsgálva, a "nem várjuk el, hogy a megoldás belül legyen" gondolata tökéletesen illeszkedik a mélytanulási algoritmusok mélyen gyökerező logikájához és történelmi fejlődéséhez. A "Szezámot kinyitni!" olyan feladat, amelyet nem lehet programozni, vagy ami azt illeti, nem lehet "megtanulni". A biológiai inspiráció és a terület történeti fejlődésének romantikus-teleologikus beszámolóí (Rosenblatt, 1958; Hinton & al. 2006) ellenére tény, hogy ami "tanulásnak" áll, az valójában adaptáció és önfejlesztés. A matematikai struktúra önmagát módosítja, miközben interakcióba lép a külvilágból érkező adatárammal és megbirkózik vele (Litvinski 2018; Lheureux & al. 2017). A backpropagáció, a rekurzív hurkok és más finomságok így nemcsak tükrözik, hanem életbe is léptetik a flux valóságát. Egy másik módja az ilyen bizonytalanságok szemlélésének az AI klasszikusabb szimbolikus megközelítése és a mai konnekcionista vagy neokonnekcionista elmozdulás közötti ellentmondás (Cardon & al. 2018). Míg az első a dedukcióra, az explicit modellezésre, az absztrakt szabályokra és a programozható nyelvekre támaszkodott egy logikai és formális érvelési mód megteremtése érdekében, addig a második az indukcióra épül, ahol a kapcsolt hipotézisek és közelítések "opcionált" észleléseket és előrejelzéseket hoznak létre arról, hogy mi történik az adatokban, amennyiben az adatok a kiszámíthatóság jobb arányaira fordítanak (Mackenzie Sudmann 2017, 2018). A nem lineáris számítások rétegei így valamiféle "mély", de sekély architektúrát informálnak, amely nem feltétlenül képez megmagyarázhatatlan AI-t, hanem a megmagyarázhatóság határait tolja tovább. Ha nem is teljesen fekete, a jelenlegi mesterséges intelligencia doboza enyhén szólva nagyon szürke. Ez látható azokban a problémákban is, amelyekkel a tudósok mostanában szembesülnek a kis léptékű elméletek reprodukálhatóságával kapcsolatban, ahol a publikálás jelenlegi gyakorlata általában megakadályozza őket abban, hogy mind a forráskódot, mind a tréningadatbázist megosszák, vagy hogy foglalkozzanak a hiperparaméterek önkényes beállításával kapcsolatos veszélyekkel, vagy akár a tréningértékek generálásának folyamatában rejtlő elkerülhetetlen véletlenszerűséggel (Hutson 2018). Ennyi adat kezelése és masszírozása sosem könnyű feladat, különösen nem kísérleti környezetben, és még kevésbé a való

világban, ahol a "garbage in, garbage out" mondás továbbra is alaposan érvényes. Az elfogultsággal kapcsolatos többszörös problémák manapság ugyanabba a körbe tartoznak.

kategóriába tartozik, és egy utolsó példaként szolgálhatna itt, nevezetesen, hogy a doboz a maga definíciója szerint nem lehet a megoldás valami önmagánál nagyobb dologra.¹

A fent vázolt belső problémák talán össze sem hasonlíthatók azzal, ami Bucher "a dobozolás megfontolására" irányuló meglátásaival kapcsolatban felmerül. Valójában a társadalomtudományokban általában és különösen az STS-ben hosszú múltra tekint vissza az, hogy nagy figyelmet fordítanak mindarra, ami egy adott technológia körülvesz (Bijker & al. 2012). A mélytanulási algoritmusok konkrét esetében még fontosabb, hogy ne feledjük, hogy "nagyobb, sokkal összetettebb összeállításokba ágyazódnak", amelyek soha nem szűnnek meg befolyásolni alakjukat és tartalmukat (Gillespie 2014: 3). A kérdés tehát az, hogy hogyan lehet értelmet adni az ilyen formáló nyomásoknak és az általuk létrehozott átláthatatlanságnak. Egy adott kontextus vagy egy adott "ökológia" komplexitásáról van szó, mégis amellet kívánunk érvelni, hogy az ilyen dobozolást leginkább a *hálózat* megközelítésen keresztül lehet figyelembe venni. Ahogyan máshol is elhangzott, "[...] nem egy doboz van, hanem több doboz. Az algoritmusok átláthatatlansága pontosabban az átláthatatlanság különböző formáiban fejeződik ki, amelyek mindegyike sajátos módon az emberi és nem emberi szereplők sokaságának *közbeiktatásától* függ" (Roberge & Seyfert 2018: 2; Latour 1987). Először is nehéz nem elismerni az innováció e területén belüli intenzív munkamegosztást - ez a helyzet gyakran azt jelenti, hogy a fejlesztők úgy dolgoznak egy adathalmazon, hogy nem tudják teljesen, kinek, milyen célból és miért. Itt az ügynöki tevékenység sok kis kéz között oszlik meg. Másodszor, egy hálózatos megközelítés a mélytanulási eszközök és technikák tényleges megvalósítását inkább nyájassnak, mint sorban állónak tekintené, inkább egymáshoz igazodva, mint együttműködve. Ezt jól dokumentálta például az algoritmikus pénzügyekkel foglalkozó szakirodalom, ahol a versengő érdekeltek "algotrading" eszközöket vetnek be az at-tack vagy a védelmi manőverek megerősítésére (Seyfert 2018; Knorr-Cetina & Preda 2011; Castelle & al. 2016). Hogy a társadalomtudományok a jövőben jobban odafigyelnek-e majd mindezen szereplők "dobozolásra" irányuló erőfeszítéseinek együttes, pillangószerű hatásaira, azt még nem tudjuk. Tekintettel azonban a jelenség alulvizsgálatosságára, sürgősen nagyobb figyelmet kell fordítanunk azokra az irányítási, rendezési és döntési folyamatokra, amelyek alakítják azt, hogy a mélytanulási algoritmusok mire jutnak a való világban. Egyszerűbben fogalmazva, a mesterséges intelligencia politikai gazdaságtana ma az egyik legnagyobb és legátláthatatlanabb dobozunk. Ebben a fejezetben a folyamatban lévő vitához kívánunk hozzájárulni azzal, hogy elemezzük, mi a tétje a társadalmi-technikai kormányzás eme új formájának, azaz mik a feszültségek, küzdelmek, a tudás, a hatalom kooptálására irányuló erőfeszítések stb. A montreali mesterséges intelligencia központot esettanulmányként használva, és a 2016-2018 közötti időszakra vonatkozó

- I Az IBM nemrégiben bejelentette, hogy hozzáférést biztosít egy több mint kétmillió képet tartalmazó könyvtárhoz az arcfelismerés gyakorlásához, abban a reményben, hogy a nagyobb pontosság segíthet az előítéletek visszaszorításában. Az olyan nem kormányzati szervezetek, mint az *Amerikai Polgári Szabadságjogi Unió* ebben és más hasonló esetekben az az álláspontjuk, hogy a jobb arcfelismerés még mindig rossz hír a kisebbségek számára, akiket a legkülönbözőbb társadalmi helyzetekben diszkriminációval kell szembenéznük. Lásd Browne 2019.

etnográfiai vizsgálat² során arra fogunk összpontosítani, hogy az érdekelt felek hogyan alkalmaznak több stratégiát és erőforrást, beleértve a legitimitás szimbolikusan terhelt médiaműveleteken keresztül történő kiépítését. A miénk tehát empirikus, ugyanakkor hálózati megközelítésünk elméletileg megalapozott; ebben az értelemben reméljük, hogy válaszolni tudunk a kiemelkedő tudósok felhívásaira, hogy a kritikai gondolkodást *helyben végzett* tanulmányok révén fejlesszük (Kitchin 2014; Mackenzie 2018).

I. Kormányzatiság - mi van vele, és mit változtat a mélytanulás tanulmányozásán?

A mai politikai gazdaság növekvő komplexitását övező viták, valamint az, hogy ennek hogyan kellene a hatalom új értelmezésében megnyilvánulnia, nagymértékben érthetőbbé válnak, ha Michel Foucault kormányzati koncepcióját követjük (2004a, 2004b). A "kormányzás problémájának" ez a gyakorlatorientált elemzése lehetővé tette a tudósok számára, hogy újra a "szüntelen tranzakciókra összpontosítsanak, amelyek változó módon módosítják, elmozdítják, felborítják vagy alattomosan eltolják a finanszírozást, a beruházási módokat, a döntési központokat, az ellenőrzés formáit és típusait, a helyi és központi hatóságok közötti kapcsolatokat stb.". (Foucault, idézi Lascoumes 2004: 3; a mi fordításunk.) A francia filozófusnak valóban kulcsszerepe volt - legalábbis - e három logika azonosításában: i) a hatalom és a tudás elválaszthatatlanok egymástól, ii) hogyan vezetik be ezek a mobil és hálózatos dinamikát, és iii) hogyan teszi mindez lehetővé, hogy a hatalomról úgy gondolkodjunk, mint a technológiák által létrehozott és technológiák összességéként. Foucault-val az a nehézség, hogy soha nem írt megfelelően a digitálisról. A közelmúltban Mackenzie volt az, aki arra törekedett, hogy a koncepciót az általa "gépi tanulónak" nevezett, azaz a naiv Bayes-osztályozók, a döntési fák, a neurális hálózatok és egy sor más, a mesterséges intelligencia tág kategóriájába sorolható dolog tanulmányozására alkalmazza (2013; 2017; 2018). Mindez szerinte egy olyan "adatgyakorlatnak felel meg, amely az ember-gép viszonyok átrajzolásával újra konfigurálja a hatalom és a tudás helyi központjait" (2017: 9). Az, hogy a kutatás, a fejlesztés és a megvalósítás hogyan szerveződik; ki által, milyen céllal, milyen eszközökkel és diskurzusokon keresztül, más és más Londonban és a Szilícium-völgyben, vagy éppen Kínában és Kanadában. Hasonlóképpen, az, hogy a hatalmi viszonyokat és a hatalom elosztását hogyan alakítja specifikusan a szervezeti és intézményi struktúrája, részterületenként - pénzügy, katonaság, közlekedés stb. Mackenzie tehát nagyon hasznos azáltal, hogy ilyen ökoszisztémikus, ha nem is ökológiai szemléletet nyújt. Ugyanakkor könyvében fennáll annak a veszélye, hogy túlságosan is hangsúlyozza a technológia belső vizsgálatát,

2 Összesítjük a gépi tanulással foglalkozó szakemberekkel készített interjúkat¹², a tudományos újságírókkal készített további 4 interjúkat, valamint a helyi frankofón és angol nyelvű újságok és havi kiadványok több mint 400 cikkét.

és így csak részben képes "figyelembe venni a dobozolás" - hogy ismét Boucher utaljak.

Mi kellene ahhoz, hogy képesek legyünk olyan betekintést nyújtani, amely egyszerre lenne helyi és "építészeti" - azaz képes lenne bemutatni, hogyan történnek bizonyos konstrukciók és átalakulások kívülről befelé? Az egyik ilyen lehetőség a kormányzati jelleg és a ma "kormányzásnak" nevezett dolgok közötti különbség vizsgálata, valamint az, hogy valójában az utóbbi az előbbi témája. Montrealban és valószínűleg máshol is nagy lendületet kap a kormányzással kapcsolatos diskurzus, mivel maga a fogalom egyfajta üres jelzőként szolgál, amelynek taktikusan lehet jelentést adni. A kormányzás, akárcsak a haladás vagy az innováció, könnyen jelenti a "jó kormányzást", és a montreali központ építésében részt vevő számos érdekelt fél összemossa a kettőt, hogy a mélytanulás piacorientált fejlesztésének intézményi és állami támogatását erősítse, amelynek részleteit rövidesen bemutatjuk. Egyelőre elég, ha annyit mondunk, hogy a québeci metropoliszok mélytanulási kormányzásának ötlete arra törekszik, hogy nagyjából mindenkit "partnerként" bevonjon a kollektív önmenedzselés és a céltudatos társadalmi változás játékába. A játékban az van, amit olyan tudósok, mint Walters úgy azonosítottak, mint "[a] hangsúlyt az önirányító hálózatokra", amely nagymértékben támaszkodik "a kibernetika és a komplexitáselmélet képi világára" (2004: 29-30; lásd még Simard hasonló elméleti 1979 megközelítését Québecre alkalmazva). A hatalmat és a tekintélyt itt lehetővé tevő tényezőként fogják fel: lehetővé teszik az erőforrások körforgását, nem pedig azok korlátozását vagy korlátozását. Amint az alábbiakban világossá válik, minden, ami az etikával kapcsolatos - az ipar által támogatott partnerség az AI-ról vagy a montreali nyilatkozat az AI felelősségteljes fejlesztéséről -, az önszabályozás eszméje és annak sajátos módja, ahogyan az a törvényhozás *mellőzésére* irányuló laza, beteges törekvéssé válik, beszennyezi. A hatalom és a politika nem tűnt el, ami azt illeti; bár a kormányzás a legjobb fényben, mint könnyűsúlyú *kormányzás a távoból*, a lényeg az, hogy a kormányzás egy hatékony, bár kevésbé tanulmányozott formájaként jelenik meg.

Mi az, ami a montreali mélytanulási központot méltóvá teszi a sci-entific analízis? A válasz részben azzal a ténnyel függ össze, hogy Québec egy meglehetősen kis társadalom, amely jól fejlett, de még mindig az elit - társadalmi, politikai, gazdasági, kulturális stb. elit koncentrációja jellemzi. Ami azt a történelmi kontextust illeti, amelyben a tartomány a legfrissebb "mesterséges intelligencia ébredését" kezelte, fontos felidézni a *kanadai* CIFAR szerepét a mélytanulási kutatások támogatásában, még akkor is, amikor a technika még nagyon nem volt divatos (Hernandez 2014; Cardon & al. 2018; Enge- mann and Sudmann 2018). Az olyan sztárkutatók, mint Hinton (Torontói Egyetem), Bengio (Université de Montréal) korábbi tanítványai és kisebb mértékben³ LeCun (NYU és most Facebook) egyszerre örökösei és legjobb támogatói a ma már 125 millió kanadai dolláros pán-kanadai AI-stratégiának. Amikor például a beszélgetés a

- 3 LeCun az elmúlt években kevesebbet dolgozott Kanadában, és többet Franciaországban és az USA-ban, bár még mindig fontos médiavisszhangot kap.

a québeci frankofón médiaközegben kialakult egy folyamatban lévő "AI-forradalom", Bengio maga is a montreali AI-ökoszisztéma egyik vezetőjeként mutatkozott be, amely maga is a legkorszerűbb AI-innováció egyik világhírű központjaként mutatkozott be (lásd Bourgault 2017). Ezt a sajátos dinamikát tovább bizonyítja az a tény, hogy a neve a *La Presse* című montreali napilapban 2016 májusa és 2017 júliusa között megjelent 161, a mesterséges intelligencia fejlesztéseire összpontosító cikkből 126-ban szerepel (Bourgault 2017). A lényeg itt az, hogy a québeci AI-mezőt vizsgálva két összefüggés világosan megjelenik: egyrészt a forradalom kialakuló retorikája és egy karizmatikus vezető felemelkedése között; másrészt a mélytanulás és az AI körüli hangsúlyos hype és a helyi szereplők képessége között, hogy gyorsan mozgásba hozzák a vonatkozó intézményeket. "A hype-nak alacsony az informatív tartalma" - jegyzi meg helyesen a tudós Guice - "de közvetlenül az információ relevanciáját mondja ki egy társadalmi kontextusban" (1999: 85). A montreali központ megerősítése érdekében a korábbi québeci gazdasági és innovációs miniszter, Anglade megjegyezte, hogy kormánya nem fogja "szórni" az állami beruházásokat (Rettino-Parazelli 2017). Ez vezetett először egy tanácsadó bizottság, majd egy kezdetben 100 millió C\$-os költségvetéssel felszerelt AI-klaszter létrehozásához. A klaszter két legérdekesebb ténye, hogy a pénzeszközök 80%-át a Bengio által irányított, az Université de Montréal vezette MILA (Montreal Institute for Learning Algorithms) számára fordította, miközben mindezt Breton-Université de Montréal dékánja - és az ismert üzletember Boivin - vezette, aki néhány hónappal később a MILA igazgatótanácsának vezetője is lett (lásd IA.Québec 2018). Ez azt sugallja, hogy ebben a konkrét kontextusban és ebben a rövid időszakban a jó kormányzás a hatékonyságot jelentette, míg egy szélesebb, reflexívebb és kritikusabb perspektíva inkább azt kérdezte volna ki, hogy mit jelent a jó kormányzás az elitek körforgása szempontjából, és hogy a hatékonyság maximalizálására irányuló erőfeszítéseknek miért kell még mindig igazolniuk és legitimálniuk magukat legalább a megfelelően lefolytatott adminisztratív folyamatok látszatán keresztül.

A montreali esetben a "dobozolás" egy másik módja a következőképpen néz ki nézzük meg a fent említett, a mesterséges intelligencia felelősségteljes fejlesztéséről szóló montreali nyilatkozat elindítására irányuló erőfeszítések és a mesterséges intelligencia társadalmi hatásainak nemzetközi 2018, megfigyelőközpontjának késői létrehozása közötti kapcsolatot. Mindkettő, amelyet a kormány és a legfontosabb tudományos intézmények teljes mértékben támogatnak, a nyilvánossággal, a különböző érdekelt felekkel stb. együttesen kialakított "tudás" episztemológiai álláspontjára hivatkozik, amely a gyakorlatban képlékeny, ha nem is sekélyes jelzőként szolgál. A Nyilatkozat például tíz elvből álló listát javasol, amelyek mind általánosabbak és elvontabbak a másiknál, némelyikük túlságosan naiv vagy ellentmondásban áll a mélytanulás jelenlegi gazdasági valóságával.

A 6.2 például kimondja, hogy "az AI fejlesztésének hozzá kell járulnia a csoportok és emberek közötti, hatalmi, vagyoni vagy tudásbeli különbségeken alapuló uralmi viszonyok felszámolásához" (IA felelős 2017). A Megfigyelőközpont a maga részéről még csak születőben van, de már a maga

alkotmányában is jelzi a társadalomtudományok gyenge megértését.

a társadalmi hatások tanulmányozásában betöltött szerepét, például több tag érkezett az informatikából, mint a szociológiából és a kommunikációkutatásból. Fontos, hogy a Megfigyelőközpont és a Nyilatkozat közös vonása a kormányzás fent bemutatott kibernetikai szemlélete. Egyrészt a tudástermelés irányítása eltérli a fékek és ellensúlyok vagy a karhatalmi szabályozási elvek minden fogalmát, amelyek a modernitás eszméjének központi elemei. Másrészt úgy tűnik, hogy minden jelenlegi erényjelző törekvés, beleértve a Nyilatkozatot, a Partnerség a mesterséges intelligenciáért és hasonlókat, úgy jelenik meg, mint amit Wagner "menekülés a szabályozás elől" (2018).⁴ Mindent egybevetve, a québeci kormányzat a montreali központ fejlesztésében való részvétele nem akadályok és gátak felállítását jelenti, hanem inkább azt, hogy a "pozitívnak" tartott dolgok áramlását, nevezetesen minden olyan fordulatot, amely a kormány vagy az érdekeltek részéről valamilyen cselekvési formát mutat, elővezeti és elősegíti, tudva, hogy annak legitimitása, hogy ki ad, visszahat arra, hogy ki kap, és fordítva.

II. A "Triple Helix" remix és a nyílt tudomány szerepe

Ezen a ponton csábító lenne kijelenteni, hogy számos szereplő kezdeményezései ellenére, beleértve a québeci kormány jelentős gesztusait is, még mindig "minden a régiben". Ez azonban legalább két szempontból félrevezető lenne. Először is, bár pontos azt mondani, hogy a Klaszter, a Nyilatkozat és a Megfigyelőközpont mind részt vesznek egy bizonyos közfelfogás kialakításában mindenről, ami a mesterséges intelligenciával kapcsolatos, nem lehet ezt egy olyan ideológiának megfeleltetni, amely elrejtene bármiféle csupasz igazságot.⁵ Más szóval, kritikusként lenni annyi, mint megkérdőjelezni, hogyan készül a doboz, nem pedig felgyújtani. Másodszor, a "szokásos üzletmenet" kifejezés aláássa, hogy a mélytanulás és a kapcsolódó mesterséges intelligencia technikák alkalmazása mennyire megváltoztatja a tartomány hatalmi-tudásbeli topográfiáját, nevezetesen azt a kulcsfontosságú szerepet, amelyet az egyetemeknek kell betölteniük. A kormányzati megközelítésnek ezért figyelnie kell a struktúráltságra, valamint az itt felmerülő feszültségekre - vagyis azokra a történelmi és földrajzi finomságokra, amelyek a québeci felsőoktatást egyszerre teszik észak-amerikai, de a francia univerzalista megközelítés által is mélyen befolyásolt dologgá,

4 Sokan valóban észrevették, hogy az iparági szereplők egyre gyakoribb felhívásai az "etikus mesterséges intelligenciára" gyakran korrelálnak a "túlságosan kényszerítő" kormányzati szabályozás elleni folyamatos kampányokkal; lásd Wagner, 2018 és Greene & al., 2019. Úgy tűnik, hogy e téma egyik legújabb variációja az ipar által támogatott szabályozás (Simonite, 2019); már a "szabályozási foglyul ejtés" vádjai is elhangzott (Biddle, 2019).

5 Egy ilyen hagyomány a korai Habermasban talál emblemikus alakot, amikor például azt mondta, hogy "[...] [az ideológiák] a modern tudomány köntösében megjelenve és az ideológiakritikából levezetve igazolásukat, a hatalom hagyományos legitimitációit váltják fel". Az ideológiák egyidősek az ideológiakritikával" (Habermas 99).1971,

és hogy a 60-as és 70-es évektől kezdődően hogyan tett jelentős lépéseket a hozzáférés demokratizálása felé.⁶ Érdekes módon az Université de Montréal és a McGill Egyetem, mindkettő a MILA élvonalában, történelmileg "elit" iskoláknak számítanak, miközben még mindig nagy állami támogatást élveznek. Maga a MILA nemcsak azért fontos, mert a kormány által támogatott klaszter pénzének nagy részét vonzotta, hanem azért is, mert megtestesíti a szakirodalom által hármasspirálnak nevezett elmozdulást, sőt, finomodását - az innováció olyan sematikus modelljét, amelyben a vállalati szereplők összefonódnak az egyetemi és kormányzati szereplőkkel (Etzkowitz & Leidesdorff 2000) - egy "négyes spirál"-ra, amelyben a start-upok is kulcsfontosságú stratégiai partnereknek számítanak. A kicsiket és a nagyokat összekeverik, így Montrealban nem ritka, hogy az olyan nemzetközi vállalatokat, mint a Microsoft, a helyi Maluubával, a Facebookot a feltörekvő FAIR-MTL-nek vagy a Google-t az embrionális DeepMindnek tekintik - a média pedig még a legkisebb befektetéseket is ünnepli.⁷ Mindez egy olyan ökológiai mentalitásban vesz részt, amely szerintünk összemosódik a szimbolikus - a központ természeténél fogva pozitív - és a gyakorlati - a Hoffman és mások által "akadémiai kapitalizmusnak" nevezett legújabb trendnek köszönhetően (2017; lásd még Slaughter and Rhodes 2010). A valóság viszont kevésbé felel meg a québeci felsőoktatási rendszerre gyakorolt korai francia hatásnak, mint inkább a "Szilícium-völgy-izáció" vagy a "Stanford-izáció" módjának, Saltertől kölcsönzött kifejezésekkel (2018).

Az összes mesterséges intelligencia-fejlesztés közös jellemzője, hogy a "nyílt tudomány" egy sajátos etikája vagy modellje vezérli és attól elválaszthatatlan (Leonelli 2013; Mirowski 2018). A kutatók az információk megosztását *prima facie* önmagában is előrelépésnek tekintik; a felfedezéseket és az innovációkat az egész "közösség" hivatott felkarolni és hasznára fordítani, ami így egy ökológiai és kibernetikai mentalitású, amely ismét egy bizonyos kaliforniai "lázadásban" gyökerezik.⁸ Ma ezek a normák igen hatékonynak bizonyulnak, különösen a következő három dimenzió tekintetében. Először is, az egyetemi karrier vagy a magán K+F karrier választásának kötelezettsége kevésbé válik kérdéssé, ha az ember szabadon publikálhat, ami ma már megengedett, ha nem is ösztönözve van a legtöbb alap kutatáshoz kötődő ipari laboratóriumban - sőt, nem ritka, hogy a Facebook vagy a Deep Mind tudósai egyetemhez kötődő kutatókkal együtt írnak publikációkat. A mesterséges intelligencia területén tapasztalható képzett munkaerőhiány ismeretében az ilyen nyílt tudományos gyakorlatokat eszközként alkalmazzák.

6 Az Université du Québec egésze alatt működő tíz, a terület egészén hálózatba tömörült intézmény e tekintetben emblemikus.

7 Lásd például a helyi ökoszisztémát érintő beruházások összefoglalóját a 2017 Mathysban. 2017.

8 Itt arra szeretnénk utalni, amit Saxenian (1994) és mások úgy jellemeztek, mint a Szilícium-völgy akadémiai és ipari szereplőinek jellegzetesen innovatív módját, hogy lenyűgöző sebességgel új szervezeti formákat hozzanak létre. A különböző szakmai kategóriákhoz tartozó, de közös

kulturális háttérrel rendelkező szereplők közötti együttműködési kapcsolatok létrehozásának ez a sajátos módja sokak szerint a technológiai innováció vezetésére vonatkozó kiváltságos felfogás élére állt napjainkban - lásd még Storper & al. 2015.

a "tehetségek vonzásáért folytatott háborúban" való eligazodáshoz szükséges repertoár részeként (Hernandez & King 2016; Metz 2016a; 2017). Másodszer, az emberek és az ötletek korlátlan áramlásának lehetővé kell tennie, hogy a vállalatok nyomon követhessék az egyetemi kutatások legjavát. A gyakornokok, a támogatások és a kisebb-nagyobb filantróp adományok hozzájárulnak ahhoz, hogy biztosítsák a hozzáférést az informatikai laboratóriumokhoz, és ott érjék el a kutatókat, ahol vannak. Egy olyan város számára, mint Montreal, ez nagyon hasznosnak bizonyult, még ha ebből a decentralizációból és nyitottságból nem is lehet azt a következtetést levonni, hogy a központja "plaque-tournante" - elvégre mások, mint Párizs, Szingapúr, Pittsburg stb. is profitáltak belőle. Harmadszor, ez a nyitottság nemcsak földrajzi, hanem időbeli is, mivel a különböző spirálok, különösen a vállalatok és az egyetemi laboratóriumok közötti alkalmazkodás többé-kevésbé valós időben történik. A kutatás üteme itt ugyanolyan fontos, mint a kereskedelmi forgalom üteme, amely egy algoritmikus architektúrát API-vá, egy innovációt ajánlási rendszerrel stb. alakít át. Miközben a nyitottság a lehetőségek ablakaiba és a jó időzítés a vállalatok közötti éles versenybe torkollik, különösen fontos megérteni, hogy az egész modellt fenntartó logika valójában a "stratégiai nyitottság" logikája (Ananny & Crawford 2018). Az egyetemeken előállított tudásról kiderül, hogy egy új és kevésbé vizsgált értelemben "nyitott az üzlet számára", különösen az olyan kontextusokban, mint Québec, jelentkező szélesebb körű következményei tekintetében.

Az a tény, hogy a mélytanulás és a kapcsolódó AI-technológiák egy jelentős A gazdagság, a presztízs és a hatalom elmozdulása számos és annál inkább empirikus példát talál, amelyekre néhány pillanat múlva rátérünk. Egyelőre azonban úgy tűnik, hogy a szükséges átmenet azt jelenti, hogy meg kell kérdőjelezni a négyes spirál és a nyílt tudomány nexus révén mozgásba hozott "kizsákmányolható episztemológia" (Levy & Johns 2016) szélesebb körű jelentőségét. E kutatás részeként egy sor interjú készült a montreali mesterséges intelligenciával foglalkozó személyekkel, akiknek többsége - két-három kritikusabb személyiség kivételével - nagyjából egybehangzó véleményt fogalmazott meg. Az első egy, az egészségügyben dolgozó informatikustól érkezett. Az ő kritikája a kutatásfinanszírozás quebec-i és kanadai átalakulásának strukturális elemeire mutat rá:

Ez dollármilliók koncentrációja, olyan, mintha egyetlen számra fogadnál a kaszinó rulettjén. Számos különböző típusú kutatást végeznek, nem mind a mélytanulás vagy a nagy adattörzsek [az AI-ban], de ezek is innovatívak - de nem ezekre fogadsz, csak a mélytanulásra fogadsz. Mindenkit ugyanabba az irányba terelsz, és elfelejted, hogy az innováció nem feltétlenül ugyanabba az irányba megy. Szükséged van arra is, hogy hagyj néhányat, hogy biztos legyél benne, hogy az újrakeresés a maga teljességében némileg diverzifikált. Ezt én fenyegetésnek látom. Ez mindent ugyanabba az irányba fog sodorni [...]. Valójában mindenki belerohan ebbe.

Ez a megjegyzés egyfajta közvetítő vagy utaló jelként szolgálhat arra, hogy hogyan és miért vesznek részt még a tudományos intézmények is abban a fajta önbeteljesítő jóslatban, amely a mély tanulást valósággá teszi. Az intézmények, a hype és a "stanfordizált" kutatásra gyakorolt nyomás ismét kéz a kézben járnak. Egy másik fontos példa lenne a CI-29 FAR kutatási tanszékek késői odaítélése néhány 2018, prominens Facebook munkatársnak, mint például Pineault (McGill) vagy Vincent (Université de Montréal) (CIFAR 2018). A kevésbé divatos kutatási áramlatok esetében ez természetesen olyan utat rajzol ki, amelyben a finanszírozáshoz való nehéz hozzáférés keveredne a hallgatókhoz való megerőltető hozzáférés megfelelőjével - legalább két másik, állami egyetemeken dolgozó informatikus a mintáinkból legalább két másik informatikus beszélt arról, hogy manapság alig van végzős hallgatójuk. Hosszabb távon nézve a helyzet valószínűleg csak még ciklikusabbá és hátrányosabbá válik.

Egy másik kapcsolódó kérdés, amely az itt érintett "kihasználható", sőt fegyverként használt episztemológia jelentésével kapcsolatban felmerül, az adatbázisok kezelésével kapcsolatos: kié, hogyan és milyen célokra adják ki őket. Az olyan nagy technológiai cégek számára, mint a Google, amely nemrégiben nyílt forráskódúvá tette a GPipe-ot, vagy a Microsoft, amely felvásárolta és most a Githubot üzemelteti, a gigantikus adatkönyvtárak egyszerre performatív és legitimáló szerepet töltenek be. F törvényeik és korlátaik alig vagy egyáltalán nem kerülnek napvilágra - például az a tény, hogy az ilyen vállalatok még mindig agresszívan követnek szabadalmakat (Simonite 2018) -, különösen, ha összehasonlítjuk az ökológiai és kibernetikai előnyökkel, amelyeket ezeknek a platformoknak tulajdonítanak, és amelyeket a médiában széles körben dicsérnek. Ezután mezzo vagy helyi szinten a dolgok már nagyobb kihívást jelentenek. A probléma az, hogy a képzésre szánt nyílt adatok nem egészen ugyanazok, mint a "valódi" adatok vagy az értéknövelt adatok. Például interjúalanyunk, aki mélytanulási alkalmazásokkal dolgozik az egészségügyben, ragaszkodott ahhoz, hogy veszélyes dinamika van kialakulóban, ahol az induló vállalkozások keresnek bármilyen jelentős alapot, amelyhez hozzáférhetnek a mélytanulási szolgáltatásokért cserébe, vagy legalábbis olyan szerződést kötnek, amely lehetővé teszi számukra, hogy adatokat osszanak meg egy harmadik féllel. Az olyan helyeken, mint Montreal, a profitszerzés azt jelenti, hogy olyan ügyfeleket - biztosítók, bankok, klinikák, biotechnológiák stb. - találnak, akik még nem szoktak hozzá a mélytanulási technikákhoz, egy olyan jogi környezetben, amely még mindig bizonytalan a magánélet védelmének legjobb módját illetően, vagy az esetleges jogsértések szabályozását illetően.⁹ Valószínűleg mégis a különböző egyetemi laboratóriumok mikroszintjén a legszembetűnőbb a különbség az adatokkal "rendelkezők" és "nem rendelkezők" között. Az olyan sztárkutatók, mint a montreali Bengio - vagy éppenséggel a torontói Hinton - a szoros kapcsolataik miatt vonzzák a finanszírozást és a hallgatókat.

⁹ Fontos párhuzamot kell vonni azzal, ahogyan a Theranos, a bukott unikornis, kihasználta a biotechnológiai ipar fontos szabályozási mozgásterét, hogy e feltörekvő terület egyik legnagyobb

(ha rövid életű) sikertörténetévé váljon. Az, hogy a növekvő befektetési körök fenntartása érdekében olyan korai támogatók, mint James Mattis tábornok, Larry Ellison Oracle-alapító, Rupert Murdoch médiamogul vagy Betsy DeVos jelenlegi oktatási miniszter hírnevére támaszkodhatott, pontosan a terület viszonylagos újdonságának és a megfelelő szabályozás hiányának eredményeként értendő (O'Brien 2018).

a Google és hasonlók. De mi a helyzet a terület többi, kevésbé ismert kutatójával? A probléma az, hogy ők szinte soha nem jutnak hozzá a széleskörű kereskedelmi alkalmazásokhoz szükséges adatokhoz. Hogy egy utolsó példát mondjak ebben a részben, csapatunk 2017 nyarán találkozott egy másik tudóssal, és beszélgettünk az általános közösségi érzésről, és arról, hogy mit jelent a nyílt tudomány a terület magán- és közszereplői számára a kommunikáció egyik módja. A válasza lakonikus volt: "Ez csak kamu. Ez csak kamu. Megosztják az algoritmust, de az adatokat nem, ezzel nem lehet mit kezdeni. [Ami a "nyitott" jelentését illeti], ez csak egy szó, mert nem tudom használni".

III. A mélytanulás újradefiniálja a magán-állami partnerséget

Ha azt mondjuk, hogy a mesterséges intelligencia politikai gazdaságtanának mai fejleményei elmosják a magán- és közoktatásnak tekintett területek közötti korábbi különbségeket, vagy hogy ez a felsőoktatás "stanfordizálódását" jelenti olyan helyeken, mint Québec, akkor nem szabad egy utópisztikus múlt iránti nosztalgiának engedni. Valóban leíró és agnosztikus megközelítésre van szükség ahhoz, hogy - ahogy Hoffman fogalmazott - "a kereskedelmi logikák bonyolult, finom és néha ellentmondásos módjai, amelyek az akadémiai kultúrában elterjedtek" (2017: 727). A lényeg az, hogy Montrealban és valószínűleg máshol is a kétértelműség önmagában a kormányozottság egyik formája. A gyengébb intézményi autonómiát több együttműködésre fordítják; a "hub", a "klaszterezés", az "ökoszisztéma" és hasonlók szemantikai régiójába tartozó divatszavakat ismételtetik és ünneplik, amit aztán nehéz megfejteni a public relations törekvésekből (lásd például 2018, Turkina). Ennek a fejlődésnek a fordulópontja volt a MILA januári áttelepülése 2019 Mile-Ex-be, egy posztindusztriális belvárosba Montrealban. A laboratórium áthelyezése akkor történt, amikor a "Québeci Mesterséges Intelligencia Intézet" státuszba emelkedett, és a Mile-Ex-i Cité de l'IA élére került, ahol több kis- és nagyvállalat is új létesítményt hozott létre ugyanabban az épületben vagy annak közvetlen környezetében. O Mile-Ex, az átalakított textilgyár, ahová a laboratórium beköltözött, már a feltörekvő startup, az Element AI, a Royal Bank of Canada Borealis nevű AI-ágazata, a francia katonai vállalkozó Thales AI-kutató részlege és a para-állami Adatértékesítési Intézet (IVADO) irodáit fogadta be, és a Microsoft Maluuba is közeli szomszédja (Bachand 2018; Dubuc 2018). Fontos, hogy a Cité létrehozásának ötlete a kormánystratégiájának negyedik pillérét képviseli minden téren - az ipari klaszter, a nyilatkozat és a megfigyelőközpont mellett -, de a munkahelyek, a beruházások, az irodaterületek négyzetméterei és hasonlók tekintetében mért súlya miatt azt lehet állítani, hogy ez a legfontosabb. Az ottani felelősök meglehetősen jól ismerik a köztes pozíciójukkal járó előnyöket. A médiának az áthelyezéssel kapcsolatban adott interjúkban ugyanis szívesen kérdezték a

következőket

további állami finanszírozásért: "A kutatókat Montrealba vonzani azzal, hogy azt mondjuk nekik, hogy már csak két évünk van a finanszírozásra, nem fog működni. Egy szélesebb körű, hosszabb távú elképzelés részévé kell válnunk. Tíz- és százmilliós nagyságrendben vagyunk" (Pisano idézi Rettino-Parazelli 2019).

Ami a gyakorlati, de nem hivatalos köz- és magánszféra keveredését illeti, Québecben talán nincs párja az Element AI-nak, a MILA igazgatója, Bengio által társalapított, gyorsan növekvő vállalatnak. Ahogyan azt egy médiakommentátor helyesen megfogalmazta, "az üzleti modelljét nem könnyű megérteni" (idézi a *Mercure* 2016); nemcsak hogy történelmi mennyiségű kockázati tőkét vonzott anélkül, hogy megfelelő termék lenne a piacon, hanem folyamatosan a közjó és a társadalmi előnyök etikai irányultságú diskurzusa alapján működik, miközben aktív szereplőként pozicionálja magát a logisztika, a biztosítás és a banki tevékenység meglehetősen hagyományos és profitorientált területein is (*The Economist* 2017). Úgy tűnik, maga Bengio mindkét szinten hajlandó játszani, mivel valódi erőfeszítéseket tesz a saját területének etikai és társadalmi szempontból történő fejlesztésére, miközben oroszlánrészt vállalata kereskedelmi sikerének, a nyilvánvalóan az első kanadai AI Unicorn egyikévé váló vállalatának oroszlánrészét dedikálja (George-Cosh 2018; Vara 2018). Ráadásul előadásokban, Power Pointokban és máshol néha összekeveri saját számos köz- és magánhovatartozását, ami most már emblematikus egy nagyobb problémára, nevezetesen arra, hogy a közszférában létrehozott érték és vagyon hajlamos eltávolodni a közszférától. Az Element AI lényege - és a kezdeti értékelésének egyik oka -, hogy a sztárakadémikusokhoz való hozzáférést kihasználva "üzleti megoldásokat" dolgoz ki a magánszektorbeli ügyfelei számára. A *Journal of Small Business & Entrepreneur* című folyóiratban megjelentek szerint a vállalatnak "több mint 20 világhírű mesterséges intelligencia tudósból álló oktatói hálózata van, akik Can- ada legjobb egyetemi laboratóriumaiban dolgoznak. Ezek a professzorok nemcsak a kutatással kapcsolatos munkát végzik a vállalat számára, hanem [...] értékes tanácsokat is adnak [...]. Ez az egyedülálló megállapodás hozzáférést biztosít az Element AI számára a legmodernebb kutatásokhoz" (Turkina 2018: 2). Ismét, ami ebben az idézetben van, re- lálja mindazt, ami a fentebb tárgyalt, megvalósítás alatt álló új modell kibernetikus voltát. Hélixek forognak, a nyitottság a hozzáférést jelzi, a tempó egyenlő a körforgással és az inno- vációval stb. egy olyan mozgásban, amelyet bizonyára nehéz, de nem lehetetlen nyomon követni.

Míg az Element AI lényegét tekintve kibernetikus, addig ökológiai szempontból egy nagyon is gyakorlatias módon. A MILA közelsége alakítja a körülötte lévő városi teret, és méterben mérhető. Természetesen az ilyen közelség nem csak a Cité de l'IA-ban található meg; számos inkubátorház Észak-Amerikában, Európában és máshol is alkalmazza ezt a modellt azzal az indoklással, hogy hozzájárul az ötletek és erőforrások határokon átvéelő megosztásához. A probléma azonban némileg más, amikor a köz- és a magánvagyon összemosódásáról van szó. Emblematikus ebből

a szempontból az Element AI Facebook-bejegyzése, amelyben üdvözölte a MILA "a szomszédságba" való érkezését, és amely egy Bengio-képet mutatott be, miközben nyomatékosan hozzátette: "találkozunk a lépcsőházban" (Element AI 2018). Az ilyen metafora általában egy többé-kevésbé legális térre utal; olyanra, amelynek kódjai és határai többé-kevésbé homályosak. Ki megy fel, mi megy

le, mikor és hogyan? Különösen a gyakornokoskodó hallgatók esetében a kifejezett korlátozások hiánya mind egyénileg, mind pedig a kohorszok tekintetében mindig is problematikus. Ezek az állami és magán részmunkaidős vagy kétszintű kötődések összesítve egy olyan modellt erősítenek meg, amely kevésbé ellenőrzött és kiegyensúlyozott, különösen annak hosszú távú társadalmi-politikai és gazdasági hatásai és azok nyilvános megvitatásának és módosításának lehetőségei tekintetében.

Jelenleg a köz- és a magánszféra ilyen jellegű keveredése az, ami a montreali mesterséges intelligencia központ számos rétegét kolonizálja - annak ellenére, hogy Bengio bizonyos kérésekkel tiltakozik a valóság ellen, amihez valójában ő is hozzájárul.¹⁰ A montreali központban a külföldi vállalatok által a közelmúltban eszközölt befektetések hulláma kéz a kézben jár az újonnan "társult" tudósok által egyre inkább elfogadott új szervezeti elrendezéssel, nevezetesen a kettős tagsági modellel. Ez az egyetem és az ipar közötti együttműködési forma, amelyet a jog, a menedzsment és az orvostudomány területéről importáltak - és amely valószínűleg a biotechnológiai iparban a legerősebb; lásd Mirowski 2012 - lehetővé teszi a tudósok számára, hogy megtartsák egyetemi professzori kinevezésüket, miközben legalább részmunkaidőben elkötelezik magukat új vállalati munkaadójuk mellett (Serebrin 2017a; 2017b). Az e fejezet során elemzett listához még hozzá kellene tennünk a MILA-hez kötődő tudósok számos nevét, mint például Larochelle az Uni- versité de Montréal és a Google Brain, Precup a McGill és a DeepMind vagy Pal az Element AI és az Université de Montréal munkatársainak nevét. A kettős affiliációt a terület szereplői újszerű megoldásként indokolják, ahol a tudósok továbbra is képesek az oktatást és az alapkutatót folytatni, miközben részt vesznek az ipari K+F-ben is, míg korábban ez a részvétel az egyetemi oktatói és alapkutatói tevékenységtől való teljes visszavonulással járt volna (Plamondon Emond 2017). A modell növekvő elterjedése tehát két különböző, de egymással párhuzamosan zajló dinamika találkozásánál működik. Egyfelől a vállalati szereplők egyre inkább tudatában vannak annak, hogy üzleti modelljükhez elengedhetetlenül szükséges az "ökoszisztéma fenntarthatóságának" megőrzése - azaz a mesterséges intelligenciával foglalkozó kutatók jövőbeli generációinak folyamatos képzése és az alapkutatói törekvések további előmozdítása (LeCun 2018) -, valamint a II. fejezetben leírtak szerint a speciális humán erőforrások azonnali kisajátításának ellentmondásos kényszere között egyensúlyozni (Metz 2016a; 2017). Másrészt a tudósok olyan korlátokra reagálnak, amelyek többnyire ösztönzőként jelennek meg: az alternatív magánfinanszírozási források csábító lehetősége mellett a kutatóknak azzal is meg kell küzdeniük, hogy a legmodernebb vállalati számítási infrastruktúrákhoz és a legszélesebb körű saját adatbázisokhoz való hozzáférés valóban olyan technológiai eszközök, amelyekre egyre nagyobb szükség van a legmodernebb mélytanulási kutatások folytatásához. Ez az új közpénzekből finanszírozott

- 10 Lásd Shead kritikáját, 2018. amely szerint a nagy technológiai vállalatok egyre inkább koncentrálnak a technológiai eszközöket és a speciális emberi erőforrásokat, miközben aligha utasítják el a velük való együttműködés lehetőségeit - lásd Mathys és 2017. Vara. 2018.

vate összeszerelés és annak sajátos módja, hogy egyszerre sok megoldást kínáljon, a szükségszerűség egyik formájaként, nem pedig esetleges döntések és elvek kaszkádjaként van előre megírva, igazolva és legitimálva. Ebben az értelemben azonban mégis erősen politikai és a kormányzás egyik formája.

Következtetés

Ez a fejezet azzal kezdődött, hogy elismerte, hogy a fekete doboz egy erőteljes, de elvárásolt figura, hogy elgondolkodjon az olyan *készülő* technológiákról, mint a mélytanulás és tágabb értelemben a mesterséges intelligencia. Miközben van egy olyan sürgetés, amelynek ösztönöznie kell a társadalomtudományi kutatást, mindazonáltal fontos, hogy a dolgokat jól csináljuk, egy bizonyos adag ag- nosztikus és kritikai reflexivitással. Ennek fényében megpróbáltuk követni Bucher hármas tanácsát: "ne féljünk a fekete doboztól", "ne várjuk, hogy a megoldás benne van" és "fontoljuk meg a doboz dobozba zárását". És hogy ment? Hogyan alkalmaztuk az elméleti koncepciókat a gyakorlati valóságban, és hogyan segíthet egy olyan eset jobb megértése, mint a montreali csomópont, szélesebb körű és kritikusabb gondolkodásban? A válasz egy részét az I. fejezetben kaptuk meg, ahol az érvelés szerint a tét leginkább a mesterséges intelligencia jelenlegi és jövőbeli politikai gazdaságtana, vagyis az, hogy a tudástermelés automatizálása hogyan alakítja át a hatalmi viszonyokat, és hogy a mélytanulásban részt vevő különböző szereplők hogyan vesznek részt abban, amit Crandall a "kooperatív küzdelmek" sajátos formájának nevez (2010). Jelentős erőforrások, köztük pénz, állami támogatás, médiamegjelenések stb. áramlanak és halmozódnak, amelyeknek részleteit pontosan meg kell érteni a kormányzás e sűrű és feszült rendszerében. A mélytanulás és a mesterséges intelligenciával kapcsolatos technológiák által előidézett új normális állapot rendezetlen és ambivalens lesz, ha ez nem így van már most is. A fejezetben végig hangsúlyoztuk, hogy a hatalom minden eddigénél inkább tranzakció, és hogy amit az "irányítás" jelent ilyen körülmények között, az a kibernetika és az ökológia új értelméhez kapcsolódik, amely a mutualizmus és a parazitizmus minden típusát figyelembe veszi. A II. fejezetben ezt a montreali példán keresztül mutattuk be, különös tekintettel arra, hogy a montreali példa a kormányzati, egyetemi, befutott és feltörekvő vállalati szereplők közötti spirálok sajátos forgómozgását mutatja. Míg a szereplők ismételten azt hangoztatják, hogy létezik egy "közösség", és hogy a Klaszter, a Megfigyelőközpont, a Deklaráció és a *Cité de l'IA* egy egységes egészet alkot, mi egy kevésbé optimista, realisabb elemzést javasolunk. A nyitott tudomány egy ilyen eset, mivel a "nyitott" az aggregációt jelenti, és mivel a divatos oktatási modellben fontos változást jelez. A felsőoktatás privatizációja olyan helyeken, mint Quebec, ahol a mélytanulás és a mesterséges intelligenciához kapcsolódó technológiák fontos szerepet játszanak. Ez volt a III.

szakasz fő következtetése. Akár "kettős hovatartozásnak", akár a "találkozunk a lépcsőházban" modellnek nevezzük, az egyértelmű, hogy az előnyök jelenleg nem egyenlően oszlanak meg újra, és nagyon rossz esélye van - legalábbis a-

oretikusan - hogy a jövőben is így lesz. Végző soron talán ez a kibontakozás, a maga sokféle fordulatával és fordulatával, képezi egy másik, még kialakulóban lévő ökológia megfelelő tárgyát, amelyet kritikai mesterséges intelligencia-tanulmányoknak nevezhetnénk.

Hivatkozások

- Ananny, Mike és Crawford, Kate (2018 [2016]): "Seeing without Knowing: A transzparenciaideál korlátai és alkalmazása az algoritmikus elszámoltathatóságra." *New Media & Society* 20/3, pp. 973-89. <https://doi.org/10.1177/1461444816676645>.
- Bachand, Olivier (2018): "Une Cité de l'intelligence Artificielle Dans Le Mile-Ex à Montréal". Radio-Canada, letölthető: <https://ici.radio-canada.ca/nouvelle/1081709/technologie-montreal-saint-urbain-intelligence-artificielle> (accessed March 11, 2019).
- Biddle, Sam (2019): "Bízzunk-e a mesterséges intelligencia szabályozásában a kongresszus által, ha a Facebook támogatja?" *The Intercept*, letölthető: <https://theintercept.com/2019/03/07/artificial-intelligence-facebook-ibm-rokhanna/> (accessed March 9, 2019).
- Bijker, Wiebe E., Hugues, Thomas Parke és Pinch, Trevor (szerk.) (2012 [1987]): *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*. Cambridge, Mass: MIT Press.
- Bourgault, Mathieu (2017): "Összefoglaló a québeci mesterséges intelligencia médiakörnyezetéről". *Kiadatlan kézirat*, pp. 1-30.
- Browne, Ryan (2019): "IBM Hopes 1 Million Faces Will Help Fight Bias in Facial Recognition (Az IBM reméli, hogy 1 millió arc segít az arcfelismerés elfogultságá ellen)." *CNBC*, letöltve: <https://www.cnn.com/2019/01/29/ibm-releases-diverse-dataset-to-fight-facial-recognition-bias.html> (hozzáférés: 2019. március 9.).
- Bucher, Taina (2016): "Sem fekete, sem doboz: Ways of Knowing Algorithms". In *Innovatív módszerek a média- és kommunikációkutatásban*, szerkesztette: Sebastian Kubitschko és Anne Kaun, pp. 81-98. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-40700-5_5.
- Burrell, Jenna (2016): "Hogyan "gondolkodik" a gép: Understanding Opacity in Machine Learning Algorithms (A gépi tanulási algoritmusok átláthatatlanságának megértése)". *Big Data & Society*, 3/1., pp. 1-12. <https://doi.org/10.1177/2053951715622512>.
- Cardon, Dominique, Cointet, Jean-Philippe és Mazières, Antoine (2018): "La revanche des neurones: L'invention des machines inductives et la controverse de l'intelligence artificielle". *Réseaux* 211/5, pp. 173-220. <https://doi.org/10.3917/res.211.0173>.
- Castelle, Michael, Millo, Yuval, Beunza, Daniel és Lubin, David C. (2016): "Honnan jönnek az elektronikus piacok? Regulation and the Transforma-

- tion of Financial Exchanges." *Gazdaság és Társadalom* 45/2, pp. 166-200. <https://doi.org/10.1080/03085147.2016.1213985>.
- CIFAR (2018): "29 kutatót neveztek ki a kanadai CIFAR művészeti intelligencia tanszékek első kohorszába". CIFAR, letölthető: <https://www.cifar.ca/cifarnews/2018/12/03/29-researchers-named-to-first-cohort-of-canada-cifar-artificial-intelligence-chairs> (hozzáférés: 2019. március 9.).
- Crandall, Jordan (2010): "A számítási műveletek térbeli térbeli megjelenítése: A megavárosok nyomon követése, érzékelése és a megavárosok". *Theory, Culture & Society*, 27/6, pp. 68-90. <https://doi.org/10.1177/0263276410382027>.
- Dubuc, André (2018): "Microsoft Déménage Ses 200 Employés Dans Le Mile-Ex". La Presse, letölthető: http://plus.lapresse.ca/screens/60fda222-94fb-49ee-ac0a-05a933c166a0__7C__0.html?utm_medium=Twitter&utm_campaign=Internal+Share&utm_content=Screen (elérés: 2019. március 11.).
- Element AI (2018): "Találkozunk a lépcsőházban". *Facebook-poszt*, január 29., letölthető: <https://www.facebook.com> (elérés: 2019. március 9.).
- Engemann, Christoph és Sudmann, Andreas, szerk. (2018): *Machine Learning. Medien, Infrastrukturen und Technologien der Künstlichen Intelligenz*. Bielefeld: transcript.
- Etzkowicz, Henry és Leydesdorff, Loet (2000): "Az innováció dinamikája: A nemzeti rendszerektől és a "Mode 2"-től az egyetem-indusztriális-önkormányzati kapcsolatok hármasspiráljáig". *Research Policy* 29/2, pp. 109-23. [https://doi.org/10.1016/S0048-7333\(99\)00055-4](https://doi.org/10.1016/S0048-7333(99)00055-4). [https://doi.org/10.1016/S0048-7333\(99\)00055-4](https://doi.org/10.1016/S0048-7333(99)00055-4).
- Foucault, Michel (2004a): Foucault: "Naissance de La Biopolitique". In *Cours Au Collège de France 1978-1979*. Párizs: Gallimard-Seuil.
- Geiger, R Stuart (2017): "Beyond Opening up the Black Box: Investigating the Role of Algorithmic Systems in Wikipedian Organizational Culture" (Az algoritmusok szerepének vizsgálata a Wikipédia szervezeti kultúrájában)." *Big Data & Society* 4/2, pp. 1-14. <https://doi.org/10.1177/2053951717730735>.
- George-Cosh, David (2018): "Element AI célja az unikornis státusz elérése kanadai rekordfinanszírozással: Források." BNN Bloomberg, lekérdezve: <https://www.bnnbloomberg.ca/element-ai-aims-for-unicorn-status-with-record-canadian-financing-sources-1.1101206> (hozzáférés: 2019. március).
- Gillespie, Tarleton (2014): "Algoritmus [tervezet] [#digitalkeywords]." *Kultúra digitálisan* (blog), letölthető: <http://culturedigitally.org/2014/06/algorithm-draft-digi-talkeyword/> (hozzáférés: 2019. március 9.).
- (2017): "A platformok és a platformok általi szabályozás". In *SAGE Handbook of Social Media*, szerkesztette: Jean Burgess, Thomas Poell és Alice Marwick, pp. 254-78. London: Sage.
- Greene, Daniel, Hoffman, Anna Lauren és Stark, Luke (2019): "Better, Nicer, Tisztább, igazságosabb: Az etikus mesterséges intelligencia és a gépi tanulás mozgalmának kritikai értékelése." Proceedings of the 52nd Hawaii International Conference on System Sciences.

- Guice, Jon (1999): Guice: "A jövő tervezése: a tudomány és a technológia új irányzatainak kultúrája". *Research Policy*, 28, pp. 81-98.
- Habermas, Jürgen (1971): *Habermas: Toward a Rational Society: Habermas: Student Protest, Science and Politics*. London: Heinemann.
- Hernandez, Daniela (2014): "Meet the Man Google Hired to Make AI a Reality". *Wired*, letölthető: <https://www.wired.com/2014/01/geoffrey-hinton-deep-learning/> (hozzáférés: 2019. március 9.).
- Hernandez, Daniela és King, Rachael (2016): "Universities' AI Talent Poaching by Tech Giants". *Wall Street Journal*, letölthető: <http://www.wsj.com/articles/universities-ai-talent-poached-by-tech-giants-1479999601> (elérés: 2019. március 9.).
- Hinton, Geoffrey, Osindero, Simon és The, Yee-Whye (2006): "A Fast Learning Algorithm for Deep Belief Nets". *Neural Computation* 18/7, pp. 1527-1554.
- Hoffman, Steve G. (2017): "Kétértelműségek kezelése a tudás határán: Research Strategy and Artificial Intelligence Labs in an Era of Academic Capitalism". *Science, Technology, & Human Values* 42/4, pp. 703-40. <https://doi.org/10.1177/0162243916687038>.
- Hutson, Matthew (2018): "A mesterséges intelligencia reprodukálhatósági válsággal néz szembe". *Science* 359/6377, pp. 725-726. <https://doi.org/10.1126/science.359.6377.725>. <https://doi.org/10.1126/science.359.6377.725>.
- IA. Québec (2018): "Effervescence de l'intelligence Artificielle". Comité d'orientation de La Grappe En Intelligence Artificielle, letölthető a <http://ia.quebec/> (hozzáférés: 2019. március 9.).
- IA responsable (2017) : "Déclaration de Montréal Pour Un Développement Responsable de l'IA. " Déclaration de Montréal IA Responsable, letölthető: <https://www.declarationmontreal-iaresponsable.com/> (hozzáférés: 2019. március 9.).
- Kitchin, Rob (2014): "Big Data, új episztemológiák és paradigmaváltások". *Big Data & Society* 1/1, pp. 1-12. <https://doi.org/10.1177/2053951714528481>. <https://doi.org/10.1177/2053951714528481>
- Knorr-Cetina, Karin és Preda, Alex (szerk.) (2011 [2006]): *A pénzügyi szociológia*. *Piacok*. Oxford: Oxford University Press.
- Lascoumes, Pierre (2004): "La Gouvernamentalité : De La Critique de l'État Aux Technologies Du Pouvoir". *Le Portique*, 13-14. szám, pp. 1-16.
- Latour, Bruno (1987) : "Tudomány a cselekvésben: Cambridge, Massachusetts: Hogyan kövessük a tudósokat és mérnököket a társadalomban? Harvard University Press.
- LeCun, Yann (2018): "A Facebook vezető mesterséges intelligencia kutatója szerint a Szilícium-völgynek szorosabban kell együttműködnie az akadémiákkal a mesterséges intelligencia jövőjének megteremtése érdekében." *Business Insider*, letölthető: <https://www.businessinsider.com/facebook-yann-lecun-dual-affiliation-model-ai-experts-2018-8> (hozzáférés: 2019., március). Leonelli, Sabina (2013): "Miért ragaszkodunk jelenleg a tudományos adatok nyílt hozzáféréséhez? Big

Data, a tudástermelés és az ideiglenes biológia politikai gazdaságtana". *Bulletin of Science and Technology Studies* 33/1-2, pp. 6-11.
<https://doi.org/10.1177/0270467613496768>.

- Levy, Karen EC és Johns, David Merritt (2016): "When Open Data Is a Trojan Horse: The Weaponization of Transparency in Science and Governance" (Amikor a nyílt adatok trójai falónak számítanak: Az átláthatóság fegyverré tétele a tudományban és a kormányzás területén). *Big Data & Society* 3/1, pp. 1-6. <https://doi.org/10.1177/2053951715621568>.
<https://doi.org/10.1177/2053951715621568>
- Lheureux, Alexandra, Grolinger, Katarina, Elyamany, Hany F. és Capretz, Miriam A. M. (2017): "Gépi tanulás nagy adatokkal: Challenges and Approach- es." *IEEE Access* 5, pp. 7776-7797.
- Litvinski, Oleg (2018): "Az algoritmikus átláthatatlanság társadalmi mechanizmusairól". *Kiadatlan kézirat*, pp. 1-32.
- Mackenzie, Adrian (2013): "A programozó szubjektumok az anticipáció rendszerében: Software Studies and Subjectivity (Szoftvertanulmányok és szubjektivitás)". *Subjectivity* 6/4, pp. 391-405.
- (2017): *Machine Learners: Archaeology of a Data Practice*. Cambridge, MA: The MIT Press.
- (2018): "Az API-tól a mesterséges intelligenciáig: platformok és azok opacitása". *Information, Communication & Society*, June, pp. 1-18. <https://doi.org/10.1080/1369118X.2018.1476569>.
- Mathys, Catherine (2017): Mathys: "Le Rebelle de l'intelligence Artificielle". L'actualité, lekérdezve a <http://lactualite.com/techno/2017/11/10/le-rebelle-de-lintelligence-artificielle/> (hozzáférés: 2019. március 9.).
- Mercure, Philippe (2016): "Montréal, Future Plaque Tournante de l'intelligence Artificielle?" *La Presse*, letölthető: <https://www.lapresse.ca/af-faires/economie/2016/10/26/01-5034411-montreal-future-plaque-tournante-de-lintelligence-artificielle.php> (hozzáférés: 2019. március 9.).
- Metz, Cade (2016a): Metz: "Giant Corporations Are Hoarding the World's AI Talent". *Wired*, letölthető: <https://www.wired.com/2016/11/giant-corporations-hoarding-worlds-ai-talent/> (hozzáférés: 2019. március 9.).
- (2016b): "A GOOGLE MEGNYITJA MONTREALI AI LABORT, HOGY ELKAPJA A RITKA GLOBÁLIS TALENT." *Wired*, letöltve: <https://www.wired.com/2016/11/google-opens-montreal-ai-lab-snap-scarce-global-talent/> (hozzáférés: 2019. március 9.).
- (2017): "A GOOGLE SZÁMÁRA AZ AI TEHETSÉGGUTATÓ VERSENY EGYENESEN KANADÁBA VEZET." *Wired*, letöltve: <https://www.wired.com/2017/03/google-ai-talent-race-leads-straight-canada/> (hozzáférés: 2019. március 9.).
- Mirowski, Philip (2012): "A tudomány modern kommercializálása a Ponzi-rendszerek halmaza". *Social Epistemology* 26/3-4., pp. 285-310.
- (2018): "A nyílt tudomány jövője(i)". *Social Studies of Science* 48/2, pp. 171-203. <https://doi.org/10.1177/0306312718772086>.
- O'Brien, Sara Ashley (2018): "Holmes körülvette a TheraNost hatalommal rendelkező emberekkel." *CNN Business*, letöltve: <https://money.cnn.com/2018/03/15/technology/elizabeth-holmes-theranos/index.html> (elérés: 2019. március 9.).

- Pasquale, Frank (2015): *A Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge, Massachusetts: Harvard University Press.
- Plamondon Emond, Étienne (2017): "Intelligence Artificielle: Un Pied à McGill et l'autre Chez Un Géant Des Technos". Le Devoir, lekérdezve a <https://www.>

- ledevoir.com/societe/education/511873/intelligence-artificielle-un-pied-a-mc-gill-et-l-autre-chez-un-geant-des-technos (hozzáférés: 2019. március 9.).
- Rettino-Parazelli, Karl (2017): "L'intelligence Artificielle, Moteur Économique. Le Devoir, letölthető: <https://www.ledevoir.com/economie/500340/dominique-anglade-au-devoir> (hozzáférés: 2019. március 9.).
- (2019): "Nouveaux Bureaux, Nouveaux Besoins Pour Mila." "LeDevoir, retrieved from https://www.ledevoir.com/economie/546567/nouveaux-bureaux-nouveaux-besoins?fbclid=IwAR3XDUBbIUrqGIt5oXUfZTdp8GqYtY6kbVQsG18pp_slKzpPQFgv_s2IQ (hozzáférés: 2019., március).
- Roberge, Jonathan és Seyfert, Robert (2018 [2016]). "What are algorithmic cultures?" in *Algorithmic Cultures: Essays on Meaning, Performance and New Technologies*, Robert Seyfert és Jonathan Roberge (szerk.), pp. 1-25, New York: Routledge.
- Rosenblatt, Frank (1958): "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain (Az információ tárolásának és szervezésének valószínűségi modellje az agyban)". *Psychological Review* 65/6, pp. 386-408. Salter, Christopher (2018). Személyes közlés: "AI Talks: The New Silicon Valley - Az északi ley, Tényleg?", december Montreal, Concordia Egyetem.
- Saxenian, AnnaLee (1994): *Saxen Saxen: Regional Advantage: Szilícium-völgy és a 128-as út: Kultúra és verseny a Szilícium-völgyben*. Cambridge, Mass: Harvard University Press.
- Serebrin, Jacob (2017a): "Facebook to Open AI Lab in Montreal, Headed by McGill Professor". Montreal Gazette, letöltve: <https://montrealgazette.com/news/local-news/facebook-to-open-ai-lab-in-montreal-headed-by-mcgill-professor> (elérés: 2019., március).
- (2017b): "Google-Linked AI Company to Open Research Lab in Montreal". Montreal Gazette, letölthető: <https://montrealgazette.com/business/local-business/google-affiliated-ai-company-deepmind-to-open-research-lab-in-montreal> (hozzáférés: 2019., március).
- Seyfert, Robert (2018): "Automatizálás és affektus: A Study of Algorithmic Trading." In *Affect in Relation - Családok, helyek, technológiák. Essays on Affectivity and Subject Formation in the 21st Century*, Birgitt Röttger-Rössler, Jan Slaby (szerk.), pp. 197-218. New York: Routledge.
- Sudmann, Andreas (2018): "A mesterséges intelligencia médiapolitikai dimenziójáról: Deep Learning as a Black Box and OpenAI.", in: Rethinking AI. Neural Networks, Biopolitics and the New Artificial Intelligence, *Digital Culture & Society* 4/1, pp. 181-200. (elérhető az academia.edu oldalon)
- Shead, Sam (2018): "Bengio szerint az egyetemek több elismerést érdemelnének." Forbes, letölthető: <https://www.forbes.com/sites/samshead/2018/11/13/ai-pioneer-yoshua-bengio-says-universities-deserve-more-credit-for-their-ai-research/#5cb7d54273ab> (hozzáférés: 2019., március).
- Simard, Jean-Jacques (1979): "Québec et frères, inc. La cybernétisation du pouvoir". *Recherches sociographiques*, 20/2, pp. 239-261. <https://doi.org/10.7202/055840ar>.

- Simonite, Tom (2018): "Despite Pledge Pending Openness, Companies Rush To Patent AI Tech". Wired, letölthető: <https://www.wired.com/story/despite-pledging-openness-companies-rush-to-patent-ai-tech/> (hozzáférés: 2019. március 9.).
- (2019): "AZ AMAZON CSATLAKOZIK A MICROSOFT FELHÍVÁSÁHOZ AZ ARCFELISMERÉSRE VONATKOZÓ SZABÁLYOKÉRT. "OGNITION." Wired, letöltve: <https://www.wired.com/story/amazon-joins-microsofts-call-rules-facial-recognition/> (hozzáférés: 2019. március 9.).
- Slaughter, Sheila és Rhoades, Gary (2010): *Gades: Academic Capitalism and the New Economy: Piacok, állam és felsőoktatás*. Baltimore: Johns Hopkins University Press.
- Storper, Michael, Kemeny, Thomas, Makarem, Naji Philip, and Osman, Taner (2015): *The Rise and Fall of Urban Economies: Lessons from San Francisco and Los Angeles*. Innováció és technológia a világgazdaságban. Stanford, Kalifornia: Stanford Business Books.
- The Economist (2017): "Egy hibrid startup AI-szolgáltatásokat kínál a vállalkozásoknak." The Economist, letölthető: <https://www.economist.com/business/2017/06/22/a-hybrid-startup-offers-ai-services-to-business> (hozzáférés: 2019. március 9.).
- Turkina, E. (2018): "A hálózatépítés jelentősége a vállalkozói tevékenységben: Montreal's Artificial Intelligence Cluster and Its Born-Global Firm Element AI". *Journal of Small Business & Entrepreneurship* 30/1, pp. 1-8. <https://doi.org/10.1080/08980101.2017.1402154>.
- Vara, Vauhini (2018): Vuhara: "Can This Startup Break Big Tech's Hold on A.I.?" Fortune, lekérdezve a <http://fortune.com/longform/element-ai-startup-tech/> (2019. március 9-i dátummal).
- Wagner, Ben (2018): "Az etika mint menekülés a szabályozás elől: Az etikai mosakodástól az etikai vásárlásig?" In *Being Profiling. Cogitas Ergo Sum.*, Amsterdam University Press, pp. 1-7. M. Hildebrandt.
- Walters, William (2004): Néhány kritikai megjegyzés a "Kormányzásról": "Néhány kritikai megjegyzés a 'Kormányzásról'". *Studies in Political Economy* 73/1, pp. 27-46. <https://doi.org/10.1080/19187033.2004.11675150>.

Csökkenés és részvétel

Stefan Rieger

Három évvel ezelőtt a kaliforniai Mountain View-ban található titkos Google X laboratórium kutatói mintegy 10 millió állóképet vettek ki a YouTube-videókból, és betáplálták őket a Google Brainbe - egy 1000 számítógépből álló hálózatba, amelyet úgy programoztak, hogy az emberi kisgyermekhez hasonlóan szívja magába a világot. Három nap múlva a Google Brain magától döntött úgy, hogy vannak bizonyos ismétlődő kategóriák, amelyeket azonosítani tud: emberi arcok, emberi testek és ... macskák. (Jones 2014: 146)

1. Mély tanulás

Ekkor már mindenütt az algoritmusok mindenhatóságáról beszélnek. Ez a diskurzus megszakítás nélkül folyik, és látszólag lehetetlen megállítani - nem utolsósorban azért, mert az algoritmusok csendesen és észrevétlenül működnek a háttérben (vö. Bunz 2012; Seyfert/Roberge 2017). A befolyásukról szóló viták közül sokan az átláthatatlanságuk státuszával foglalkoznak, és azért, hogy a cégek elutasítására összpontosítanak, hogy átláthatóvá tegyék őket, olyan érveket hoznak játékba, amelyek egy másik korszakból származó relikviáknak tűnnek. Míg akkor a kritikusok középpontjában egy lejáratott kultúrpar tevékenysége állt, ma az eddig elképzelhetetlen mennyiségű adat gazdaságosítását tekintik jogsértésnek. Az adatok gazdasági értéke egy olyan médiakritika tárgyává vált, amely elvesztette az előző évszázad egyik kedvenc témáját: a kritikus és autonóm médiahasználót (vagy azt, akit egykor annak tekintettek). Az olyan nagyvállalatok, mint a Google, az Amazon vagy a Facebook algoritmusai joggal tűnnek úgy, mintha az utóbbi százévtől cselekvési, autonómiai, ellenállási és felforgatási potenciálját magába olvasztották volna (vö. Sudmann 2017). Ez a folyamat olyan sikeres volt, hogy még olyan ellenmozgalmakhoz is vezetett, amelyek nem lazán siránkoznak a magánszféra végéről mint a digitalizáció járulékos káráról, hanem inkább olyan napirendeket fogadtak el, amelyek lelkesen támogatják annak felszámolását (vö. Rieger 2018). Az úgynevezett *Post-Privacy Spackeria* internetes exhibicionistái számára az adatvédelem nem több, mint egy történelmileg datálható maradvány, egy csökevény az elmúlt millenniumból: "A magánszféra olyan 1980-as évekbeli". (Reißmann 2019., pag.)

Az algoritmusok használatának alkalmazási területei, amelyek a maguk részéről a tárgya egy rövid evolúció, és amelynek optimalizálása volt orientált nem

utolsósorban

a természet előírásainak való megfelelés felé, mindenütt jelen vannak, és olyan sokfélék, hogy nem lehet őket teljes egészében felmérni: legyen 'szó a mindennapi életben az arcok felismeréséről a késleltetett megfigyelés vagy a jövőbe mutató kriminalisztika érdekében, a viselkedéssorozat azonosításáról vagy a biopolitikáról, a képek és szövegek szerzőségének tisztázásáról (vö. Rodri-guez et al. 2019; Rehman et al. 2019), a műalkotásoknak egy adott korszak stílus szerinti osztályozását vagy a feltételezhetően azonos kéz által írt aláírások összehasonlítását, a tudomány üzleti életébe való beleszólást és a tudás látszólag nem számítható objektumaival való szembesülést a big data és az algoritmizáció segítségével (vö. Rieger 2019), a "com- putational turn" és a "humanities computing" során egyes diszciplínák önképét és önértékelését befolyásolva, az utóbbi diszciplínákat a reflexió különböző formáival társítva, és ezáltal alapvető változásokhoz járulva hozzá magán a humán tudományokon belül (vö. Hall 2013), vagy más módon beavatkozva a dolgok rendjébe - az ilyen tevékenység jellemzően a mesterséges intelligencia, a mesterséges neurális hálózatok és a mélytanulás folyamataira támaszkodik. A nagy mennyiségű adat kezelésének módja tudásserkentő játékká vált, és még a foucault-i diskurzuselemzés számára is új lehetőségeket nyitott meg, ami a technológiai de- vülések esetében ritkán alkalmazható (vö. Engemann/Sudmann 2018). A mesterséges intelligencia lehetőségei egyenesen Foucault azon alapvető intuíciójának a kezére játszanak, hogy "az empirikus tudás egy adott időben és egy adott kultúrában *valóban* rendelkezett egy jól meghatározott szabályszerűséggel", és hogy "a nem formális tudás története maga is rendszer volt". (Foucault 2002 [1966]: x) A *dolgok rendje* című könyve során Foucault a *tudás* egy episztemológiailag szigorú (de szaknyelven szólva alig megvalósítható) *pozitív tudattalanját* igyekezett feltárni, és ezzel annak a feltételezésnek adott hangot, hogy a nem formális tudás mögött is van egy "jól meghatározott szabályszerűség" - egy formális kód. Lehetséges lenne tehát, hogy e tudás tudományát másképp dolgozzuk fel: egy algoritmikus diskurzuselemzés tárgyává válhatna, és távol maradhatna az egyéni megértéstől és felfogástól. A kulturális analitika által alkalmazott hozzáférési módokban a tudás ilyen pozitív tudattalanját technikai sebességre hozzák, és szabályszerűségek és ismétlődések formájában teszik láthatóvá. Az adatbányászat és a szövegbányászat olyan mintázatokat és így olyan tudásformákat tesz láthatóvá, amelyek nem feltétlenül merülnek ki a szándékos kérdésekben. Itt mindaz, amit az emberi intelligencia tudományos nárcizmusában valódi tevékenységi területének tekint - a dolgok rendezése és osztályozása, a hasonlóságok azonosítása és a genealógiák létrehozása -, az algoritmusokra hárul. Ebben az esetben tehát a tudomány nem a véletlennek van kiszolgáltatva a tudás előállítására irányuló erőfeszítései során; az azonosságok és különbségek feldolgozása inkább automatikusan történik - algoritmikus és nem antropogén támogatással.

Ez azonban nem csak a tudományokat érinti, azok széles tárgykörével és a igényt tartanak a velük kapcsolatos komplexitásra. Az algoritmusok tevékenysége még az alábbiakra is kiterjed

I A szóban forgó kulcsszavak evolúciós algoritmusok, evolúciós vagy genetikai programozás.

az alsó érzékszervek, amelyek hosszú ideig alig kaptak figyelmet a kultúrtörténetben, de azóta a különböző natu-ralizációs mozgalmak erőfeszítéseinek köszönhetően reflektorfénybe kerültek (vö. Kortum 2008). Mint szinte minden más, a szagok érzékelése is algoritmusokra bízható - azzal a következménnyel, hogy ahol a szaglási adatok nagy mennyiségben és nagy sebességgel (több szereplő egyik kedvenc kifejezésével élve: *valós időben*) automatikusan feldolgozhatók, ott egy ismerős veszély fenyeget. A szagok esetében ezt a veszélyt "szagmegfigyelésnek" nevezték el. Jeremy Bentham panoptikuma mellett, hogy mindent lát, most már mindent megszagolhat is (vö. Stark et al. 2018a). A konszenzus ezzel kapcsolatban úgy tűnik, hogy a szagok ilyen rendszerét más biometrikus tevékenységek kinövésének kell tekinteni, és ennek megfelelően ellenezni kell. Természetesen a következő csak egy retorikai kérdés: Valóban ezt a fajta szagmegfigyelést akarjuk? (Stark et al. 2018b: 18) És úgy tűnik, konszenzus van abban is, hogy az ilyen jellegű automatizált tevékenységeknek alapvető reflexió tárgyát kell képezniük a "veillance" természetét illetően annak minden változatában (ez utóbbiak közé tartozik ma már a "sousveillance" és a "metaveillance" is) (vö. Kammerer/Waitz 2015). Sőt, ezt a gondolatot még programszerűen is kifejtették - olyan címekekkel ellátott művekben, mint "Declaration of Veillance (Surveillance is a Half-Truth)" (Mann 2015).

Itt kell megemlíteni az olyan főszereplőket, mint Steve Mann (2016) vagy José van Dijck (2014), az előbbit a felügyelet meghaladó felügyelet fogalmának bevezetéséért, az utóbbit az adatfikáció fogalmáért, amely az adatpolitika normalizálódását és annak hatókörét egy újfajta valutaként írja le. Az adatfikációval és az ilyen adatfeldolgozás átláthatatlan módszereivel kapcsolatos alapvető gyanakvással úgy tűnik, hogy a korábbi idők médiakritikai lendülete megmaradt, és nem kapitulált a magánszféra teljes feladásának követelése előtt. *Post-Privacy: Prima leben ohne Privatsphäre* című könyvében például Christian Heller internetes aktivista az utóbbit értékeli, még ha vannak is érvek a teljes lemondás mellett. Példaként említi, hogy az algoritmusok képesek meghatározni az egyének szexuális orientációját a társadalmi viselkedésükből - természetesen az érintettek biztonságára való tekintet nélkül:

A szexuális irányultsága magánügy, és ennek így is kell maradnia. A Massachusetts Institute of Technology (MIT) feltalálóinak figyelembevétel nélkül hozta létre azonban a beszámolóját. Ott kifejlesztettek egy eljárást, amellyel nagy valószínűséggel azonosítani lehet a férfiak homoszexualitását a Facebook-profiljuk alapján, még akkor is, ha nem posztoltak fényképet, vagy nem tüntettek fel semmilyen preferenciát. Mindössze a Facebookon lévő szociális környezetüket kell elemezni, amelyet mindenekelőtt a barátokkal, rokonokkal és ismerősökkel való kapcsolattartásra használnak. Gyakran előfordul, hogy a profilok tartalmazzák a barátok listáját, amely bárki számára látható az egész világon (lehetőség van arra, hogy ezt az információt privátá tegyék, de kevesen veszik a fáradságot, hogy ezt megtegyék). Az MIT kutatói felfedezték,

hogy közelítő előrejelzéseket lehet tenni arra vonatkozóan, hogy egy férfi diák meleg-e annak alapján, hogy a Facebook-ismerősei között mekkora azoknak a férfiaknak az aránya, akik saját profiljukon melegnek vallották magukat (Heller 2011: 12).

2. A részvétel stratégiái

A mesterséges intelligencia mindenütt jelen van, függetlenül attól, hogy ismerjük-e, megismerhetjük-e, vagy egyáltalán meg akarjuk-e ismerni. A megfigyelési tanulmányok vagy a kritikus kódok tanulmányozásának aggályaitól eltekintve, a mindennapi felhasználói viselkedést gyakran az algoritmusok tevékenysége és a biztonsági kérdések iránti alapvető és vakmerő közömbösség határozza meg. Amint az nemcsak az emberek belépési kódjaiból és jelszavaiból (könnyen megfejthető számsorozatok, születésnapok, rokonok és háziállatok beceneveinek használata), hanem abból is kitűnik, hogy hajlandóak felfedni fogyasztói preferenciáikat és egyéb szokásaikat, ez a viselkedés a könnyelmű és gondatlan hanyagságot példázza. Van azonban egy másik szempont is, amely meghatározza, hogy mi történik, és egyre inkább meghatározza, hogyan viszonyulunk a mesterséges intelligenciához, amely talán kevésbé látható, és első pillantásra távol áll a konkrét politikai cselekvéstől. Míg az algoritmusok mindenütt monopolizálják az autonómiát, míg azok önrendelkezve vagy részben önrendelkezve működnek, míg azok exe

egyre nagyobb hatékonysággal és egyre nagyobb kapacitású hardverrel vágják az automatizált tudás nagyszabású rendszerét, és miközben - ahogyan azt többször is olvashatjuk - észrevétlenül és az operatív látencia módjában végzik tevékenységüket, a használat és a részvétel, a társadalmi elkötelezettség és az elfogadásért folytatott kampány szintjén egy sajátos ellenmozgás zajlik. Ez a folyamat azért különös, mert látszólag felborítja a technológiát általában és különösen a digitalizációt körülvevő nagy narratíva rendjét. A körülöttünk lévő technológiáról és az inter- net kapcsolódó politikájáról szóló nagy narratíva jellemzően a mennyiségi növekedés elvéhez kötődik. Ezt nagy számok formájában lehet elbeszélni, és úgy írják meg, mint a haladás történetét, a növekvő komplexitás teljesen könnyörtelen diadalát.

Ezért még inkább szembeűnő, hogy a technikai (vagy talán jobb lenne azt mondani, hogy társadalmi-technikai) infrastruktúrákkal való foglalkozásban olyan tendenciákat látunk, amelyek más irányba haladnak, és a növekedés ellenkezőjére épülnek - vagyis arra, amit itt "Csökkentés és részvétel" címmel tárgyalunk. A megszokott sikertörténet megszakítása és a technikailag lehetséges komplexitás szándékos csökkentése figyelemre méltó - és különböző módokon az internetpolitika, a demokratizálódás és annak kérdése körül forog, hogy kinek kell egyáltalán hozzáférnie (és milyen módon). Ami különösen figyelemre méltó, az az interneten való részvételre és így a megszólításra alkalmasnak tekintett dolgok alapvető kiterjesztése. E bővülés során, ahogyan azt a

későbbiekben bemutatjuk, különböző és további ágensek kerültek a részvétel pozíciójába - olyan ágensek, akik az ember-számítógép interakció (HCI) domináns szempontjain kívül helyezkednek el, és akik támogatják a redukció melletti érvelést, vagy legalábbis jelzik az ehhez kapcsolódó gesztusokat. Azok, akik némileg szisztematikusan a terv részévé váltak, olyan felhasználók, akik sajátos profiljukkal valószínűleg megtestesítik a redukció kérdését. Ezek a különleges felhasználók fenotípusosan sokfélék, és így nem utolsósorban a gyermekek és a kihívásokkal küzdő emberek egyre nagyobb figyelmet kapnak, mint a különleges felhasználói profillal rendelkezők szélsőséges esetei: "A HCI-metódusok módosítása gyakori, amikor az interakciótervezést "extrém" emberi felhasználók számára tervezik". (Hirskyj-Douglas et al. é. n2016., pag.)

Ez azonban nem csupán arról szól, hogy az embereket aszerint különböztetjük meg, hogy milyen

fejlődési szakaszok (gyermekek) vagy sajátos kihívások (süketek, vakok, autisták, idősek; kognitív vagy egyéb kihívásokkal küzdő személyek) szerint. Az emberi lényeken túl a komplexitás kívánt csökkentése új szereplőket is a játékba hoz. Ebből a szempontból említésre méltó az "állat-számítógép interakció" (ACI). Clara Mancini, ennek a mozgalomnak az egyik vezetője, szeret rámutatni, hogy az ilyen megközelítések mögött több van, mint pusztán kazuisztika, és hogy többről van szó, mint pusztán anekdotákról a skypeoló kutyákról és a csevegő macskákról (vö. Ritvo/Allison 2014; Pongrácz et al. 2016; Golbeck/Neustaedter 2012). Mancini programja inkább egy olyan rendszert képvisel, amely alapvetően a megváltozott szociális

formák ("fajok közötti közösségek"), és elkötelezett a "több fajra vonatkozó tudatosság" előmozdítása mellett (vö. Mankoff et al. 2005). Törekvését egy átfogó szisztematikus megközelítés ígéretével és egy újonnan kialakuló tudományág magabiztosságával ötvözi, ahogyan az a kiáltványából és annak pozitív fogadtatásából lenyűgözően kitűnik (vö. Mancini 2011; Hirskyj-Douglas et al. 2018).

Ez a kötelezettség nem utolsósorban társadalmi felelősségvállalás kérdése. A fajok közötti határok túllépéséhez hasonlóan ez is a részvétel tágabb értelemben vett elképzeléseinek köszönhető (vö. Kelty 2016; Stahl 2014). Ez a figyelmesség a szemiotikai rendszerek sajátosságainak mélyebb figyelembevételével és a saját etikai pozíciók újragondolásával jár együtt (vö. Mancini 2017)2011.. A fajok keveredésének és a politikai ébredésnek a bátorításával az olyan megközelítések, mint az ACI, a transzhumanizmus vagy poszthumanizmus néven ismert nagyobb szellemi mozgalom részét képezik. Ez utóbbit a differenciálódásról programszerűen lemondó alakok határozzák meg. Ezt a lemondást példázza Donna Haraway *Stay-ing with the Trouble* című könyve: *Making Kin in the Chthulucene - különösen a critter* szó használatában, amely gondolkodásának középpontjában áll. Mint megjegyzi, ez a kifejezés a lények (és gépek) sajátosan széles skálájának helytartójaként szolgál: "Ebben a könyvben a "critter" szó a mikrobákra, növényekre, állatokra, emberekre és nem emberekre, sőt néha még a gépekre is

vonatkozik". (Haraway 2016: 169n1) Hasonlóan érvel Rosi Braidotti (2013) filozófus is, aki szerint minden faj egyformán sebezhető az antropogén klímaváltozás fenyegetéseivel szemben, ezért a fajok közötti együttműködést sürgeti, amelynek - mint erőteljesen kéri - a politikai napirend részét kellene képeznie.² Braidotti szerint itt az ideje, hogy az emberek új társadalmi kötelekeket hozzanak létre - nemcsak más fajokkal, hanem a techno-másokkal is, amelyeket hajlamosak vagyunk távol tartani és a működési funkciójukra redukálni. Szerinte csak így lehet majd biztosítani közös túlélésünket, mint az azonos fenyegetettséggel szembenező közösség.³ Amit Braidotti képvisel, az a szociális alapvető dedifferenciálódása, ami hasonló az ontológiai dedifferenciálódásához Haraway kritikusok definíciójában (vagy nem-definíciójában).

A dedifferenciálódás ilyen alakzatai, amelyek a poszt- és transzhumanizmus elméleti álláspontjának középpontjában állnak, és így igyekeznek elkerülni az antropocentrizmus szokásos vádját, szükségszerűen intuitív gesztusokkal társulnak - ez a megállapítás egyesíti a nyitottság és a terjeszkedés számos mozgalmát az ember-számítógép interakció dominanciájával szemben. A politikai tartózkodás hosszú szakaszai után a napirendre kerültek olyan kategóriák is, mint a respon-

ség, etika és részvétel. Mindennek a hangneme követelőző, szerénytelen és programszerű; ahogy azt Braidotti maga is elismeri, türelmetlen és aligha mentes a pátoztól. Hogy az egyéni részvételi offenzívák aggályai miként kapcsolódnak össze bizonyos elméleti megfogalmazásokkal, azt például Fredrik Aspling munkásságában láthatjuk. A svéd szociológus az antropocentrizmus elkötelezett kritikusa, és a poszt- és transzhumanizmus szoros szövetségésének tartja magát:

A nem emberi fajok fokozott bevonása a digitális technológia által támogatott interaktív kontextusokba, amelyet több faj és számítógép közötti interakcióként lehetne megfogalmazni, új interakciós lehetőségekhez és formákhoz vezet, és következképpen szükségessé teszi annak újragondolását, hogy ez mit jelent és mit jelenthet az interakció szempontjából. (Aspling 2015: 1)

A több fajra kiterjedő kölcsönhatás így a kölcsönhatás új fogalmának működési alapjává válik. Ezen az alapon Aspling napirendre tűzi a befogadás fogalmát, és arra ösztönzi az embereket, hogy vegyék figyelembe a különböző fajok sajátos igényeit és jellemzőit:

A nem emberi fajokkal való kiegészítés kihívást jelent a hagyományos interakciós megközelítések és elméleti keretek számára a HCI-ben. Az emberen túl kell gondolkodni, és szembe kell nézni a más fajok bevonásával járó kihívásokkal, amelyek eltérő megismeréssel, tapasztalatokkal, érzékekkel, képességekkel, időbeli ütemezéssel, igényekkel és szükségletekkel rendelkeznek. A további fejlődéshez megfelelő megközelítésekre és elméleti alapokra van szükségünk, hogy jobban megértsük az interakciók ezen új formáinak kialakuló dinamikáját. A nem

emberi fajokra fordított figyelem a HCI-ben (pl. az állatok, mint legitim felhasználók, akiknek és akikkel tervezni kell) a poszthumanizmus és az antropocentrizmus kritikájának analógiájára történik. (Ibid; Aspling et al. 2018)

Az ember-számítógép interakción túllépő és új ágenseket és folyamatokat integráló kérdés része annak, amit olyan fogalmakkal tárgyalnak, mint a poszt- és transzhumanizmus, valamint olyan új korszakos megnevezésekkel, mint az antropocén vagy a khthulucén (vö. Haraway 2016). Az ontológiai megnyílás gesztusai mellett, amelyek Haraway munkásságában kiemelkedő szerepet játszanak, a társadalmi interakciók megnyílásának gesztusai is jelen vannak tehát. Az ACI (animal-computer interaction), a PCI (plant-computer interaction), a CCI (child-computer interaction) vagy az RCI (robot-computer interaction) fejlődése is ezt képviseli. Az interakciós formák megfelelő részlegekre való felosztásának logikája éppúgy szembeötlő, mint a fajok közötti együttműködés mindenféle formájának előmozdításának szempontja. Interaktív re- láció érvényesül a különböző egyes felhasználói csoportokra vonatkozó ismeretek között. Ezek a kapcsolatok lehetővé teszik, hogy ezek a csoportok tanuljanak és profitáljanak egymásból: "A cél az összekapcsolt gondolkodás erősítése, miközben az ACI és a HCI, valamint azok részterületeinek ex-változtatható összekapcsolási módszerei kerülnek előtérbe.

beleértve a gyermek-számítógép interakciót (CCI) és az ember-robot interakciót (HRI)." (Hirskyj-Douglas et al. 2016: n. pag.; vö. Hourcade/Bullock-Rest 2011; Hourcade et al. 2018) Ezek az interaktív kapcsolatok és az egymástól való tanulás ezen aktusa ("megvitatva, hogy ezek a területek mit tanulnak egymástól, feltérképezve hasonlóságukat és különbségeiket") közös tervezési kritériumokat eredményez. Ez utóbbi kritériumok pedig szem előtt tartják a különleges vagy szélsőséges felhasználókat - mint gyermekeket, mint kognitív vagy érzékszervi korlátokkal rendelkező embereket, mint autistákat és így tovább (vö. Gennari et al. 2017; Eisapour et al. 2018; Lindsay et al. 2012; Satterfield et al. 2016).

A "részvételi tervezés" támogatói által megfogalmazott maximák közül több is szintaktikai sajátosság. Ma már gyakran találkozunk kettős prepozícióval ellátott kifejezésekkel; a speciális felhasználók előzetes bevonása érdekében például a programozókat most már arra ösztönzik, hogy *értük és velük együtt* dolgozzanak. A prepozícióknak ez a kettős használata fontos a mozgalom számára, ezért gyakran találkozunk vele. Figyelemre méltó a *with* prepozíció szokatlan használata is. Ebben a kontextusban gyakran a *becoming* szóhoz kapcsolódik, ami a posztstrukturalizmus egyik bővületének tárgya volt. Az "együtt-valóvá válás" (állatokkal, növényekkel, kövekkel) e módját, amely egyaránt érinti a molekuláris, valamint a technikai és mesterséges dolgokat, a globális túlélésért folytatott küzdelem kulcselemének tartják ("az állatokkal és tech- objektumokkal való együtt-valóvá válás szükségességét a túlélés érdekében" (Davis 2016: 210).

A macskák és a gyerekek - de a kihívásokkal, fogyatékkal vagy nagyon egyéni igényekkel rendelkező emberek is - a kifejezetten számukra készült

speciális interfészek megbecsült célcsoportjaivá váltak (vö. Maaß/Buchmüller 2018; Westerlaken/Gualeni 2016). Részvételük a komplexitás csökkentésén keresztül valósul meg - és ez, mint már megjegyeztem, egy olyan területen, amelyet egyébként a növekvő komplexitás gesztusai határoznak meg. Nem utolsósorban egy további gesztus határozza meg, amely a játékoság szisztematikus integrálásával jár; valójában az a benyomás marad, hogy a játékoság a napirend, és hogy maga a játék központi szerepet játszik a fajok közötti korlátok felszámolásában (vö. Nijholt 2015).⁴ Az ilyen intézkedések miatt szinte úgy tűnik, mintha az algoritmusok professzionalizálódását egy infantáziamozgalom kísérné - egyfajta ellenmozgásként. Ez a redukció gesztusaihoz kötődik, vagy legalábbis e képlet alapján értelmezhető. A formula programszerűsége annak a felfedezésnek köszönheti magát, hogy ahol a technológiáról beszélnek, ott egy másik narratíva is kifejezésre jut. A redukció mellett szóló ilyen érvekhez, amely funkcionális értelemben fogant és nem tiszteletlenül értendő, a HCI pri- mátusát felbontó és diverzifikáló helyszínnek felé kell tekintenünk. A gyermekekre való közelmúltbeli összpontosítás és az Aspling által említett erőfeszítések, amelyek a gyermek-számítógép interakciók és az ACI interakcióinak összekapcsolására irányulnak, tehát több mint pusztán tünet: Ezek egy sajátos gyakorlatot alakítanak ki. Atipikus szövetség-

2 Ez nem csak az interfészek, hanem az adatok és a hozzájuk kapcsolódó gyakorlatok kialakítására is vonatkozik (vö. Anderson et al. 2017).

es most válnak láthatóvá és lehetővé, amint az az ACI-ről szóló cikk következő címéből is kitűnik: "Cicákról és kölykökről: Reflections on Participatory Design with Small Animals and Small Humans". (Chisik/Mancini 2017) Ez a szervezett összpontosítás a gyerekekre és a macskákra mint az intuíció egy kívánt formájának képviselőire ex-emplifikálja a részvételi tervezés néhány aggályát. A cél egy olyan felhasználóbarát interfésztervezés létrehozása, amelyet nem kell fáradságosan elmagyarázni, hanem inkább intuitív, önmagyarozó és hallgatólagos tudásra épülő. A részvételi design mind az emberekkel, mind az állatokkal és a gépekkel tárgyal. A kis állatokra és az emberekre összpontosítva a csökkentést kézzelfoghatóvá teszi. Mi több, a redukciót a részvétel alapkövévé teszi.

3. Aszimmetriák

Az interakcióra, együttműködésre vagy kommunikációra szánt tervek természetessé tétele felületek használatát igényli, és operatív dimenziókkal rendelkezik (vö. Norman 2010). Így nem merül ki a dedifferenciálás gesztusaiban, hanem inkább stratégiai megfontolásokkal jár együtt. Ezek egyike a multiszenzoros felfedezés - vagy ahogy Caon és társai (2018) nevezték, a "multiszenzoros roham". Az érzékek megrohamozása mindenekelőtt a tapintás és a haptika egyre nagyobb mértékű felfedezése révén valósulhatott meg - egy olyan felfedezés, amelynek során maguk a számítógépekről és algoritmusokról való beszéd emberei is megváltoztak. A természetessé válás gesztusai, amelyeket az interfészekről szóló vitában leírtak és kritizáltak is, nemcsak a hardverrel való bánásmód problémáit érintik, hanem a programozás modorát is (vö. Bruns 1993; Hornecker 2008). Nemcsak a számítógépnek van szüksége masszív stratégiákra az érzékek befogadására; a programozási tevékenység is nyomás alatt áll, hogy új arculatban teremtsen meg önmagát. Fel kell hagynia kognitív szolipszizmusával, és a szimbólumokkal való pusztán munkán túl tapintható vállalkozássá kell válnia. Így ismét a test lesz a természetes garanciája egy olyan átfogó részvételi formának, amely erőfeszítés nélkül, intuitív módon, az önmegértés átlátható módjában valósulhat meg, vagy kellene, hogy megvalósuljon.

Nemcsak a számítógép-használat és az írástudás a kérdés, hanem egy olyan életvilág, amely al-

lows technológia létezését bármely adott környezeti formában. Erre mostanra már rengeteg példa van, és szisztematikus szinten ezek általában közösek bizonyos jellemzőkben, leginkább az új csatornák kialakulásában, a különböző érzékek integrálásában és a kommunikáció új formáinak megjelenésében. Ez utóbbiak aszimmetrikus kommunikációs jeleneteket szabadítanak fel - olyan jeleneteket, amelyek érvénytelenítik a kommunikációelmélet általános elképzeléseit. A példák bősége a háziállatok távoli gondozására szolgáló alkalmazásoktól a mesterséges környezetben termesztett növényekkel való interakcióig terjed (vö. Lee et al. 2006; Kuribayashi et al. 2007).

Megváltoztatják a társadalmiság formáit. E terület egyik elméletileg legambiciózusabb szereplője a japán kutató, Hill Hiroki Kobayashi.⁵ Célja, hogy túllépjen a kommunikáció és az interakció olyan paradigmáján, amelyet kizárólag az (értelmi képességeik teljes birtokában lévő) ember és a nyelvi kommunikáció egy bizonyos formája alapján mérnek. Kobayashi "ember-számítógép-bioszféra interakció" (HCBI) fogalma gyakorlatilag korlátlan működési területtel rendelkezik. Nemcsak a szereplők szféráját változtatja meg, hanem szükségszerűen a kommunikáció módját is: "A HCBI kiterjeszti az ember-számítógép interakció (HCI) tárgyát a megszámlálható emberekről, tárgyokról, háziállatokról és növényekről egy megszámlálhatatlan, összetett és nem nyelvi auditív bioszférára." (Kobayashi 2010:

n. pag.) Az antropocentrikus nézőpont elhagyása éppúgy pro- gramm, mint a megszámlálhatóság elvét elkerülő ágensekkel való együttműködés (vö. Kobayashi 2014). Ily módon az artikulált beszédet meghaladó lehetséges kifejezési formák központi szerepet kapnak. A viselhető eszközök használatával kapcsolatban, amelyek célja, hogy közelebb hozzák az embereket a természethez ("Wearable Forest-Feeling of Belonging to Nature" a cikke címe), Kobayashi a következőket írja: "Így a viselhető számítógépes rendszerek olyan intermédiummá váltak, amely a bioszférában élő különböző fajok tele- jelenlétét fejezi ki oly módon, hogy nem nyelvi kifejezésüket minden résztvevő érzékeli és megérti, ami a nyelvtudomány minden szabályát megszegi". (Kobayashi 2008: 1133)

Az ilyen alkalmazások helyszíne tehát közel van az élethez, és korántsem korlátozódik a művészeti installációkra. Hogy mennyire élethűek lehetnek, azt a LumiTouch nevű eszköz mutatja be. A LumiTouch első pillantásra úgy néz ki, mint egy közönséges képkeret. Az egyik feltűnésmentes keret az interneten keresztül kapcsolódik egy másikhoz, és képes olyan jeleket kiváltani, amelyek megfelelnek valakinek a puszta érintésének. Az érintés típusától (intenzitásától, gyakoriságától, időtartamától) függően különböző fényminták és színkonstellációk szabadulnak fel, amelyek egy indí- vualizált kódhoz társíthatók. Tervezői szerint ez utóbbi alkalmas egy sajátos kifejezési forma megvalósítására, és így ösztönzi egy privát érzelmi nyelv kialakulását (vö. Kaye/Goulding 2004). A LumiTouch megváltoztatja a kommunikációelmélet leegyszerűsített (mert idealizált) modelljeit, és a képkeret megérintésének aktusa hasznos előnyökkel jár a fogyatékkal élők számára. A de- signátorok az aszimmetrikus csere olyan formáit képzelik el, amelyekhez a kommunikációs partnerek egyikének nem kell kognitív vagy fizikai képességeinek teljes mértékben ura lennie: "Azok az emberek, akik hosszú ideig nem képesek aktívan kommunikálni (pl. betegek vagy idősek), használhatják a LumiTouch passzív átvitelét." (Chang et al. 2001: 314) A motoros képességek vagy a kognitív képességek túlterhelésének lehetősége bizonyos helyzetekben, például ha valaki ágyhoz kötött, ellensúlyozható a kevésbé megterhelő kommunikációs rendszerekkel: "Hasonló-

3 Lásd a honlapját a <http://hkhkobayashi.com> címen (elérés 20192., június) és Nijholt 2015.

ly felhasználók, akiknek nincs meg a szükséges kézügyességük vagy koncentrációjuk a számos gomb megnyomásához, értékelhetik ezt a rendszert a kis számú, egyszerű megragadható bemenet miatt." (Ibid.)

Ennek a növénytartó-alapú érzékelő eszköznek a kialakítása több fontos elemet tartalmaz, amelyek történetesen a fajok közötti kommunikáció és interakció szempontjából is központi szerepet játszanak. Az implicit és így tudattalan részvétel módjában a rendszer lehetővé teszi az emberek számára, hogy zökkenőmentesen és olyan módon vegyenek részt a távoli családtagok mindennapi életében, hogy azt ne érezzék kényszernek vagy zavarónak: "Ezeket a cserefolyamatokat úgy tervezték, hogy beleolvadjanak a felhasználó mindennapi életébe". (Ibid.) Három érzékelő segítségével különböző hangjelzéseket képes továbbítani, így a rendszer explicit üzenetek közvetítésére is használható. Ennél az explicit módnál azonban fontosabb a használatának implicit jellege, amely feltűnésmentes részvételével aláveti magát a média logikájának és annak, ahogyan egyre észrevétlenebbül beleolvad a környezetünkbe. Ezt ma már zökkenőmentesen, feltűnésmentesen, csendesén, simán és mindenütt teszik, és ezek a tulajdonságok újradefiniálják azokat a módokat, ahogyan az elméletalkotóknak általában véve a médiáról gondolkodniuk kell. Ez egy olyan médiaelmélet lenne, amely megszabadulva attól a paradigmatis elképzeléstől, hogy a média az ember kiterjesztése vagy szerves kivételése, ehelyett olyan jelzőkkel lehetne leírni, mint a mindenütt jelenlévő, *zökkenőmentes* és *nyugodt* (vö. Weiser/Brown 1996). Ez egy olyan médiaelmélet lenne, amely a kommunikáció bensőségesebbé tételével kapcsolatos kérdésekre összpontosít, és felöleli saját, bensőségesen feltöltött tárgyait. Ez a tendencia az affektusokkal feltölthető dolgok kifejlesztése felé csak erősödik. Ezt egy azonosítható napirend vezérli, és nem a

a kazuisztika által (vö. Choi et al. 2014; Kaye/Goulding 2004). Alapja - a redukció - olyan programmá vált, amelynek eredményei a mindennapi élet részévé válnak.

4. Finis (hominis)

A gyerekektől és a macskáktól eltekintve, mindezekből az derül ki, hogy milyen szükségletek és helyszínek vannak egy olyan kommunikációra és együttműködésre, amelyet úgy terveztek, hogy aszimmetrikus, de diszkriminációmentes legyen. A fent bemutatott alkalmazások nem a technikai környezetek kezelésének módjait kívánják optimalizálni, hanem inkább azt remélik, hogy alternatív és kevésbé összetett módokat kínálnak a használatukra (vö. Rieger 2019). Így a szemlélet is eltolódott a korábbi makacs, egy adott felhasználótípusra való irányultságtól (vö. Satchell/Dourish 2009). Zárásként két dolgot kell még elmondani: Először is, a gyerekek és a macskák, amelyeket itt egy szélesebb jelenség képviselőiként mutattam be, funkcionális használatba kerülnek. Amiről szó van, az nem egy olyan minimalizálási ajánlat, mint amilyen a féktelen macskatartalmakat meghatározza, hanem inkább a komplexitás stratégiailag követett csökkentésének funkcionális megfelelője. E törekvések között olyan elfogadási kampányok szerepelnek, amelyek a speciális felhasználókat is bevonják, és a tervezés fókuszát a részvétel minden lehetséges formája felé tolják el. Ez utóbbiak egyike a naturalizáció gesztusa (vö. Andreas et al. 2018).

A második pont arra a kérdésre vonatkozik, hogy ki irányítja a hálózatot. Erre a kérdésre első pillantásra egyszerű a válasz, és ennek semmi köze az észrevétlen algoritmusok erejéhez, hanem inkább az online tartalomhoz. Nem más, mint a Google Brain mélytanulási folyamatai hozták felszínre, hogy mennyiségi szempontból a macskák uralják az ott zajló eseményeket (vö. Guerin/Vasconcelos 2008). A projekten dolgozók nagy mulatságára algoritmsaikból kiderült, hogy valóban a macska az internet ura - egy olyan feltételezés, amellyel Alexander Pschera (2016) is eljátszik, bár valamivel kevésbé tréfásan, az "állatok internetének" szentelt cikkében. Az internet már jó ideje nem csak az embereké. Ez a helyzet most még olyan szójátékokban is tükröződik, amelyek - bármennyire is buták - mégis alátámasztják a részvételi tervezés etikai érveit: "Munkánk a kutatóársakra összpontosít, és magában foglalja, a *manco*s részvételi tervezést, a *labradoros* tesztekét és a kutyakamerás megfigyelést." (Mankoff et al. 2005: 253; vö. Trindade et al. 2015) Vagy éppen a macskák tekintetében: "A digitális média modern korában nehéz tagadni, hogy a macskák bekarmolták magukat az internet korszellemébe". (Myrick 2015: 175)

A cím, amelyet ennek az esszének választottam - "Csökkentés és részvétel" - szaván fogadja a más fajok és létformák bevonására vonatkozó követeléseket. Az ilyen követelések célja a cselekvőképességgel és episztemikus relevanciával rendelkezők körének bővítése. A többfajú közösségek új szereplőknek, a

kommunikáció és az együttműködés új formáinak, a tervezés és a részvétel új típusainak, új felelőségeknek és társadalmi formáknak adnak otthont: emberek és állatok, növények és kövek, műtárgyak és biofaktumok, gépek és médiumok, élő és élettelen, valós és virtuális, kiterjesztett és nem kiterjesztett, szimulált és modellezett, megnövelt és csökkentett (vö. Leistert 2017). Ezért csak következetes, hogy a szereplők e szférájában az algoritmusok talán nem találják meg a nyugalmukat, de minden bizonnyal megtalálják a helyüket.

Hivatkozások

- Anderson, Theresa Dirndorfer et al. (2017): "Adatjáték: Participatory Visualisation to Make Sense of Data" (Részvételi vizualizáció az adatok értelmezéséhez)." In: Proceedings of the Association for Information Science and Technology 54/1, pp. 617-18.
- Andreas, Michael et al. (2018): "Unterwachen und Schlafen: Einleitung". In: Michael Andreas et al. (szerk.), Unterwachen und Schlafen: Anthropophile Medien nach dem Interface, pp. 7-31.
- Angelini, Leonardo/Caon, Maurizio (2015): "Towards an Anthropomorphic Lamp for Affective Computing". In: Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction, New York: ACM, pp. 661-666.
- Aspling, Fredrik (2015): "Állatok, növények, emberek és a digitális technológia: Multispecies-Computer Interaction: Exploring and Understanding Multispecies-Computer Interaction". In: Proceedings of the 12th International Conference on Advances in Computer Entertainment Technology, New York: ACM, pp. 1-55.
- Aspling, Fredrik et al. (2018): "Az állatok megértése: A Critical Challenge in ACI." In: Proceedings of the Tenth Nordic Conference on Human-Computer Interaction, New York: ACM, pp. 148-60.
- Benneke, Johannes et al. (2018) (szerk.): Das Mitsein der Medien: Prekäre Koexistenzen von Menschen, Maschinen und Algorithmen, Paderborn: Wilhelm Fink.
- Braidotti, Rosi (2013): Cambridge: The Posthuman: Polity.
- Bruns F. Wilhelm (1993): "Zur Rückgewinnung von Sinnlichkeit: Eine neue Form des Umgangs mit Rechnern". In: Technische Rundschau 29, pp. 14-18.
- Bunz, Mercedes (2012): Die stille Revolution: Berlin: Wie Algorithmen Wissen, Arbeit, Öffentlichkeit und Politik verändern, ohne dabei viel Lärm zu machen: Suhrkamp.
- Caon, Maurizio et al. (2018): "Towards Multisensory Storming". In: Proceedings of the 2018 ACM Conference Companion Publication on Designing Interactive Systems, New York: ACM, pp. 213-18.
- Chang, Angela et al. (2001): "LumiTouch: LumiTouch: Egy érzelmi kommunikációs eszköz." In: Extended Abstracts on Human Factors in Computing Systems, New York: ACM, pp. 313-14.
- Chisik, Yoram/Mancini, Clara (2017): "Cicákról és kölykökről: Reflections on Participatory Design with Small Animals and Small Humans (Gondolatok a kisállatokkal és kisemberekkel való részvételi tervezésről)." In: Proceedings of the Conference 2017 on Interaction Design and Children, New York: ACM, pp. 753-56.
- Choi, Yongsoon et al. (2014): "Ring*U: A Wearable System for Intimate Communication Using Tactile Lighting Expressions (Egy viselhető rendszer

intim kommunikációhoz taktilis fénykifejezések használatával)". In: Proceedings of the 11th Conference on Advances in Computer Entertainment Technology, New York: ACM, n. pag.

- Davis, Heather (2016): "Molekuláris intimitás." In: James Graham et al. (szerk.), *Climates: Architecture and the Planetary Imaginary*, Zürich: Lars Müller, pp. 205-11.
- Eisapour, Mahzar et al. (2018): "Egy virtuális valósággyakorlat részvételi tervezése enyhe kognitív károsodással élő emberek számára". In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, New York: ACM, n. pag.
- Engemann, Christoph/Sudmann, Andreas (2018) (szerk.): *Machine Learning. Medien, Infrastrukturen und Technologien der Künstlichen Intelligenz*, Bielefeld: transcript.
- Foucault, Michel (2002 [1966]): *Foucault: A dolgok rendje*: New York: A humán tudományok archeológiája: Routledge.
- Gennari, Rosella et al. (2017): "The Participatory Design Process of Tangibles for Children's Socio-Emotional Learning". In: Simone Barbosa et al. (szerk.), *Végfelhasználói fejlesztés: 6th International Symposium*, Cham: Springer, pp. 167-82.
- Golbeck, Jennifer/Neustaedter, Carman (2012): "Pet Video Chat: Monitoring and Interacting with Dogs over Distance (Kutyák megfigyelése és interakciója a távolságon keresztül)". In: *Extended Abstracts on Human Factors in Computing Systems*, New York: ACM, pp. 211-20.
- Guerin, Frank/Vasconcelos, Wamberto (2008) (szerk.): *The Reign of Cats and Dogs: The Second AISB Symposium on the role of Virtual Creatures in a Computerised Society*, London: AISB.
- Hall, Gary (2013): Gary Gary: "Toward a Postdigital Humanities: Cultural Analytics and the Computational Turn to Data-Driven Scholarship". In: *American Literature* 85, pp. 781-809.
- Haraway, Donna J. (2016): *A bajban maradni: Making Kin in the Chthulucene*, Durham, NC: Duke University Press.
- Heller, Christian (2011): Heller, Heller: *Post Privacy: Prima leben ohne Privatsphäre*, München: Beck.
- Hirskyj-Douglas, Ilyena et al. (2016): "Ahol a HCI és az ACI találkozik". In: *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*, New York: ACM, n. pag.
- Hirskyj-Douglas, Ilyena et al. (2018): "Hét évvel a kiáltvány után: Irodalmi áttekintés és kutatási irányok az állat-számítógépes interakció technológiákhoz". In: *Multimodal Technologies and Interaction* 2/2, pp. 1-25.
- Homepage Kobayashi, Hill Hiroyuki: <http://hkhkobayashi.com> (hozzáférés: 2019. június 2., június 2.).
- Hornecker, Eva (2008): "Die Rückkehr des Sensorischen: Tangible Interfaces and Tangible Interaction". In: Hans Dieter Hellige (szerk.), *Mensch-Computer-Interface: Zur Geschichte und Zukunft der Computerbedienung*, Bielefeld: transcript, pp. 235-56.

- Hourcade, Juan Pablo/Bullock-Rest, Natasha E. (2011): "Univerzális kölcsönhatások: Kihívások és lehetőségek." In: *Interactions* 18/2, pp. 76-79.
- Hourcade, Juan Pablo et al. (2018): "Child-Computer Interaction, Ubiquitous Technologies, and Big Data". In: *Interactions* 25/6, pp. 78-81.
- Itoh, Yoshihiro et al. (2002): "TSUNAGARI' Communication: A családtagok közötti kapcsolat érzésének elősegítése". In: *Conference on Human Factors in Computing Systems*, New York: ACM, pp. 810-11.
- Jones, Nicola (2014): "A tanuló gépek." In: *Nature* 505, pp. 146-48.
- Kammerer, Dietmar/Waitz, Thomas (2015): "Überwachung und Kontrolle: Einleitung in den Schwerpunkt." In: *Zeitschrift für Medienwissenschaft* 13, pp. 10-20.
- Kaye, Joseph/Goulding, Liz (2004): "Intim tárgyak". In: *Proceedings of the Fifth Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*, New York: ACM, pp. 341-44.
- Kelty, Christopher (2016): "Participation." In: Benjamin Peters (szerk.), *Digitális kulcsszavak: A Vocabulary of Information Society and Culture*, Princeton, New Jersey: Princeton University Press, pp. 227-41.
- Kobayashi, Hill Hiroki et al. (2008): "A természethez való tartozás érzése. In: *Proceedings of the 16th ACM Conference on Multimedia*, New York: ACM, pp. 1133-34.
- Kobayashi, Hill Hiroki (2014): "Ember-számítógép-bioszféra kölcsönhatás: A humán-központú interakción túl." In: Norbert Streitz/Panos Markopoulos (szerk.), *Distributed, Ambient, and Pervasive Interactions: Proceedings of the Second International Conference, DAPI 2014*, Cham: Springer, pp. 349-58.
- Kobayashi, Hill Hiroki (2010): *Doktori disszertáció: Tokyo University*, n. pag.
- Kortum, Philip (2008) (szerk.): *HCI Beyond the GUI: Design for Haptic, Speech, Olfactory, and Other Nontraditional Interfaces*, Oxford: Elsevier Science.
- Kuribayashi, Satoshi et al. (2007): "Plantio: Interaktív edény a növények kifejezéseinek bővítésére." In: *Proceedings of the International Conference on Advances in Computer Entertainment Technology*, New York: ACM, pp. 139-42.
- Lee, Shang Ping et al. (2006): "A Mobile Pet Wearable Computer and Mixed Reality System for Human-Poultry Interaction through the Internet (Egy hordozható számítógép és vegyes valóság rendszer az ember és a baromfi közötti interakcióhoz az interneten keresztül)." In: *Personal and Ubiquitous Computing* 10/5, pp. 301-17.
- Leister, Oliver (2017): "Social Bots als algorithmische Piraten und als Boten einer techno-environmentalen Handlungskraft". In: Robert Seyfert/Jonathan Roberge (szerk.), *Algorithmskulturen: Über die rechnerische Konstruktion der Wirklichkeit*, Bielefeld: transcript, pp. 215-34.
- Lindsay, Stephen et al. (2012): "Empátia, részvételi tervezés és a demenciában szenvedő emberek". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York: ACM, pp. 521-30.

- Maaß, Susanne/Buchmüller, Sandra (2018): "A kulturális szondák döntő szerepe az idősek számára és az idősekkel való részvételi tervezésben". In: *i-com* 17/2, pp. 119-35.
- Mancini, Clara (2017): Cincini: "Towards an Animal-Centred Ethics for Animal-Computer Interaction." (Az állatközpontú etika felé az állat és a társas interakcióban). In: *International Journal of Human-Computer Studies* 98, pp. 221-33.
- Mancini, Clara (2011): "Állat-számítógép interakció (ACI): Az interaktív-interaktív-interaktív kapcsolat: Változó perspektíva a HCI, a részvétel és a fenntarthatóság terén." In: *Interactions* 18/4, pp. 60-73.
- Mancini, Clara (2018): Állat-számítógép interakció (ACI): "Animal-Computer Interaction (ACI): A Manifesto." In: *Interactions* 18/4, pp. 60-73.
- Mankoff, Demi et al. (2005): "A fajok közötti tudatosság támogatása: Using Peripheral Displays for Distributed Pack Awareness (Perifériás kijelzők használata az elosztott csomagtudatossághoz)". In: *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology*, New York: ACM, pp. 253-58.
- Mann, Steve et al. (2015): (A felügyelet egy féligazság): "Declaration of Veillance (A felügyelet egy féligazság)". In: *2015 IEEE Games Entertainment Media Conference*, Piscataway, New Jersey: IEEE, pp. 1-2.
- Mann, Steve (2016): Mann: "Surveillance (Felügyelet), Sousveillance (Alulnézet) és Metaveillance (Maga a látás)". In: *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Piscataway, New Jersey: IEEE, 2016), pp. 1-10.
- Miyajima, Asami et al. (2005): "Tsunagari-kan' Communication: Távközlési környezet kialakítása és gyakorlati tesztelés külön élő családtagokkal". In: *International Journal of Human-Computer Interaction* 19, pp. 253-76.
- Myrick, Jessica Gall (2015): "Emotion Regulation, Procrastination, and Watching Cat Videos Online: Who Watches Internet Cats, Why, and to What Effect?" (Ki nézi az internetes macskákat, miért és milyen hatásra? In: *Computers in Human Behavior* 52, pp. 168-76.
- Nijholt, Anton (2015) (szerk.): *More Playful User Interfaces: Interfaces that Invite Social and Physical Interaction*, Szingapúr: Springer.
- Norman, Donald A. (2010): "A természetes felhasználói felületek nem természetesekek". In: *Interactions* 17/3, pp. 6-10.
- Pongrácz, Alexandre et al. (2016): "Egy kutya a Skype segítségével". In: *Proceedings of the Third International Conference on Animal-Computer Interaction*, New York: ACM, n. pag.
- Pschera, Alexander (2016): "Das Internet der Tiere: Natur und 4.0die conditio humana." In: *Zeitschrift für Medien- und Kulturforschung* 7, pp. 111-24.
- Rehman, Arshia et al. (2019): "Automatikus vizuális jellemzők az író azonosításához: A Deep Learning Approach." In: *IEEE Access* pp7., 17149-57.
- Reißmann, Ole (2011): Spackeria: "Internet-Exhibitionisten 'Spackeria':

'Privatsphäre ist so was von Eighties'. " In: Spiegel Online március 10.,
<http://www.spiegel.de/netzwelt/netzpolitik/internet-exhibitionisten-spackeria-privatsphaere-ist-ist->.

- sowas-von-eighties-a-749831.html (interjú Julia Schrammal; hozzáférés: 2019. május 26.).
- Rieger, Stefan (2019): "Virtuális humán tudományok". In: Dawid Kaszproicz/Stefan Rieger (szerk.), *Handbuch Virtualität*, Berlin: Springer.
- Rieger, Stefan (2019): *Die Enden des Körpers: Versuch einer negativen Prothetik*, Wiesbaden: Springer.
- Rieger, Stefan (2018): "Alles, was zählt: Megfigyelések egy számszerűsített szelfi által". In: *Psychosozial* 152, pp. 47-56.
- Ritvo, Sarah E./ Allison, Robert S. (2014): "A nem-emberi ember-számítógép interakcióval kapcsolatos kihívások: Liking". In: *Proceedings of the 2014 Workshops on Advances in Computer Entertainment*, New York: ACM, n. pag.
- Rodriguez, Catherine S. et al. (2019): "Kétlépcsős mélytanulási megközelítés a képzőművészeti festmények osztályozásához". In: *IEEE Access* pp7., 41770-81.
- Satchell, Christine/Dourish, Paul (2009): "Beyond the User: Use and Non-Use in HCI". In: *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group*, New York: ACM, pp. 9-16.
- Satterfield, Debra et al. (2016): "An Analysis of Data Collection Methods for User Participatory Design for and with People with Autism Spectrum Disorders". In: Marcus Aaron (szerk.), *Design, felhasználói élmény és használhatóság: Design Thinking and Methods*, Cham: Springer, pp. 509-16.
- Seaman, Bill (2011): *Neosentience*: Bristol: The Benevolence Engine, Bristol: Intellect.
- Seyfert, Robert/Roberge, Jonathan (2017) (szerk.): *Algorithmuskulturen: Über die rechnerische Konstruktion der Wirklichkeit*, Bielefeld: transcript.
- Stahl, Bernd Carsten (2014): "Participatory Design as Ethical Practice: Fogalmak, valóság és feltételek." In: *Journal of Information, Communication and Ethics in Society* 12, pp. 10-13.
- Stark, Emily et al. (2018a): "Medicine Has Gone to the Dogs: Deep Learning and Robotic Olfaction to Mimic Working Dogs" (Mélytanulás és robotikus szaglás a munkakutyák utánzására). In: *IEEE Technology and Society Magazine* 37/4, pp. 55-60.
- Stark, Emily et al. (2018b): "Odorveillance and the Ethics of Robotic Olfaction" (Szagmegfigyelés és a robotikus szaglás etikája). In: *IEEE Technology and Society Magazine* 37/4, pp. 16-19.
- Sudmann, Andreas (2017): Sudmann: "Deep Learning als dokumentarische Praxis". In: *Sprache und Literatur* 48/2, pp. 155-70.
- Trindade, Rui et al. (2015): "Tökéletes bűnözés: Exploring Animal Computer Interaction through a Digital Game for Humans and Cats." (Állati számítógépes interakción feltárása egy digitális játékon keresztül emberek és macskák számára). In: *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, New York: ACM, pp. 93-63.

van Dijck, José (2014): "Datafication, Dataism and Dataveillance: Big Data a tudományos paradigma és az ideológia között". In: *Surveillance & Society* 12/2, pp. 197-208.

- Weiser, Mark/Brown, John Seely (1966): "A nyugodt technológia eljövendő korszaka". In: Xerox PARC október 5., <https://calmtech.com/papers/coming-age-calm-technology.html> (hozzáférés: 2019. június 4.).
- Westerlaken, Michelle/Gualeni, Stefano (2016): "Becoming With: Towards the Inclusion of Animals as Participants in Design Processes". In: Proceedings of the Third International Conference on Animal-Computer Interaction, New York: ACM, n. pag.

A mesterséges intelligencia politikai vonatkozásai

Dan McQuillan

Bevezetés

A mesterséges intelligencia radikális politikájára van szükségünk, azaz a mesterséges neurális hálózatok politikájára. A mesterséges intelligencia politikai technológiaként működik, de a jellemzésére tett jelenlegi erőfeszítések az etikáról szóló liberális kijelentések formáját öltik. A mesterséges intelligenciában az elfogultság kérdéseit a méltányosság kérdéseiként kezelik, mintha a társadalom már most is egy egyenlő játéktér lenne, amelyet csak fenn kell tartani. Az átláthatóságot és az elszámoltathatóságot elegendőnek tekintik a mesterséges intelligencia problémáinak kijavításához (ACM FAT 2019), mintha jól működő és valóban demokratikus politikába vezetnék be. A mesterséges intelligenciát politikai foglalkozásként való szemmel látható elutasítása a megszorítások megoldásaként való népszerűsítésével szemben áll. Az Egyesült Királyságban például az alulfinanszírozott állami egészségügyi rendszerrel szóló vitákat olyan magas szintű nyilatkozatok fűszerezik, hogy "az AI lehet az, ami megmenti az NHS-t" (Ghosh 2018). A megszorítás nem természeti katasztrófa, hanem politikai döntés a pénzügyi intézeteknek a közkiadások rovására történő támasztására. E döntéshozók reménye az, hogy a gépi gondolkodás megoldja a rejtélyt, hogy a növekvő szükségleteket meredeken csökkentett erőforrásokkal kezeljék. Eközben a tényleges mesterséges intelligencia működési jellemzőinek más politikai hatásai is vannak, mint például a tisztességes eljárás megszüntetése. A backpropagation által végrehajtott hatalmas párhuzamos iterációk átláthatatlanságot vetnek a mesterséges intelligenciára, mivel optimalizálásait nagyon nehéz visszavezetni az emberi érvelésre (Lipton 2016). A fontos társadalmi és politikai döntéseket befolyásoló algoritmikus ítéletek így kikerülnek a diskurzusból.

A mesterséges intelligencia politikai dimenzióit nem lehet megsejteni az absztrakt filozófiai etika nem oldja meg. Konkrét technikai műveletekből, például a vektorok feletti összegzésből adódnak, konkrét társadalmi körülmények között. Az etikus mesterséges intelligencia eszméje egy olyan információs művelet, amelynek célja az algoritmikus hatásokkal kapcsolatos közvélemény féltrevezése lecsillapítása és piaci előnyökkel való pozicionálása (Hern 2018). A mesterséges intelligencia valódi veszélyei akkor bukkannak fel, amikor keveredik korunk politikai áramlataival. A mesterséges intelligencia politikai összefonódását jól mutatja egy, a közelmúltbeli Világgazdasági Fórumon készült fénykép, amelyen a

populista és szélsőjobboldali brazil politikus, Jair Bolsonaro ül ebéd közben az Apple vezérigazgatója, Tim Cook és a Microsoft vezérigazgatója, Satya Nadella között (Slobodian 2019). A mesterséges neurális hálózatokra azért van kereslet, mert a nagy mennyiségű adat és a feldolgozási teljesítmény összefolyása, a GPU-k formájában, a

lehetővé tette számukra, hogy olyan területeken, mint a képfelismerés, háttorzongató eredményeket érjenek el. A mesterséges intelligencia nagykorúvá válásának éveit azonban egyúttal a neoliberális válság és a szélsőjobbpoli politika globális térnyerésének éveit is. A sürgető kérdés az, hogy a neurális hálózatok konkrét technikai műveletei hogyan erősítik, erősítik vagy terjesztik ki ezeket a politikai áramlatokat, és hogyan szolgálhatják (ha egyáltalán) a társadalmi igazságosság céljait.

Határok

A mélytanulás vonzereje abban rejlik, hogy képes előrejelzéseket készíteni az adatok túláradásából. A rétegek súlyait olyan iterációkkal optimalizálják, amelyek a veszteségfüggvényt a minimumba vezetik, míg az olyan aktiválási függvények, mint a ReLU, kíméletlenül hatnak minden egyes neuronra, hogy eltávolítsák a gyengébb jeleket (3Blue1Brown. n.d). Az általános hatás a statisztikai bizonyosság előállítása; a súlyok hálója, amely a kusza bemeneti adatokat egyértelmű osztályozásokká alakítja át. Ez úgy történik, hogy korrelációkkal helyettesítjük az oksági mechanizmusok megállapítására tett kísérleteket, és ezt nem korlátozza semmilyen tágabb konzisztenciakeret. A mechanizmusban nincs a "józan ész" eleme, amely különbséget tesz a világról szerzett megtestestült tapasztalatokon alapuló sejtések között. A neurális hálózatok neoplatonikusak; a világ rejtett matematikai rendjét állítják, amely a közvetlen tapasztalatnál magasabb rendű (McQuillan 2017). A politika abban lép be, ahogyan ezek a rendezések tágabb mechanizmusokba vonódnak be. Ahelyett, hogy az emberi következményekkel való szélesebb körű törődés alapján korlátozná a statisztikai tekintélyt, a mesterséges intelligencia átvételének jelenlegi versenyét az hajtja, ahogyan annak céltudatos optimalizálása rezonál a hatékonyság vagy a részvénytulajdonosok értékének maximalizálására irányuló intézményi célokkal. A mesterséges intelligencia működése összhangban van azzal a neoliberalizmussal, amely a világot a piaci mechanizmusba bevitt inputok atomizált halmazaként fogja fel, amely szükségszerűen optimális eredményt fog produkálni.

A mesterséges intelligencia matematikai regresszióinak célja a döntési határok meghúzósa,

hogy egy bemenet egyértelműen az egyik vagy a másik oldalra sorolható legyen. Ennek összekapcsolása a külső világ kockázati kérdéseivel, még akkor is, ha a célok feltételezhetően jóindulatúak, a mesterséges intelligenciát az irányítás rendszerévé teszi. A számításalapú kategorizálás a gépi és emberi döntések láncolatát indítja el, amelyek valós következményekkel járnak, és az erőforrások vagy lehetőségek elosztását vagy elvonását vonják maguk után. Ezek a számszerű ítéletek mély tanulási folyamatokba ágyazva, a megfelelő eljárástól vagy diskurzustól elfedve törvényszerű erővel bírnak, anélkül, hogy a törvényekből fakadnának. Így a mesterséges intelligencia előrejelzési határai

folyamatos részleges kivételes állapotokként jelennek meg (Agamben 2005). A gépi ész számításait saját feltételeik szerint vitatni képes ex- pertise néhány vállalatnál és egyetemen erősen centralizálódott. A többiek számára a gépi tanulás számítási tekintélye olyan helyzetekhez vezet, ahol a személyes tanúságtétel leértékelődik a számítási felismerésekkel szemben. A mesterséges intelligencia az epizodikus igazságtalanság motorjává válik, olyan felismeréseket követelve, amelyek felülírják a megélt tapasztalatokat (Fricker 2007).

A bürokrácia

A huszadik századi bürokráciához hasonlóan a mesterséges intelligencia is arra készül, hogy a vállalatok és a kormányzat legitimációjának egységesítő logikájává váljon. Jelenleg a mélytanulás teljesítménye arányos a felhasznált számítási teljesítmény nagyságával: a 2012-es AlexNet képfelismerő áttörés és a Google DeepMind rendszere között, amely legyőzte a Go nagymestert, a szükséges számítási teljesítmény 300 000-szeresére, mintegy 2000 Petaflops/s-napra nőtt (Amodei/Herandez 2018). A mélytanulás hardveres és szoftveres csővezetékei stratégiai jelentőségűvé válnak, és a meglévő példányok, például az Amazon Web Services egyre kevésbé különböztethetők meg a kritikus nemzeti infrastruktúrától (Konkel 2016). De bár a mesterséges intelligencia a szerverfarmok elkerített anonimitásában materializálódik, a gondolkodás és a cselekvés között rejlik a kihatása. A társadalmi döntésekre alkalmazott mélytanulás a Bourdieu-féle habitus konkrét megnyilvánulása lesz; a strukturált struktúrák hajlamosak arra, hogy strukturáló struktúráként működjenek (Bourdieu 1990). Nem arról van szó, hogy a kulcsfontosságú döntéseket gépekre bizzák anélkül, hogy az ember részt venne a folyamatban; sokkal inkább arról, hogy a kényszerhelyzetben döntéseket hozó emberek olyan empirikus kockázati rangsorokat kapnak, amelyek levezetését nem tudják megkérdőjelezni. A mesterséges intelligencia a Hannah Arendt által leírt értelemben vett gondolatlanságra ösztönöz; az utasítások kritikájának képtelenségére, a következményekre való reflexió hiányára, a helyes utasítás végrehajtásában való hit iránti elkötelezettségre (Arendt 2006).

A jósláson keresztül ez a rendezés kiterjeszti a bürokratikus kormányzást arra, hogy a szándék vagy tendencia területe, amelyet szolgáltatásként vagy beavatkozásként igyekszik megelőzni. A bűncselekményt megelőző osztályozások és a "szélsőségességet megelőző" formák elterjedése mellett, ahogyan azt az Egyesült Királyság Prevent Stratégiája (Sian 2017) prototípusa mutatja, lesznek olyan osztályozások is, amelyek jóindulatot és hatékony erőforrás-elosztást követelnek, mint például a cukorbetegség vagy a demencia előtti osztályozás (LaMattina 2016). Az elővigyázatosságra való törekvés társadalmi félelmeket és piaci érdekeket foglal magába azzal a céllal, hogy kiküszöbölje azt, ami nem kívánatos. A probléma az, hogy a mesterséges intelligencia redukcionista. Csak a kontextus azon aspektusaiból képes tanulni, amelyek matematikailag is megragadhatók, és egyetlen célt kap, amelyre optimalizálnia kell. Ezért a mesterséges intelligencia arra irányuló kísérlete, hogy megmagyarázza, mi történik, az egész rendszert bizonyos alkotóelemekre és azok kölcsönhatásaira redukálja. Ráadásul a társadalmi kérdésekre alkalmazott prediktív mélytanulás azt feltételezi, hogy a tulajdonságok individualizáltak és veleszületettek, miközben elhomályosítja a közös társadalmi okok hátterét. Ez egy apartheid bürokráciát terjeszt ki az élet minden olyan aspektusára, amelyet adatok érintenek.

Instabilitás

A mesterséges intelligencia nem alkalmazható a komplex társadalmi és kulturális problémák egy részére, a rendkívül szűk és korlátozott szempontokon kívül. Ez nem csak azért van így, mert a működésmódja átgondolatlanságra és redukcionizmusra ösztönöz, hanem azért is, mert a mélytanulás a szó szoros értelmében túlmutat a társadalmi és politikai komplexitáson. Lényegében egyszerűen csak egy mintakeresési technika, amely meglepően jól működik az észlelési osztályozásban és néhány más, jól körülhatárolt alkalmazásban, például a játékokban. Azonban még a mesterséges intelligencia eme szívtérületein is vannak szisztematikus problémákra utaló jelek. Számos ellenpélda van arra, hogy egy képhez gondosan megválasztott zaj hozzáadása, amely az emberi érzékelés számára nem tűnik többnek jelentéktelen fehér pontok elszórásánál, arra kényszeríthet egy neurális hálózatot, hogy tévesen osztályozzon egy nyilvánvaló képet (Goswami et al. 2018). Talán még ennél is jelentősebb, hogy egy nemrégiben megjelent tanulmány szerint amélytanulás képfelismerése gyakran összeomlik, amikor a háromdimenziós térben szokatlan helyzetbe elforgatott, gyakori ingerekkel szembesül. Az egyik példában a hálózat helyesen ismer fel egy iskolai teherautót, de amikor egy valós képet lát egy ilyenről az oldalára fordítva, tévesen hókotróznak minősíti azt. A szerzők arra a következtetésre jutottak, hogy bár a mély neurális hálózatok jól működnek a képosztályozásban, a valódi tárgyfelismeréstől még messze vannak, és a tárgyról alkotott képük meglehetősen naiv. Következtetésük az, hogy "a mély neurális hálózatok (DNN-ek) nem képesek általánosítani az OoD (out-of-distribution) bemenetekre, beleértve a természetes, nem-adverzális bemeneteket, amelyek gyakoriak a valós környezetben" (Alcorn et al. 2018). Ez egy fontos, de aligha meglepő megfigyelés. Egyetlen neurális hálózatnak sincs semmiről olyan megértése, egy absztrakt modell vagy ontológia formájában, amelyet szabadon alkalmazhatunk új helyzetekre. Vagyis a neurális hálózatok pontosan arra a fajta adaptív és analógiás gondolkodásra nem képesek, ami még a kisgyermeket is jellemzi. Vezető AI mérnökök kijelentései, miszerint a neurális hálózatok "most vagy a közeljövőben" képesek lennének "bármilyen mentális feladatot" elvégezni, amit egy ember "kevesebb mint egy másodperc gondolkodással" (Ng 2016), nemcsak nevetségesek, hanem valójában veszélyesek. Ha amélytanulás nem képes felismerni tárgyakat nem-kanonikus pózokban, akkor ne várjuk el tőle, hogy a mindennapi, józan ész szerinti gondolkodást végezze, egy olyan feladatot, amelyre eddig soha semmilyen képességet nem mutatott. Még kevésbé szabadna alkalmaznunk azt zűrös társadalmi-politikai kontextusokban, és elvárunk, hogy olyan felismeréseket vonjon le belőle, amelyeket korábban a diskurzusra bízunk.

Azonban nem csak azért kell mélyen tartanunk, mert nem vagyunk a mélyben. a szociálisan érzékeny helyzetekből való tanulás. A céltudatos optimalizálás, amely miatt a mesterséges intelligencia olyan jól rezonál a neoliborális szemléletre, végzetes etikai terhet hordoz magában. A hasznossági függvények,

mint például a mélytanulás backpropagationje, etikai szempontból mély vízbe kerülnek, amikor független, redukálhatatlan célok vannak, amelyeket egyszerre kell követni. Az etikusoknak vannak olyan tételeik, amelyek szerint lehetetlen, hogy egy optimalizálás jó eredményt hozzon létre egy populáció számára anélkül, hogy megsértené az alábbiakat

etikai szándékainkat. Például a pusztá összeadás paradoxona azt mutatja, hogy ha egy társadalmi jóléti függvényt a boldog emberek bármely populációjára optimalizáljuk, akkor ex- isz egy sokkal nagyobb, nyomorúságos életet élő populáció létezik, amely "jobb" (a teljes jólét szempontjából optimalizáltabb), mint a boldog populáció (Eckersley 2018). Nem meglepő módon ezt a paradoxont visszataszító következtetésnek is nevezik. Bár úgy tűnhet, hogy ez az erkölcsfilozófia elvont, analitikus logikájából ered, ne feledjük, hogy éppen ez a lényeg: az intézményesített neurális hálózatokon keresztül egy elvont és számító logikát alkalmazunk a társadalmi világra. A hasonló etikai érvelés egy egész sor kellemetlen paradoxont eredményezett, mint például a "szadista következtetés" és a "nagyon antiegalitárius következtetés". Ezek alapvető összeegyeztethetlenséget sugallnak a különböző utilitarista célok, például a teljes jólét maximalizálása, az átlagos jólét maximalizálása és a szenvedés elkerülése között. A mesterséges intelligenciában alkalmazott célfunkció meggondolatlan követése etikailag mérgező következményekhez vezet, még akkor is, ha az eredeti funkció látszólag jóindulatú, nem is beszélve arról, ha az a kapitalista profítcélt szolgálja.

A lehetséges

Az átgondolatlanság szintén a mesterséges intelligencia megoldásához vezető út elején lép be. A mesterséges intelligencia mindig a Bergson által "kész problémának" nevezett problémák megoldásának szolgálatában áll. Azaz a gépi tanulást olyan problémákra alkalmazzák, amelyek olyan nem vizsgált feltételezéseken alapulnak, mint a kulturális előítéletek és intézményi célok, valamint azok a mélyebb előítéletek, amelyek magába a nyelvbe ágyazódnak. A kész problémával az a probléma, hogy a lehetséges megoldások egy sorát feltételezi, amelyek egybeesnek a probléma adott kifejezésével. Bergson úgy érvelt, hogy ha valaki elfogad egy kész problémát, "éppúgy mondhatná azt is, hogy minden igazság gyakorlatilag már ismert, hogy modellje az állam adminisztratív irodáiban van szabadalmaztatva, és hogy a filozófia egy kirakós játék, ahol a probléma az, hogy a társadalom által adott darabokból megépítsük azt a tervet, amelyet nem hajlandó megmutatni nekünk". (Henri Bergson, *La pensée et le mouvant*, idézi Solhjdju 2015). A cselekvőképesség, az adott valóság megváltoztatására való képesség ehelyett a probléma megtalálásának és felvetésének kérdése. Ez azért más, mert "a probléma megállapítása nem egyszerűen feltárást jelent, hanem feltalálást". Solhjdju szerint Isabelle Stengers ezt úgy fejezi ki, mint "a lehetséges és a valószínű közötti különbséget", ahol a valószínű az, "ami a valóságoshoz képest csak egyetlen dolog, a létezés hiányzik" (Solhjdju 2015). Ez már megkonstruálható ugyanazzal a fogalmi állványzattal, amelyből a problémát felépítettük, és a megfejtése egyszerűen valószínűségi dedukció kérdése. A lehetséges ezzel szemben valami kiszámíthatatlan és nem

kiszámítható dologról szól; egy kreatív aktusról, amely nem pusztán a meglévő igazságok átrendezése. A tét nem a jelenlegi mesterséges intelligencia valószínűsége, hanem a politikai gondolkodás és cselekvés lehetősége. Mi

úgy kell megközelítenünk a mesterséges intelligenciát, hogy a valószínűségekkel szemben a lehetséges oldalára álljunk.

A gépi tanulás lehetőségeinek újratemtése azt jelenti, hogy a programmatika és a politika nem osztható tényezőként dolgozik, mérnöki problémákat old meg, miközben továbbra is a társadalmi hatásokra összpontosít. Ez egyszerre igényel matematikai szintű precizitást és nyitottságot a különböző lehetséges valóságok iránt, amelyeket fel lehet fogalmazni. Egy ilyen tudományág egyik megközelítését a tudomány feminista modellje kínálja, mint amelyet Roy, Harding és Spanier (Roy 2004) ír le. Vagyis a tudományos módszertan kiterjesztett formája, amely magában foglalja a probléma keletkezését és a vizsgálat célját. Azoknak, akik nem elnyomó gépi tanulást akarnak kifejleszteni, nem szabadna egy problémát adottnak elfogadniuk, hanem a probléma eredetének, vagyis a strukturális erőknél a felvetését és a pri- oritását meghatározó tényezőknek a felkutatásával kellene kezdeniük. Egy vizsgálat céljainak feltárása a mélytanulással azt jelenti, hogy túl kell lépni azon, hogy a hiperparaméterek optimalizálásával pontosan megjósoljuk a validációs adatokat. Azt jelenti, hogy ezt a szűk technikai célt egy szélesebb körű hatásrendszer részeként kell megértenünk, fel kell tennünk a kérdést, hogy kinek a céljait szolgálja, kit zárhat ki, és hogyan hat a társadalom tágabb értelemben vett jólétére. A mesterséges intelligencia szempontjából talán a legradikálisabb, hogy a feminista megközelítés kapcsolatot hoz létre a kutató és a kutatás tárgya között, ami megköveteli, hogy céltudatosan félretegyük a báméskodó tudatosságot, amely a mesterséges intelligencia önhittségét táplálja. A feminista módszer gyakorlati megvalósításának legközvetlenebb módja az emberi tanulással kapcsolatban a kollektív kutatási struktúrákon keresztül, amelyek bevonják a "célcsoportot" a vizsgálat folyamatába, olyan struktúrákon keresztül, mint például az emberek tanácsai (Mc- Quillan 2018). Az ilyen szituált vizsgálati kollektívák jó helyzetben vannak ahhoz, hogy a valószínűség zsarnokságából f fényvonalakként újra feltalálják a problémát.

Politikai fellépés

Ezt az alternatívát a megdölgdölgdölgdölgd mesterséges intelligenciával rezonáló politikai áramlatokkal szemben kell megteremtenünk. Ez nem lesz könnyű feladat. A mesterséges intelligencia "megszorítások alatti mesterséges intelligenciaként", azaz a neoliberalizmus saját válságára adott válaszáként valósul meg. Az olyan mindennapi kegyetlenségeket, mint a fogyatékkal élőkkel szembeni jóléti megszorítások, egyre inkább elfedik a gépies osztályozások (Alston n.d.). A neurális hálózatok az episztemikus igazságtalanság és a részleges kivételes állapotok motorjaivá válhatnak. Ami még veszélyesebb, hogy a társadalmi problémák redukcionista érvelésen és in- natív tulajdonságokon alapuló optimalizálásra való leegyszerűsítése pontosan a

populista szélsőjobboldal politikáját tükrözi. A mesterséges intelligenciával kapcsolatos állítások ellentmondásosságára való rámutatásnak nincs vonzereje ezzel a politikai irányzattal szemben. Az ostobaságnak és a gyűlöletnek nincs szüksége filozófiai következetességre, csak olyan operatív hatékonyságra, amely előadja a kegyetlenség ideológiai színházát. A mesterséges intelligencia gyakorlatának olyan politikát kell kialakítania, amely ellenáll a tekintélyelvűségnek, és a közös emberségünkkel való törődést erősíti.

A mesterséges intelligencia kollektív gyakorlatának szükségessége tehát nemcsak ismeretelméleti, hanem politikai szükségszerűség is. A néptanács és a közgyűlés politikai formái visszahelyezhetik a megfelelő eljárás és az igazságosság kérdéseit az őket megillető helyre a diskurzusban. Míg az algoritmikus hatalom az általánosított absztrakciók privilegizálásából ered, a közvetlen demokrácia a helyhez kötött ismeretek mozgósításával erősíthető meg újra. Ezeket a kommunikációs technológiákra alkalmazható formákba kell átültetni. Ivan Illich a konviviális technológiára való felhívásában az "ellenfólia-kutatást" javasolta, amelynek célja "a gyilkos logika kezdeti szakaszainak felderítése egy eszközben" (Illich 1975), ahol az eszköz Illich számára a technológiák és intézmények egy meghatározott kombinációját jelenti. Az ellenfólió-kutatás egy plauzibilis programot fogalmaz meg az AI néptanácsok számára: "tisztázzák és dramatizálják az emberek és eszközeik kapcsolatát", "folyamatosan a nyilvánosság elé tárják a rendelkezésre álló erőforrásokat és használatuk különböző módokon történő következményeit", és azonosítják "az emberek azon osztályait, akiket az ilyen tendenciák a legközvetlenebbül sújtanak". Ez nem negatív, hanem pozitív program, hogy olyan feltételeket teremtsen, ahol az emberek képesek önálló cselekvésre a mások által legkevésbé ellenőrzött eszközökkel. A cél az, hogy megfelelő határokat találjunk eszközeink számára. Az eszközök korlátozása a gyülekezések mechanizmusán keresztül azt is megteremti, amit Hannah Arendt a cselekvés tereiként azonosított, ami csak a személyes találkozásból ered, és ami "a statisztikai törvények és azok valószínűségeinek elsőpró esélyeivel szemben" történik (Arendt 1998).

Ezeket a helyeket azonban nem adják szabadon. Az ellenállás formái létrehozásához szükséges. Az ellenállás egyik lehetséges formája a munkavállalók önszerveződésében rejlik, mind a mesterséges intelligenciával kapcsolatos mérnöki munka központjában, mind pedig a többi munkahelyen, amelyeket ez a technológia érinteni fog. Az előbbinek vannak apró jelei abban, ahogyan az olyan vállalatok, mint a Google, a Microsoft és az Amazon alkalmazottai kifejezték ellenérzésüket azzal kapcsolatban, hogy technikai alkotásaikat a katonai és biztonsági apparátus átveszi (Alba 2018, Conger 2018, Lee 2019), míg az utóbbi eddig csak azokra a bizonytalan helyzetű munkavállalókra korlátozódott, mint például az Uber és a Deliveroo sofőrjei, akik már "az algoritmus alatt" vannak (Möhlmann/Henfridsson 2017). A munkavállalók önszerveződésében is kulcsszerepet játszanak a gyűlés és a tanács kollektív formái, különösen a munkavállalók azon törekvésének előmozdításában, hogy tudják: "az ipari szerveződéssel az új társadalom struktúráját alakítjuk ki a régi burokan belül". Az 1970-es évek közepén az egyik nagy fegyvergyártó vállalat dolgozói alulról szerveződő gyűléseket használtak arra, hogy tervet dolgozzanak ki gyáraik szerkezetátalakítására. Programjuk, a "Lucas-terv" nemcsak a gépek tevékenységét alakította volna át a fegyvergyártásról, hanem olyan újonnan kitalált terméklehetségeket is tartalmazott volna, amelyek a környezetre gyakorolt hatás tekintetében visszatekintve korukat megelőzendőnek tűnnek (Open University 1978). Az ellenállás egy másik lehetséges formája, amely

szükségszerűen kialakulhat, a luddizmus, ahol az emberek közvetlen cselekvéssel szállnak szembe a hegemón technológia rablásaival. A történelmi ludditák ellenállása a gőzzel hajtott gépekkel szemben a

a termelés az általuk létrehozott új társadalmi alávétettségi viszonyokon alapult. A ludditák ellenállását nem a technológiával szembeni atavisztikus ellenszenv, hanem alternatív társadalmi elképzelésük motiválta (Binfield 2004). Felhívásuk az volt, hogy "tegyenek le minden olyan gépezetet, amely árt a köznépnak". Az új luddizmus egyik módja annak, hogy jellemezzük az arizonai lakosok önvezető furgonok elleni támadásait, akiknek elégük van abból, hogy a Waymo autonóm mesterséges intelligenciát tesztel a közösségükben és az utcákon, ahol a gyerekeik játszanak. A mélytanulás ismét bebizonyította azt, amit az 1970-es évek rádiók állítottak, hogy a termelés területe kiterjedt a mindennapi életre, és hogy a "társadalmi gyárban" élünk (Cuninghame 2015).

A történelmi luddizmus a munkások és közösségek szélesebb körű lázadásának része volt, amely mélyen megrázta a korszak feltörekvő ipari elitjét. A mesterséges intelligencia jelenlegi formájában egy új technokrata elit eszköze. Bármilyen stratégiák is lesznek a mesterséges intelligencia átalakítására, azok nyilvánvalóan nem jöhetnek létre a progresszív politika szélesebb körű bevonása nélkül. A neurális hálózatok formájában a mesterséges intelligencia eredendően politikai technológia, amelyet ekként kell elismerni. Ha korlátozások nélkül fogadjuk el, akkor hajlamos lesz felerősíteni a status quo igazságtalanságait, vagy akár a szélsőjobboldal ellenséges környezetében egy sötétebb normativitás felé való elmozdulás részévé válhat. Lehetőség van azonban egy olyan mesterséges intelligenciára, amely tudatosan igazodik a társadalmi igazságosság és egyenlőség eszméihez. Nem autonóm döntéshozóként, hanem a társadalmi autonómiáért folytatott mozgalom részeként. Ez az AI egy szélesebb körű strukturális megújulás részeként, amely támogatja a hatalom kivonását a hegemon intézményekből és a kölcsönös segítségnyújtáson alapuló társadalmi szerveződések alternatív strukturáinak létrehozását (Landauer/Richard 2010). Saját cselekvőképességünk visszaszerzése nem azt jelenti, hogy a mesterséges intelligenciát mint olyat támadjuk, hanem azt a rendszert, amely a saját képére termeli a mesterséges intelligenciát. Ez azt jelenti, hogy a lehetséges oldalára állunk a valószínűvel szemben. Ha visszaszerezzük a kollektív gondolkodásra való képességünket, és inkább tanulunk egymástól és egymással, minthogy a gépi tanulásra hagyatkoznánk, akkor a szolidaritás és a kollektív gondoskodás gyakorlatával léphetünk fel a gondolatlanság ellen.

Hivatkozások

- 3Blue1Brown. n.d. "Mit csinál valójában a backpropagation?" Deep Learning, fejezet (3.<https://www.youtube.com/watch?v=Ilg3gGewQ5U>). Hozzáférés: 25 Szeptember 2018.
- "A lázadás hónapja a szolgáltatási szektorban (#4.1)". n.d. Notes From Below. ([https:// notesfrombelow.org/issue/revolt-in-the-service-sector.](https://notesfrombelow.org/issue/revolt-in-the-service-sector.)) Hozzáférés március 12. 2019.
- ACM FAT (2019): ACM Konferencia a méltányosságról, elszámoltathatóságról és

átláthatóságról (ACM FAT)". (<https://fatconference.org/>). Hozzáférés: 2019. július 11.

Agamben, Giorgio (2005): *Aggigeni Aggiben: A kivétel állapota*. Fordította: Kevin Attell. kiadás.1

Chicago: University Of Chicago Press.

- Alba, Davey (2018): "Here's How Amazon Defended Its Facial Recognition Tech To Concerned Employees At An Internal Meeting". BuzzFeed News. 8 November (<https://www.buzzfeednews.com/article/daveyalba/amazon-all-hands-facial-rekognition-ice>). Hozzáférés július 11 2019.
- Alcorn, Michael A., Qi Li, Zhitao Gong, Chengfei Wang, Long Mai, Wei-Shinn Ku és Anh Nguyen (2018): "Strike (with) a Pose: Neural Networks Are Easily Fooled by Strange Poses of Familiar Objects". ArXiv:1811.11553 [Cs], november. (<http://arxiv.org/abs/1811.11553>).
- Amodei, Dario és Danny Hernandez (2018): "AI and Compute". OpenAI Blog. Május 16. (<https://blog.openai.com/ai-and-compute/>). Hozzáférés: 2019. július 11.
- Alston, Philip. n.d. "OHCHR | Nyilatkozat Philip Alston professzor, az ENSZ szélsőséges szegénységgel és emberi jogokkal foglalkozó különmegbízottjának látogatásáról az Egyesült Királyságban". (<https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=23881&LangID=E>.) Hozzáférés: 2019. január 9.
- Arendt, Hannah (1998): Arendt, Hannah: Az emberi állapot. Chicago & London: University of Chicago Press.
- Arendt, Hannah (2006): Eichmann Jeruzsálemben: Jelentés a gonosz banalitásáról. 1 kiadás. New York, N.Y.: Penguin Classics.
- Bourdieu, Pierre (1990): Pierre Bourieu: A gyakorlat logikája. Stanford University Press.
- Conger, Kate (2018): "Google Plans Not to Renew Its Contract for Project Maven, a Controversial Pentagon Drone AI Imaging Program". Gizmodo. 2018. június 1. (<https://gizmodo.com/google-plans-not-to-renew-its-contract-for-project-mave-1826488620>). Hozzáférés: 2019. július 11.
- Cuninghame, Patrick (2015): "A harc terepének feltérképezése: Autonomous Movements in 1970s Italy". Viewpoint Magazine. 2015. november 1. (<https://www.viewpointmag.com/2015/11/01/feminism-autonomism-1970s-italy/>). Elfogadva 2019. július 11-én.
- Eckersley, Peter (2018): "(avagy miért ne legyen hasznossági függvénye az AGI-nak)". ArXiv: 1901.00064 [Cs], december. <http://arxiv.org/abs/1901.00064>.
- Fricker, Miranda (2007): Fricker: Episztemikus igazságtalanság: A hatalom és a megismerés etikája. Oxford, New York: Oxford University Press.
- Ghosh, Pallab (2018): "AI Could Save Heart and Cancer Patients", 2018. január 2., sec. Egészségügy. (<https://www.bbc.com/news/health-42357257>). Hozzáférés július 11. 2019.
- Goswami, Gaurav, Nalini Ratha, Akshay Agarwal, Richa Singh és Mayank Vatsa (2018): "Unravelling Robustness of Deep Learning Based Face Recognition Against Adversarial Attacks". ArXiv:1803.00401 [Cs], február. <http://arxiv.org/abs/1803.00401>.
- Hern, Alex (2018): "Google 'betrays Patient Trust' with DeepMind Health

Move". The Guardian, november 14sec2018,. Technológia.
(<https://www.theguardian>).

- com/technology/2018/nov/14/google-betrays-patient-trust-deepmind-health-care-move). Hozzáférés: 2019. július 11.
- Illich, Ivan (1975): *Illich: Tools for Conviviality*. Glasgow: Fontana.
- Konkel, Frank (2016): "A CIA titkos felhője hónapokról percekre csökkenti a feladatokat". *Defense One*. 2016. december 15. (<https://www.defenseone.com/technology/2016/12/cias-classified-cloud-reducing-tasks-months-minutes/133925/>). Hozzáférés: 2019. július 11.
- LaMattina, John (2016): "Is Prediabetes An Epidemic Or A Creation Of Drug Companies?" (A prediabetesz egy járvány, vagy a gyógyszergyártók kreációja?) *Forbes*, 2016. augusztus 4. (<https://www.forbes.com/sites/johnlamattina/2016/08/04/is-prediabetes-an-epidemic-an-epidemic-or-a-creation-of-drug-companies/#4302d76779c2>). Hozzáférés: 2019. július 11.
- Landauer, Gustav és Richard Day (2010): *Revolution and Other Writings: A Political Reader*. Szerkesztette Gabriel Kuhn. Oakland, CA: PM Press.
- Lee, Dave (2019): "Microsoft Staff: Do Not Use HoloLens for War", 2019. február 22., sec. *Technológia*. (<https://www.bbc.com/news/technology-47339774>). Elfogadva 2019. július 11-én.
- Lipton, Zachary C. (2016): "A modell értelmezhetőségének mítosza. ArXiv:1606.03490 [Cs, Stat], június. <http://arxiv.org/abs/1606.03490>.
- McQuillan, Dan (2017): *Mcillan: "Data Science as Machinic Neoplatonism"*. *Philosophy & Technology*, augusztus, 1-20. <https://doi.org/10.1007/s13347-017-0273-3>.
- McQuillan, Dan (2018): "Népi tanácsok az etikus gépi tanulásért". *Social Media + Society* 4 (2): 2056305118768303. <https://doi.org/10.1177/2056305118768303>.
- Möhlmann, Mareike és Ola Henfridsson (2017): "Uber Drivers Are Gaming the System and Even Going Offline En Mass to Force 'Surge' Pricing". 2017. augusztus 2. (https://warwick.ac.uk/newsandevents/pressreleases/uber_drivers_are/). Hozzáférés: 2019. július 11.
- Ng, Andrew (2016): Ng: "What Artificial Intelligence Can and Can't Do Right Now". *Harvard Business Review*, november 9. (<https://hbr.org/2016/11/what-artificial-intelligence-can-and-cant-do-right-now>). Hozzáférés: 2019. július 11.
- Open University (1978): (<https://www.youtube.com/watch?v=0pgQqqfpub-c>). Hozzáférés: 2019. július 11.
- Roy, Deboleena (2004): "Feminista elmélet a tudományban: Working Toward a Practical Transformation". *Hypatia* 19 (1): 255-79.
- Sian, Katy (2017): "Született radikálisok? Prevent, Positivism, and 'Race-Thinking'". *Palgrave Communications* 3 (1): 6. <https://doi.org/10.1057/s41599-017-0009-0>. <https://doi.org/10.1057/s41599-017-0009-0>.
- Slobodian, Quinn (2019): *The Rise of the Right-Wing Globalists*: "A jobboldali globalisták felemelkedése". *The New Statesman*, január 31. (<https://www.newstatesman.com/politics/economy/>). 2019/01/felkelés-jobboldali-globalisták). Hozzáférés július 11 2019.

Solhdju, Katrin (2015): "Valószínűségek ellenében a lehetséges oldalára állni, avagy: Hogyan örököljük a múltat?". In. ICI Berlin. (https://www.academia.edu/19861751/Taking_Sides_with_the_Possible_against_Probabilities_or_How_to_Inherit_the_Past.).

Mesterséges intelligencia

Láthatatlan ügynökségek a technológiai kultúrák ráncában

Yvonne Förster

1. Bevezetés

A demokrácia az átláthatóságról, a láthatóságról és a nyilvánosság bevonásáról szól. A görög poliszban a politikai döntéseket az *agorán* vitatták meg, egy olyan nyilvános helyen, ahol minden polgár (akkoriban csak a 30 évnél idősebb szabad férfiak) meghallgathatta és részt vehetett. A mai képvisleti demokrácia kevésbé nyilvános, de a döntési folyamatok átláthatósága rendkívül fontos. Ha egy kormány nem tudja eléggé átláthatóvá tenni a döntéseit, akkor fennáll a veszélye, hogy elveszíti az emberek bizalmát. A politikai értelemben vett átláthatóság szabályokat, láthatóságot, vitakészséget és indoklást jelent. A mesterséges intelligencia alkalmazásainak megjelenésével nemcsak a politikai szférában, hanem a társadalmi és magánélet gyakorlatilag minden területén az átláthatatlanság és a nem tudatos megismerés új formáival szembesülünk, amelyek erősen befolyásolják az emberi döntéshozatalt, viselkedést, mozgást és kommunikációt. A központi probléma az, hogy az AI-alkalmazások úgy cselekszenek, hogy nem tudnak számot adni a mögöttes okokról, sőt a mögöttes oksági folyamatok is átláthatatlanok (black box) maradnak. Ha egy hitelminősítés elemzésére használt AI elutasítja a hitelt, ez a döntés tönkreteheti a magánéletet. Ha ezután nem adnak indoklást vagy nem magyarázzák meg a lehetőségeket, ez önmagában is megingathatja az emberek bizalmát a civil társadalomban. A nem tudatos megismerésen alapuló ügynöki tevékenység mindenütt jelenlévő jelenséggé válik, és ezért etikai és fenomenológiai reflexiót igényel. Ebben az esszében azt szeretném megérteni, hogy a mesterséges intelligenciát a láthatóság és az átláthatóság szempontjából hogyan éljük meg. Ennek érdekében a fenomeno- logikus megfontolásokat Martin Heideggernek a technológia természetéről szóló reflexióival fogom összekapcsolni.

Az egyik jellemző, hogy kiváltja a spekulációkat a mesterséges intelligencia a tét itt az a tény, hogy legalábbis a felhasználó számára szinte lehetetlen megérteni, hogy a mesterséges intelligencia hogyan jut el a kimeneteihez. Az AI-alkalmazásokat gyakran fekete dobozként jellemzik (vö. Sudmann 2018a). Még ha az öntanuló algoritmusok mögött álló matematika meglehetősen egyszerű is, a bemenettől a kimenetig vezető oksági folyamatok nem igazán átláthatóak (vö. Sudmann 2018b: 63). A homályosságot általában fenyegetésként

és potenciális veszélyként fogják fel. Ez elvezet e cikk központi kérdéséhez:
Kell-e a mesterséges intelligenciának

láthatatlansága miatt a demokráciát fenyegető veszélynek tekintik? Amint érvelni fogok, ez a felszínen igaz, de azt is be fogom mutatni, hogy a technológia mindig a láthatatlanság egy bizonyos formájával jár együtt. A kérdés az, hogy ez a mesterséges intelligencia esetében új szintet ér-e el. Első lépésként definiálok, hogy mit értek láthatóság/láthatatlanság alatt nomenkológiai szempontból. Azért vezetem be ezt a nézőpontot, mert az érzékelést, a tapasztalást a technológiához kapcsolja. Ezután tisztázom, hogy ez hogyan vonatkozik az emberi és a mesterséges intelligencia viszonyára. A dolgozat utolsó részében a technológia eltűnésének kérdéseit, valamint az átláthatóság és az átláthatatlanság komplex viszonyát tárgyalom a technológiával kapcsolatban. Célom annak bemutatása, hogy a mesterséges intelligencia-rendszerek hogyan vezetnek be egy újfajta láthatatlanságot vagy átláthatatlanságot az életvilág ökológiai struktúráiba.

A láthatóság és láthatatlanság kölcsönhatásában legalább három különböző réteg van: Az egyik az érzékelés minden egyes tárgyára vonatkozik: Az észlelés perspektivikus, és így a láthatatlanság szükségszerű része. A láthatatlanság tehát az észlelés és a megismerés minden formájának alkotó része. A technológia esetében Heideggert követem abban a diagnózisban, hogy a láthatatlanságnak van egy magasabb rendű formája. Ez a technológia lényege, amely maga nem technológiai, hanem a gondolkodás vagy a feltárulkozás alapvető stílusa. A technológiának ez az elemzése párhuzamot mutat a tudat elemzésével, amely alkotmányában szintén átláthatatlan a tudatos szubjektum számára. Ennyiben nincs semmi úttörő vagy új a technológia szempontjából. A mesterséges intelligenciával az átláthatatlanság egy harmadik rétege lép a színpadra: Ez a nem-tudatos cselekvés - egy olyan cselekvés, amely nem tud okot adni, de nagyon mélyen formálja az életet. Bár a nem-tudatos cselekvés az emberekben és az állatokban is jelen van, a technológiai nem-tudatos cselekvés újdonság, mert alapvetően alakítja a társadalmi és politikai életet most és a jövőben. A láthatatlanság és az átláthatatlanság ezen aspektusainak kombinációja okozza a mesterséges intelligenciával kapcsolatos széles körű nyugtalanságot. Célom, hogy képet adjak arról, hogy a láthatóság, az átláthatóság és az átláthatatlanság különböző formái hogyan befolyásolják a mesterséges intelligencia demokráciát veszélyeztető vagy lehetővé tevő potenciálját.

2. Megjelenési feltételek: Láthatóság és láthatatlanság

Martin Heidegger *A kérdés a technológiáról* című esszéjében a technológiát a feltárás, az elrejtettnek az elrejtésbe való be- és elrejtés nélkülivé tétele módjaként írja le (vö. 11f.). Ez a felfogás mélyebb, mint a technológiának mint egy célhoz vezető eszköznek a szokásos instrumentális felfogása. A technológiai műtárgyak jellegét Heidegger szerint nem értjük meg megfelelően, ha ezt olyan eszközként fogjuk fel, amely egyszerűen csak segíti az embert bizonyos célok elérésében. Továbbá Heidegger azt is állítja, hogy a technológiát emberi cselekvésként szemlélve nem ragadja meg teljes mértékben. A technológia eszközként vagy

antropológiaként való felfogása nem téves mindkét felfogás. A technológiát aszerint ragadják meg, ahogyan azt általában megtapasztalják és használják. Mindazonáltal nem érik el annak lényegét, hogy mi is a technológia. De mi a lényege valaminek?

A dolog dologisága, az, ami által a dolog dolog? Valami, ami nem változik, míg más részei vagy aspektusai igen? Valójában nehéz fogalmilag meghatározni, hogy mit is jelent valójában valaminek a lényege.

Heidegger írásaiban legalább két fogalomról van szó: Először is, az antik görög fogalom, hogy *mi* valami (Heidegger 1977: 4); és másodsor, a "jelenlétként való megmaradás" (Heidegger 1961: 59). Esszéjében mindkét szempont releváns. Az a törekvés, hogy megértsük, *mi* a technológia, meghatározza az egész szöveget. Heidegger meg van győződve arról, hogy a válasz erre a kérdésre nem egy olyan entitás felé mutat, amely technológiai jellegű. A technológia lényege maga nem technológiai (vö. Heidegger 1977: 4). Ez azt jelenti, hogy a technológia lényege nem dolog; maga nem fizikai entitás. Továbbá úgy véli, hogy a technológia lényege egy tevékenység: valaminek a feltárása vagy *feltárhatatlanná tétele*. Állítása szerint a technológia csak mint alapvető folyamat vagy tevékenység marad fenn.

A mesterséges intelligencia körüli jelenlegi vitát is hasonló feszültség jellemzi. Az egyik oldalon az intelligens technológiákat egyszerű célokhoz vezető eszközként képzelik el. Az automatizálás, a robotika vagy a beszédfelismerés folyamatai, hogy csak néhányat említsünk, AI-alapúak. Ezek az összetett feladatok tanulási képességet igényelnek. Az öntanuló programok kívülről háborzongatónak tűnnek, de belülről talán nem annyira. Az ilyen mesterséges intelligenciák alkotói általában azt állítják, hogy nem sok intelligencia rejtőzik a programokban. Inkább egy olyan technológiai ügynökségről van szó, amely a szofisztikáltság elég magas szintjét éri el, de messze nem kreatív a képzettségi határain túl. Ez a feladatalapú funkcionális intelligencia bizonyára folyamatosan fejlődik, de egyelőre csak bizonyos határok között és az alapján, hogy az AI-t milyen inputokkal képezték ki (vö. Pontin 2018).

A közbeszédet viszont az emberi hatalmakat meghaladó szingularitás fenyegető forgatókönyvei táplálják, vagy - kevésbé futurisztikusan fogalmazva - a munkahelyek eltörlésével és az emberi viselkedés manipulálásával (pl. célzott, személyre szabott marketing) fenyegető mesterséges intelligencia. Ezek a problémák abból adódnak, hogy az AI általában átláthatatlan (ibid.), még ha lehetséges is olyan alkalmazások kifejlesztése, amelyekkel megfigyelhetők az AI tanulási folyamatai (Sudmann 2018a). Emellett az AI-alkalmazások tervezett mindenütt jelenléte széles körű vitákat vált ki a munkakultúrákra (AI a munkafolyamatok és az automatizálás optimalizálására) és a társadalmi környezetre (érzékelő alapú megfigyelőrendszerek) gyakorolt következményekről.

Ezek az aggodalmak kapcsolódnak Heideggernek a technológia lényegéről szóló vitájához. A mesterséges intelligenciával kapcsolatos aggodalmaink középpontjában talán az áll, hogy ezek a technológiák *világteremtő* technológiák (vagy legalábbis az elképzelések szerint). A technológia Heidegger szerint nem fizikai eszközök összessége, hanem mindenekelőtt a gondolkodás és a megnyilatkozó entitások stílusa. Ez az aspektus a jövő technológiáinak birodalmán

belül egyre világosabbá válik.

Ha közelebbről megvizsgáljuk Heidegger szavait a technológia lényegének leírására, a láthatósággal és láthatatlansággal való kapcsolat tagadhatatlan. A technológia lényegének olyan dologként való leírása, amely maga nem technológiai, gesztusokat tesz afelé, hogy

láthatatlanság. A technológia transzcendentális feltételei maguk nem technológiai vagy objektív jellegűek. Heidegger arra a gondolatra jut, hogy a technológia lényegében a *világteremtés* módja. Az *enframing* (*Gestell*) logikája a világot *állandó tartalékként* (*Bestand*) fogja fel, azaz a *mindenkor rendelkezésre álló erőforrások konstellációjaként*. Ezt a logikát már jóval a modern technológia megjelenése előtt is működőképesnek találja. Míg a történelem hajlamos a modern fizikát a modern technológia lehetővé tevőjének tekinteni, Heidegger úgy véli, hogy a technológia struktúrája vagy logikája már a modern fizika fejlődését is irányítja (ibid.: 22 f.). Ezt az állítását azzal indokolja, hogy a modern fizika mint olyan azon a meggyőződésen alapul, hogy a világnak megfigyelhetőnek, mérhetőnek és kiszámíthatóvá kell válnia (ibid.: 172). A kiszámíthatóságra azért van szükség, hogy a környezetet *állandó rezervátumként* kezeljük. A szükségletek és az erőforrások kölcsönhatása a jövőre és így az előrejelzésre épülő törekvés. Már a modern fizikát is az a cél vezérelte, hogy a fizikai világot előrejelzés és kiszámíthatóság révén megszelídítse, ami a mesterséges intelligencia használatában mutatkozik meg a legnyilvánvalóbban (pl. a határellenőrzésben vagy a városi CCTV alkalmazásokban használt arcfelismerés, a fogyasztói viselkedés előrejelzése vagy a munkafolyamatok optimalizálása a menedzsment AI segítségével). Ez azt jelenti, hogy a mesterséges intelligencia jelenlegi alkalmazásai a természeti erőforrások kiaknázásán túlmutatóan, és ezáltal az emberi viselkedés mélyére is kiterjesztik az *állandó tartalékok* feltárásának lehetőségét. Ennek a folyamatnak a mértéke még nem világos, még kevésbé a következményei és etikai kihívásai.

Amikor a technológia alkotja a világ értelmezhetőségét, amely a világot mint *Állandó tartalék*, mint ami mindig rendelkezésünkre áll, ugyanakkor rejt is valamit, vagy elrejt valamit. A technika (vagy inkább annak lényege, a *keretezés* folyamata) alattomos módon arra kényszeríti az embereket, hogy a világot általában a technika szempontjából értelmezhetőnek gondolják el, az instrumentális gondolkodás, az eszközök és a célok logikája által beszennyezve. Így elrejt a tárgyak jellegét, mint ami a szubjektumokkal szemben áll: "Ami az álló-tartalék értelmében áll, az már nem áll felettünk és velünk szemben mint tárgy". (Heidegger 1977: 17) A feltárulás vagy láthatóvá tétel Heidegger által leírt folyamata perspektivikus, és a perspektíva szükségszerűen elrejt az észlelt tárgyak más vagy háttérben lévő aspektusait is.

A láthatóság és a láthatatlanság több szempontból is feltételezi egymást: A technológia esetében a láthatóság és a láthatatlanság egymásra utaltsága vagy - Merleau-Pontyval (1969) szólva - a láthatóság és a láthatatlanság *chiasmája* (összefonódása vagy összefonódása) mélyebb, mint az érzékelés esetében. Az észlelés mindig szituált, tehát perspektivikus. Nincs észlelés perspektíva nélkül. És ez azt jelenti, hogy nincs láthatóság a láthatatlan nélkül. A láthatóság és a láthatatlanság dialektikája általában véve az észlelést alkotja.

A technológia anyagi tárgyként/eszközként való felfogásán túl, ami önmagában is fontos téma (vö. Verbeek 2005), Heidegger egy olyan kauzalitást

lát, amely nem merül ki a technológia instrumentális meghatározásában. A technológián keresztül a világot *mint* álló-rezervátumot látjuk. A technológia tehát látteleletet *termel*

(az élet-világ mint *állandó tartalék*), ahelyett, hogy csak (látható) tárgyakat adnánk hozzá a világhoz. Mint fentebb említettük, ezek a láthatóságok, vagy inkább az a mindent átfogó stílus, amelyben a technológia a világot arra kényszeríti, hogy általában véve technológiai jellegűnek tűnjön, valamit el is rejt, azaz láthatatlanná tesz: nevezetesen a dolgok objektív jellegét mint *Gegenstände*. Ez a láthatóság tág témakörén belül egy újabb szempontot nyit meg. Mi egy dolog, ha a dologisága vagy *Gegenständlichkeit* el van rejtve?

Ez történik, ha egy olyan szerszámot használunk, mint a kalapács: A gyakorlott felhasználó nem veszi észre a kalapácsot mint tárgyat. A kalapács inkább a test meghosszabbításává válik a használat során. Amíg a használat súrlódásmentes marad, a kalapács mint tárgy nem vonja magára a figyelmet. Tematikátlan marad, és tárgyi jellege átlátszó. A dolgoknak mint szerszámoknak ilyen használata az, amit Heidegger művein keresztül "*kézre készségnek*" vagy rendelkezésre állásnak (*Zuhandenheit*) nevez: a dolgokkal szembeni bizonyos magatartás leírása, mint ami készen áll a használatra, rendelkezésünkre áll. A technológia instrumentális attitűdje a dolgokat eszközként jeleníti meg, és elrejtja a tárgyi létüket.

3. Átláthatóság és átlátszatlanság a technológiai tárgyakban

Ha a technológiának ezt a heideggeri szemléletét lefordítjuk egy általánosabb terminológiára, akkor a láthatóság egy másik formájához jutunk: az *átláthatósághoz*. Egy eszköz vagy technológiai eszköz lehet átlátható abban az értelemben, hogy a felhasználói élmény zökkenőmentes. Az ilyen zökkenőmentes felhasználói élmény (vagy az úgynevezett "*súrlódásmentes UX*") a technológiai tervezés arany szabálya lett, és a mesterséges intelligencia az egyik eszköz ennek a célnak az eléréséhez. Egy öntanuló szoftver ideális esetben képes megtanulni a felhasználótól, hogy mit jelent a zökkenőmentes működés. Az alkalmazások használatán belüli zavarok tovább szolgálhatnak anyagként, amelyekből tanulhat, majd zökkenőmentesen áramló, gördülékenyebb működési folyamatokat hozhat létre.

Fenomenológiai szempontból a zavarok megtörik a zökkenőmentes működés mindennapi attitűdjét, és a tárgyak dologiságát és jellegét úgy tárják fel, mint ami velünk szemben áll (*Gegen-stände*). Csak ekkor lesz a felhasználóknak okuk arra, hogy ténylegesen elgondolkodjanak a technológiáról. Ez egyben a következő lehetőséget is megnyitja: A kritikai attitűd kialakulásához a megzavarás vagy a súrlódás szükséges összetevő. Az idegtudományban és az elmefilozófiában a megszakítás vagy a *predikciós hibák* beépülnek az idegi tevékenység mint *predikciós feldolgozás* modelljébe: "A prediktív kódolási sémákban az érzékszervi adatokat az előrejelzési hiba helyettesíti, mert ez az egyetlen olyan érzékszervi információ, amelyet még meg kell magyarázni. (Feldmann & Friston 2. o2010.,)". (Idézi Clark 2015: 4)

Ezzel szemben az is kiderül, hogy a funkcionális átláthatóság egyúttal az *átláthatatlansággal* is együtt jár. A működő technológia konstitutív folyamatainak és így a zökkenőmentes felhasználói élménynek rejtve kell maradnia ahhoz, hogy teljesíteni tudjon.

ezt a munkát. Ebben az értelemben a technológiai folyamatoknak *átláthatatlannak* kell lenniük: Rejtve maradnak a használat során, ha zökkenőmentesen működnek. Az átláthatóság és az átláthatatlanság úgy nyilvánul meg, mint a láthatóság és a láthatatlanság. A különbség a két fogalompár között az, hogy a láthatóság/inláthatóság esete semlegesebb módon írja le a tárgyak adottságát az érzékelésben.¹ Az átlátszóság és az átláthatatlanság inkább olyan jelentéssel bír, amely *normatív* aspektust is tartalmaz. Legalábbis ez a helyzet, ha a perspektívát a demokratizálódás vagy a mesterséges intelligencia demokráciát elősegítő potenciáljának kérdései felé tágitjuk.

Hogy ezt a gondolatmenetet részletesebben elmagyarázzam, hadd húzzak egy vonalat az átláthatóság fogalmának fenomenológiai felhasználásától a technológiai alkalmazásáig. Jean-Paul Sartre a transzparencia fogalmát a tudat, vagy szűkebb értelemben az imaginárius, azaz a képekhez és a fantáziához kapcsolódó tudatmódok leírására használja. A tudat különböző módokon konstituálja az észleléseket anélkül, hogy magát a konstituáló folyamatot érzékelhetővé tenné. Funkcionalitásában transzparens marad, ami azt jelenti, hogy nem válik az érzékelt, emlékezett vagy anticipáltként bemutatott tárgy részévé. Ennek analógiájára egy mesterséges intelligencia alkalmazás maga sem válik a tudatosság tárgyává, amikor zökkenőmentesen működik.

Ez világozássá válik, amikor az adottság különböző formáit hasonlítjuk össze. Például most éppen egy csésze tea áll az asztalomon. Az, hogy látom a csésze teát, a tudatosság aktusa, a vizuális észlelés aktusa. Ez az egyik módja annak, ahogyan a tudatosság egy dolgot megjeleníthet: a látás számára adottként, fizikailag ott, a kezem ügyében. De maga az észlelés aktusa nem tematikus, nem része a csésze szándékos tudatosságának. A tudat működése átlátható marad, és ezt kell is tennie, mert különben valami baj lehet a látásunkkal vagy az általános egészségi állapotunkkal. Ha később emlékszem a csésze teára, akkor a csésze vizuális jellemzőit az emlékezet révén reprodukálom. Az emlékezetem korrelátuma a képzeletem által előállított, amely a csészét úgy adja a tudatomnak, mintha láttam volna. A szándékos aktus ismét az emlékezés aktusaként érzékelhető, de az, hogy ez az emlékezés hogyan konstituálódik, nem tematizálódik magában az emlékezésben. A tudat működése átlátható marad. Nem tematikusak.

1 Edmund Husserl a *Thing and Space* [1907] című előadásában az észlelést szükségszerűen inadekvátnak írja le abban az értelemben, hogy szükségszerűen magában foglalja a közvetlenül nem érzékelt aspektusokat. Egy dolog térben való észlelése mindig részleges, lépésről lépésre gazdagodik a perspektíva változásával és a dolog rejtett oldalainak egyidejű kvázi-észlelésével (adumbrációjával): "Látjuk, hogy a testi dolog folytonossága "inadekvát" észlelést feltételez, olyan adumbrációkon keresztül történő észlelést, amely mindig képes a gazdagodásra és a pontosabb meghatározásra". (Husserl 1997: [101|21]) Az észlelésnek ez a fogalma szükségszerűen magában foglalja a nem észlelték észlelését is. Ez azt jelenti, hogy az emberi észlelés nem csupán a perspektívák egymás utáni hozzáadásával fogja fel a dolgokat. Inkább úgy ismerjük meg a térbeli és időbeli dolgokat, hogy a rejtett oldalakat implicit módon érzékeljük. Ez az, amit Husserl és Merleau-Ponty "appercepciónak" nevez: A nem észlelték észlelése.

a megismerés folyamatában. Ebben az értelemben ezek a folyamatok a felfedező elme számára is átláthatatlanok. Nincs tudatos hozzáféréstünk az elme belső működéséhez. És ez általában nem jelent problémát.

A mesterséges intelligencia esetében azonban más a helyzet: ha nem tudjuk, hogyan jutott el egy algoritmus a megoldáshoz, az nagyon problematikus lehet. Ha például orvosi adatokat elemezünk egy mesterséges intelligencia segítségével a rákos megbetegedésre való hajlam megállapítása érdekében, akkor tudnunk kell, hogy milyen alapon állították fel a diagnózist. Csak ezek alapján lehet döntést hozni a megelőző kezeléstről. A probléma az, hogy a mesterséges intelligencia anélkül képes előrejelzéseket generálni, hogy képes lenne megindokolni az eredményt, a minták kiválasztását vagy az alkalmazott módszert. Kategorikus különbség van a mentális állapothoz vagy egy program kimenetéhez vezető oksági folyamatok és a mentális állapotokra való indoklás és reflexió képessége között, ahogyan azt az elmefilozófia tárgyalja.

A tudatmódokat például a fenomenológiai elemzés módszereivel elemezhetjük, és egy adott szituációban a szándékosság különböző módoszatairól tarthatunk előadást. Ekkor a tudat mint folyamat elveszti átláthatóságát. Az emlékezés vagy a képzelet nem tematikus aktusa maga is egy magasabb rendű vagy- der reflexió tárgyává válik. De ekkor az átláthatóságnak egy magasabb rendű formája is megjelenik, nevezetesen a mentális állapotok konstitutív folyamataira való összpontosítás maga is az észlelés tárgyává válik, és így magának is konstituálódnia kell. A végtelen regresszus itt nagyban kirajzolódik. Edmund Husserl intencionalitás-elemzésének tanulsága az, hogy mindig van a tudatosságnak egy olyan rétege, amely maga nem lehet tudatos, mert maga is a tudatosság egy alacsonyabb szintjét vagy aspektusát alkotja. Az időbeli változás tudata például maga nem lehet időbeli, legalábbis nem úgy, ahogyan az idő tapasztalata az:

De komolyan el kellene gondolkodnunk azon, hogy nem kell-e feltételeznünk egy ilyen végső tudatosságot, amely szükségszerűen egy "tudattalan" tudatosság lenne; vagyis mint végső intencionalitás nem lehet a figyelem tárgya [...], és ezért soha nem válhat tudatossá ebben a sajátos értelemben. (Husserl 2008: 394)

A tudat tehát nemcsak az érzékelés médiumaként átlátszó, hanem bizonyos konstitutív aspektusaiban átláthatatlan is kell, hogy maradjon. Nem érthetjük meg a tudatot pusztán azáltal, hogy tudatosak vagyunk.

Az emberi tudatosságot mélyen befolyásolja a technológia, és napjainkban különösen a mesterséges intelligencia (vö. Hansen Hayles 2012, Stiegler 2012, 1998). Az emberi tudat *technogenezise*, ahogy Katherine Hayles fogalmaz, az átláthatóság/nyílékonyág egy másik dimenzióját nyitja meg. Az AI a nem-tudatos megismerés egy olyan formája (vö. Hayles 2014), amely egyre inkább mindenütt jelen van. Nincs online vásárlás mesterséges intelligencia által generált javaslatok nélkül; a közösségi média minden hírfolyamát algo-

rithmusok és még a repülőjegyárak is az időhöz, a helyhez és az eszközökhöz igazodnak. A virtuális világot többé-kevésbé kifinomult mesterséges intelligencia-alkalmazások segítségével nagymértékben személyre szabják.

Az eszközök működése nem csak abban az értelemben átláthatatlan, hogy a felhasználó nem érzékeli a tényleges számítási folyamatokat, de még kevésbé az e folyamatokkal együtt járó adatgyűjtést. Sőt, az AI-alkalmazások által generált kimenet sem feltétlenül fedi fel a mögöttes személyre szabási folyamatokat. Az internet csak nagyon korlátozott mértékben közös világ. A tartalmak nagy része a felhasználó és az AI közötti interakciók révén alakul ki, bár a felhasználó nem tudatosul ezekben az interakciókban. Ami a technológiát általában véve illeti, minden új találmánnyal együtt az emberi viselkedés és megismerés változásai is megfigyelhetők. Az írás feltalálása például mélyen megváltoztatta azt, hogy az emberek hogyan jegyzik meg a tartalmakat, és hogyan őrzik meg a kultúrák a hagyományait. Az okostelefonok elterjedése teljesen megváltoztatta az emberi kommunikáció módját. Egy egyszerű példa erre az üzenetküldő eszközökön és a közösségi hálózatokon keresztül történő kommunikáció: "a távkommunikáció [...] az arc elrejtésével, a hang eltűnésével és hangszínének eltűnésével, olyan kvázi-identitások felvállalásával jár, amelyek nem erednek hitelesen a világban való létezésünk konkrétumaiból". (El Bizri 2018: 130) Számtalan példát találhatnánk arra, hogy a feltörekvő technológiák révén új láthatóságok és egyúttal átláthatatlanságok jönnek létre.

A háttorzogató csavar a mesterséges intelligenciával jár. Két tényező fontos: Az időbeli mi- a számítási folyamatok és a prediktív kódolás kroszkálája. Az első tényező, nevezetesen a számítási folyamatok sebessége, amely az emberi megismerés számára hozzáférhetetlenné teszi azokat, olyan forgatókönyvet hoz létre, amelyben a második tényező, nevezetesen az, hogy a prediktív kódolás hogyan válik az emberi észlelést megelőző erővé. Ahogy Mark Hansen írja, amikor azt vizsgálja, hogyan működnek az életvilág szövetének központi elemévé váló számítási folyamatok a tudatosságunkon kívül eső időbeli mikroskálákon:

a számításoknak a környezetbe történő elosztása révén, a ma már tipikus technológiák, például az okostelefonok és az RFID-címkék segítségével, a tér saját ügynöki tevékenységgel válik animálttá. Ennek az animációnak az egyik döntő jellemzője, hogy az intelligens környezetekben a cselekvő személyek figyelmének középpontján kívül - vagy mellette - jelenik meg. Emiatt a kortárs élet intelligens tere egyfajta affordanciát kínál - egy nem érzékelt vagy közvetlenül érzékelt affordanciát -, amely alapvetően különbözik az affordanciáktól, ahogyan azokat James Gibson munkája nyomán a médiával kapcsolatban elméletileg értelmezték. Amikor "mi" ilyen intelligens környezetekben cselekszünk, a mi cselekvésünk olyan számítási ágensekkel párosul, amelyek cselekvése nem csak (legalábbis részben) kívül esik az ellenőrzésünkön, hanem nagyrészt a tudatosságunkon is. (Hansen 2012: 33)

Ez a leírás azon a feltételezésen alapul, hogy az emberi megismerés a környezetbe való beágyazottság révén vagy azzal összefüggésben jön létre. Hansen ezt "környezeti állapotunknak" (ibid.) nevezi, amely az egyén és környezete összekapcsolódását írja le. Ez a kapcsolódás nem statikus kapcsolat, hanem nagyon is dinamikus - egy állandó válási folyamat. Ez a folyamat on-tológiaiakra utal, amelyek vagy azt vallják, hogy a tudat a környezetbe való beágyazottságból keletkezik (vö. Merleau-Ponty 1969, Thompson 2013), vagy azt, hogy a tudat még a környezetbe is belenyúlik (a pánpszichizmus egy változata, vö. Chalmers 2013, Whitehead 1929). Anélkül, hogy belemerülnék a környezet/ökológia vitába, most ezeket a gondolatokat szeretném átültetni az intelligens környezetek és a mesterséges intelligencia által vezérelt ökológiák kontextusába.

Hadd foglaljam össze röviden a gondolatmenetet, amely ehhez a jelenlegi ponthoz vezetett. ^{Heidegger} azon felfogásával kezdtem, hogy a technológia lényege a láthatóság, vagy konkrétabban a világ érzékelhetővé tétele mint *álló-tartalék* lehetővé tevő feltétele. A technológiának ez az aspektusa önmagában nem technológiai; inkább a technológiai gondolkodás konstitutív struktúrája, és így a technológiai eszköz látható anyagiságának alapját képezi és teszi lehetővé. Innen tettem egy kitérőt abba, hogy hogyan konstituálódik az emberi észlelés, és megmutattam, hogy a vizuális észlelés mindig szituált és ezért perspektivikus. Ez azt jelenti, hogy a láthatatlanság aspektusai a látás vagy általában az észlelés konstitutív részét képezik. Érvélesem következő lépése a láthatóság és láthatatlanság viszonyát a technológiai műtárgyakra helyezte át, ahol a láthatóság és láthatatlanság helyett átláthatóságról és átlátszatlanságról beszélünk. A technológiai eszközök ugyanúgy átlátszóvá válnak a használat során, ahogyan az ember tudata is átlátszó az észlelésben (az észlelés konstitúciójának folyamata például maga nem tárgy az észlelésnek). A technológiai eszköz akkor működik zökkenőmentesen, ha nincs szükség zavarásra, és így nincs szükség a használat folyamatára való reflexióra. Ez az átlátszóság mindig átláthatatlansággal jár együtt. Bár a mechanizmusok funkcionalitást produkálnak, a számítási folyamatok rejtve maradnak, ezért a digitális technológiákat gyakran fekete dobozként jellemzik. Ez még élesebbé válik az öntanuló algoritmusok esetében, amelyeket még programozóik sem értenek teljesen.

Céлом, hogy a további fejezetekben bemutassam, hogy az átláthatóság/láthatóság és az átláthatatlanság és a láthatatlanság összefonódnak, és új lehetőségeket teremtenek. Ezen a ponton az intelligens környezetekről szóló elmélkedéssel folytatom. Az intelligens környezetek vagy a tárgyak internetévé (IoT) alakított házak olyan technológiát példáznak, amelyet az öntanuló mesterséges intelligencia irányít, és amelynek fő funkciója az előrejelzés. Az előrejelzésre azért van szükség, mert például a háztartásokon belüli IoT egymással összekapcsolt folyamatok rendkívül dinamikus együttese, amelynek mindenféle helyzethez és változáshoz alkalmazkodnia kell. Végül azzal fogok érvelni, hogy a mesterséges intelligencia által vezérelt intelligens környezetek két okból is jelentősen különböznek az alacsony technológiai színvonalú környezetektől: (1) a prediktív reagálóképesség korábban nem volt általános jellemzője a

környezeteknek, és (2) az előrejelzés

és így a preemptív funkcionalitás az ideális emberi/emberi viselkedésről alkotott elképzelés köré épül. Itt kell kezdeni a politikai vitát.

4. Intelligens környezetek: Technológiák az életvilág szövetében

Az intelligens technológiákat az életvilág szövetébe vagy a *fleetsibe* (Merleau-Ponty Rabari1969., Storper Förster2015, 2018a) szövik be, és fontos megérteni, hogy ez határozottan *nem* metafora: A városi terek számtalan érzékelőből, kamerából és monitorból állnak. Különösen az olyan megavárosokban, mint Szóul, Tokió, London vagy New York City, a város szó szerint minden szegletében vannak CCTV-k. A kijelzők mindenütt jelen vannak, ahová csak nézünk, és a levegőminőséget, a fényerősséget vagy a város zaját mérő érzékelők észrevétlenül maradnak, még akkor is, ha elkezdjük keresni őket. A csatlakoztatott eszközök növekvő sűrűsége az intelligens környezetekben egyre nagyobb igényt támaszt a nagyon kis méretű hardverek, integrált eszközök és nagy sebességű adathálózatok iránt.

Miközben a városi terek, a munkahelyek és a magánterek egyre inkább technológiai jellegűvé válnak, a hardver viszont egyre kevésbé látható. Az érzékelők látnak anélkül, hogy látnánk, és hallanak anélkül, hogy hallanánk. Ez a sajátos jelenség a jövő életvilágának elbeszéléseiben, különösen a kortárs sci-fi filmekben. Amit jelenleg IoT vagy Internet of Everything (IoE) címke alatt hirdetnek, vagy más néven jelennek meg, az a mesterséges intelligenciát és így a nem tudatos megismerést az életvilág utolsó zugaiba is kiterjeszti. A hardver eltűnésének toposza tovább növeli az AI-alkalmazások funkcionális átláthatatlanságát. A felhasználóknak alig van esélyük megérteni, hogyan épül be a mesterséges intelligencia az eszközökbe, amikor az ténylegesen működik, vagy hogyan alakítja a használat folyamatát vagy élményét. Ráadásul a fizikai megvalósítást már nem könnyű megtalálni. Ez azt jelenti, hogy a technológiai környezetek érzékszervi, re-szponzív felületté válnak, nem tudatos megismeréssel. Az érzékeny, mesterséges intelligencia által vezérelt környezetek kezelése ezért új tudás- és viselkedésformákat igényel, például a technológiai cselekvőképesség megértését. A nem-tudatos megismerés és a cselekvés meglehetősen új lehetőségeket jelent a mindennapi életben. Egyrészt az emberi viselkedésnek és mozgásnak alkalmazkodnia kell a technológiai rendszerekhez, hogy azok megfelelően működjenek. Másrészt az embereknek el kell gondolkodniuk azon, hogyan akarják, hogy ezek az új technológiák beépüljenek az életvilágukba. Pontosán ez az a pont, ahol az új lehetőségek és ezáltal az új kulturális struktúrák aktív alkalmazására van szükség.

Az intelligens környezetekbe integrált AI aktívan alakítja az érzékelést, a mozgást...

ments, érzelmek és racionális döntések (pl. választások, etikai döntések stb.). Az egyik központi probléma az, hogy a mesterséges intelligenciák elsősorban az affektusok szintjén fejtik ki hatásukat (vö. Parisi 2018). Ez az átláthatatlanság

vagy láthatatlanság egy harmadik szintjét is hozzáteszi: az AI-alkalmazások által generált lökések nem mindig érzékelhetők ekként. Recom-

A vásárlási alkalmazásokban szereplő ajánlások meglehetősen egyszerűek. A hírfolyam-generálás mögöttes struktúrái sokkal kevésbé nyilvánvalóak. A mód, ahogyan az internetről lekérdezzük az információkat, mindig a mögöttes tanulási algoritmusok előrejelzéseitől van befolyásolva. Így a hírfolyamban megjelenő világ egy megszemélyesített világ, amelyet egy mesterséges intelligencia generál, amely látszólag ismeri a felhasználót, miközben a felhasználó nem tudja, hogyan generálja a program a kimenetét. A nem-tudatos megismerés és cselekvés átláthatatlansága, ahogyan az a mesterséges intelligencia alkalmazásaiban működik, bizonytalanságot teremt a jelenlegi és a jövőbeli társadalmi életet illetően. A jelenlegi sci-fi filmek szimpatikusak egy idegesebb emberi állapot számára (vö. Förster 2016). Ott a technológiát általában rejtett erőként ábrázolják, amely kognitív evolúción megy keresztül, és végül felülkerekedik az emberiségen, vagy elavult életformaként hagyja maga mögött (pl. a *Her* (USA, 2013) és a *Transcendence* [USA, 2014]) című filmek.

Tagadhatatlan tény, hogy a technológia egyre láthatatlanabbá válik. vagy legalábbis kisebbek és jobban integrálódnak a mindennapi tárgyakra és a városi környezetbe. Még a bőr mint korlát is lassan lebomlik. A testbe integrált érzékelők egyre inkább megszokottá válnak, még ha ennek etikai dimenziója vitatott is. Svédországban például néhány embernek 3000már beültettek ilyen szenzorokat a bőre alá, hogy helyettesítsék a kulcsokat, hitelkártyákat vagy vonatjegyeket. Az elosztott mesterséges intelligencia-rendszereknek ma két markáns jellemzője van: a környezet részévé válnak (tendenciózusan beleolvadnak a mindennapi tárgyakra és felületekre, mint például a hűtőszekrények, a megfigyelő kamerák vagy az Alexa hangasszisztens), vagy pedig olyan eszközök részei, amelyek a test közelében működnek, vagy a testbe integrálódnak (intelligens szövetekkel ellátott ruhák, ékszeres vagy intelligens implantátumok). Úgy is mondhatnánk, hogy a technológia természetessé válik, ha egyáltalán volt valaha is egyértelmű különbség a mesterséges és a természetes között.

Az intelligens környezeteket nagyrészt a mesterséges intelligencia irányítja, mivel az eszközök elosztott hálózata által generált rengeteg adatot meg kell emészteni és hasznossá kell tenni. Ebben a szakaszban a láthatóságok és azok megfelelőinek összetett struktúrájával állunk szemben. A technológia mint hardver kezd eltűnni, miközben a funkció-potenciálja egyre exponenciálisabban növekszik. Ez a láthatatlanság felé mutató tendencia másod- (vagy akár harmad-) rendű átláthatóságot generál: A technológia nem csak a használatában átlátszó, hanem tárgyként is átlátszóvá válik. Ha a technológia már Heidegger megfontolásai szerint elvesztette vagy elrejtette tárgyi jellegét, akkor most az *enframing* vagy *Gestell* egy magasabb szintjére jutunk. Heidegger szerint a technológia akadályozza a világ tárgyként való szemléletét, mert arra kényszerít bennünket, hogy a világot mint *álló-tartalékot* fogjuk fel. Ez az alapjellemező a technológiára is érvényes: Nem úgy jelenik meg, mint egy szubjektummal szemben álló tárgy, és ebben a tekintetben meghaladja az emberi célokat. Eszköz egy cél érdekében, amely a

környezetet a lehetőségek kiszámítható, kiszámítható tárházaként jeleníti meg. A transzparenciának ez az új rétege a technológiai eszközök eltűnésével és az elosztott mesterséges intelligencia megjelenésével jön létre. Ez azért jelent különbséget, mert a legtöbb mesterséges intelligencia

ma elért eredmények előrejelző jellegűek, és a funkcionális folyamatokat ezen előrejelzéseknek megfelelően alakítja. A felhasználók számára az ismeretlen tényezők óriásiak: a felhasználók nem tudhatják, hogy pontosan hogyan működik a rendszer (pl. IoT) vagy akár egyetlen eszköz. Ráadásul nem tudhatják, illetve aktívan nem tapasztalhatják meg jobban, hogy milyen, mikor és mennyi adatot gyűjtenek az ezen eszközökkel kapcsolatos használatból és viselkedésből. Sokkal kevésbé tudhatják, hogy az összegyűjtött adatokat hogyan használják fel. A Big Data problémáról egyre szélesebb körben beszélnek, és az összetettség az eszközök óráról órára növekvő használatával egyre nő.

Fenomenológiai szempontból az intelligens technológiák csökkenő láthatósága (és átláthatatlansága), valamint a bennük rejlő fokozott ügynöki potenciál problematikus a mesterséges intelligencia demokratizálása szempontjából. És a probléma nem maga az AI, amelynek tényleges intelligenciája elképesztő, de folyamatosan túlértékelt is. Az AI-nak nem áll szándékában jobb önmagát, jobb társadalmat vagy jobb jövőt építeni. Az emberi lényeknek ez a célja. A filozófiától sem idegenek az ilyen, többnyire túlzó célok. A mesterséges intelligencia átlátható használatának egyik akadálya ez a tökéletességre való törekvés, amely többé-kevésbé gazdasági eszköz. Az intelligens technológiákban megvan a potenciál, hogy hasznosak legyenek, és talán még egy jobb jövőt is teremtsenek, de csak akkor, ha a kritika és a nyílt vita kultúráját sikerül kialakítani és fenntartani. Hogyan kapcsolódik ez a pont a láthatóság témájához? Hadd utaljak még egyszer utoljára Heideggerre. Ő azt állítja, hogy a technológia a világot *álló-rezervátumként* engedi megjelenni. Ma azt a kérdést kell feltennünk, hogy az emberi életek hogyan jelennek meg a technológián keresztül. Hogyan festenek az emberek az emberi életről képet azáltal, hogy olyan környezetet teremtenek, amely gépbarát? Megvannak-e az eszközeink ahhoz, hogy az intelligens környezetekben felvirágoztassuk annak rejtett arányát, hogy mit jelent embernek lenni? Hogyan teremthetünk elég szabadságot és lehetőséget a kreatív cselekvésre, amely lehetővé teszi a meglévő technológiákkal való aktív és kritikus foglalkozást? Ez azt jelentené, hogy kísérleteznünk, babrálunk kellene, és mindenekelőtt kritikus párbeszédet kellene folytatnunk az olyan iparágakkal, amelyek nagymértékben támaszkodnak az előrejelzésekre, mint például a kiskereskedelmi és biztosítási vállalkozások. A "jó" emberi életről alkotott képet (az antropocén fogalmát is figyelembe véve) alaposan meg kellene vizsgálni. Ezért alaposan meg kell figyelniük, hogyan befolyásolják a mindennapi viselkedésünket - hogyan beszélünk, mozgunk, és egyáltalán, hogyan mosolygunk vagy szeretünk - a "nudges", az előrejelzések és az elővigyázatosság. Ahhoz, hogy ezt sikeresen megteheszük, az embereknek a maguk teljes sokféleségében láthatóvá kell válniuk és hangként kell jelen lenniük a nyilvánosságban és a mesterséges intelligenciára nagymértékben támaszkodó iparágakban. A mesterséges intelligencia *politikai* dimenziója nagyon is emberi dimenzió. Az intelligens technológiákba épített *emberi* képet láthatóvá kell tenni. Csak ezután kerülhet sor megfelelő etikai vitára.

Hivatkozások

- Binfield, Kevin, szerk. (2004): *A ludditák írásai*. Baltimore: Johns Hopkins University Press.
- Chalmers, David J. (2013): "Pánpszichizmus és pánprotopsichizmus". In: *The Amherst Lecture in Philosophy* 8 (2013): 1-35 (www.amherstlecture.org/chalmers2013/).
- Clark, Andy (2015): "Embodied Prediction." In: Thomas Metzinger & Jennifer M. Windt (szerk.), *Open MIND*: 7(T). Frankfurt am Main: Frankfurt am Main: MIND Group. doi: 10.15502/9783958570115.
- El-Bizri, Nader (2018), "A hely és a tér fenomenológiája korunkban: Heideggeri utakon gondolkodva". In: Erik Champion (szerk.), *A valós és virtuális helyek fenomenológiája*, London: Routledge, 123-143.
- Förster, Yvonne (2016): "Szingularitások és szuperintelligencia: Az ember meghaladása a kortárs filmművészetben". In: *Trans-Humanities*, Szöul: Ewha Institute for the Humanities (EIH), 33-50.
- Förster, Yvonne (2018a): "A digitális bőrtől a digitális húsig: A technológia megértése a divaton keresztül." In: *Popular Inquiry* (2), Aalto University Helsinki (www.popularinquiry.com/blog/2018/8/30/yvonne-foerster-from-digital-skins-to-digital-flesh-understanding-technology-through-fashion).
- Förster, Yvonne (2018b): "Ha künstliche Intelligenz laufen lernt. Verkörperungsstrategien im *Machine Learning*." In: Christoph Engemann und Andreas Sudmann (Eds.), *Machine Learning. Medien, Infrastrukturen und Technologien der Künstlichen Intelligenz*, Bielefeld: transcript, 325-340.
- Hansen, Mark B. N. (2012): "Engineering Pre-Individual Potentiality. Technika, tranzindivíduoáció és a 20. századi média". In: *SubStance* 41/3 (Issue 129), 32-59.
- Hayles, Katherine (2012): *Hogyan gondolkodunk. Digitális média és kortárs technogénés*, Chicago, London: University of Chicago Press.
- Hayles, N. K. (2014): "Cognition Everywhere: The Rise of the Cognitive Nonconscious and the Costs of Consciousness". *Új irodalomtörténet* 45(2), 199-220.
- Heidegger, Martin (1961): New York: *An Introduction to Metaphysics*, New York: Heidegger: *An Introduction to Metaphysics*: Doubleday. Heidegger, Martin (1967): Heidegger: *Being and Time*. Oxford, Cambridge MA.: Blackwell. Heidegger, Martin (1977): Heidegger: "The Question Concerning Technology" (A technológiára vonatkozó kérdés). In: The Heidegger: Az emberiség és az emberiség: *Essays*, New York, London: Garland, 3-35.
- Husserl, Edmund (2008): Husserl: "A belső idő tudatának fenomenológiájáról (1893-1917)". In: Rudolf Bernet (szerk.), Edmund Husserl. *Collected Works IV*, Dordrecht: Springer.
- Merleau-Ponty, Maurice (1969): Evanston: *A látható és a láthatatlan*: Evanston: Northwestern University Studies.

- Parisi, Luciana (2018): Parisi Luciana: "Automata megismerés és tőke". In: Warren Neidich (szerk.). *A kognitív kapitalizmus pszichopatológiai: Harmadik rész*, Berlin: Archive Books.
- Pontin, Jason (2018): "Kapzsi, törékeny, átláthatatlan és sekélyes: The Downsides to Deep Learning". In: *Wired.com*: www.wired.com/story/greedy-brittle-opaque-and-shallow-the-downsides-to-deep-learning/ (2019.03.26.).
- Rabari, Chirag, Storper, Michael (2015): "A városok digitális bőre: városelmélet és - kutatás az érzékelt és mért város, a m i n d e n ü t t j e l e n l é v ő kom- puting és a big data korában". In: *Cambridge J Regions Econ Soc* 8 (1), 27-42. <https://doi.org/10.1093/cjres/rsu021>. <https://doi.org/10.1093/cjres/rsu021>.
- Stiegler, Bernard (1998): Stiegler: *Technika és idő, 1: The Fault of Epimetheus*, Stanford: Stanford University Press.
- Sudmann, Andreas (2018a): "A mesterséges intelligencia médiapolitikai dimenziójáról. A mélytanulás mint fekete doboz és a nyílt mesterséges intelligencia". In: Mathias Fuchs, Ramón Reichert (szerk.): *Digital Culture & Society (DCS)*, Vol. 4/1, *Rethinking AI: Neural Networks, Biometrics and the New Artificial Intelligence*, Bielefeld: transcript, 181-200.
- Sudmann, Andreas (2018b): "Szenarien des Postdigitalen. Deep Learning als MedienRevolution". In: Christoph Engemann, Andreas Sudmann (szerk.): *Machine Learning. Medien, Infrastrukturen und Technologien der künstlichen Intelligenz*, Bielefeld: transcript, 53-73.
- Thompson, Evan (2007): *Mind in Life*: Cambridge MA, London: Harvard University Press.
- Verbeek, Peter-Paul (2005): *What Things Do: Philosophical Reflections on Technology, Ügynökség és tervezés*, University Park: Pennsylvania State University Press.
- Whitehead, Alfred North (1929): *Gifford Lectures Delivered in the University of Edinburgh During the Session 1927-28*, New York: *Process and Reality an Essay in Cosmology*: Free Press.

Verseny és számítógépes látás

Alexander Monea

Bevezetés

A demokrácia és a mesterséges intelligencia kereszteződésének bármilyen elemzésének mindenekelőtt a mesterséges intelligencia és a marginalizáció már létező gyakorlatainak kereszteződését kell vizsgálnia. Az Egyesült Államokban talán nincs is kiemelkedőbb metszéspont, mint a mesterséges intelligencia és a faji hovatarozás metszéspontja. Mivel a mesterséges intelligenciát egyre inkább a gazdaság, a hadsereg, az állami bürokrácia, a térben és időben történő kommunikáció és közlekedés, egyszóval az emberiség jövőjének alapköveként pozicionálják, központi kérdéssé válik, hogy miként keresztezi egymást a mesterséges intelligencia és a faji marginalizáció már létező gyakorlata. Ezekre a kérdésekre különösen nehéz válaszolni, mivel a legtöbb kortárs mesterséges intelligencia-rendszer fekete dobozba zárt. Bár mindenképpen érdemes az AI-rendszereket működtető adathalmazok és algoritmusok nagyobb átláthatóságát szorgalmazni, ez az átláthatóság a jövőben várható, és most nem segíthet a jelenlegi AI-rendszerek működésének elemzésében. Ezt a képet csak bonyolítja az a tény, hogy az AI-rendszerek, különösen a webes léptékben működő rendszerek, működésük későbbi szakaszaiban még a mérnökeik számára is nehezen érthetőek. Egy programozó például könnyen leírhatja a kiindulási adatokat és a gépi tanulási algoritmust, amiből kiindult, de teljesen képtelen lehet elmagyarázni a rendszer által megtanult későbbi osztályozások logikáját. Ismétlem, természetesen érdemes az AI magyarázhatóságát követelni - vagyis megkövetelni az AI programozóitól és mérnökeiktől, hogy olyan rendszereket fejlesszenek ki, amelyek képesek megmagyarázni a döntéshozatali folyamatukat, vagy a legszélsőséges esetben *csak olyan döntéseket hoznak, amelyek egy ember számára világosan megmagyarázhatók* -, de ez megint csak egy előrevetített jövő, amely kevéssé segíti a faj és az AI közvetlen kérdésének megválaszolását, amelynek már most is súlyos következményei vannak.

Hogyan elemezheti tehát egy kívülálló kritikusán a faji és a faji kérdésekkel való találkozást?

és a mesterséges intelligencia a jelenkorban? Ez a tanulmány egy interdiszciplináris módszertant használ, amelyet "spekulatív kód tanulmányoknak" nevezek, amely a sajtóközlemények, a vállalati blogbejegyzések, a tudományos és technológiai szaksajtó és a technológiai szabálytalanságok jelentett eseteinek

archi- val kapcsolatos kutatását kombinálja a gépi tanulási algoritmusok által használt elérhető adathalmazok kritikai kódtanulmányokkal, a fekete dobozos kód nyílt forráskódú változatainak elemzésével és empirikus vizsgálatokkal.

a feketedobozos rendszerek kimeneteinek tanulmányozása. Egy ilyen tanulmány célja a fekete dobozos kód szigorú, de spekulatív elemzése. Az elemzésnek mindig spekulatívnak kell maradnia, mivel a tényleges rendszerek el vannak rejtve a közvetlen elemzés elől, de a módszertan biztosítja, hogy az elemzés a lehető legszigorúbb legyen a rendelkezésre álló perifériás anyagok, kódrészletek és bemenetek/kimenetek ismeretében. Ebben a fejezetben ezt a módszert arra használom, hogy megvizsgáljam, milyen számtalan módon kódolódnak bele a technológiai vállalatok igazgatótanácsaiban és kutatási és fejlesztési szárnyaiban, *valamint a* tágabb társadalmi-kulturális környezetben jelen lévő faji előítéletek az adatkészletekbe, majd a rájuk épülő gépi tanulási algoritmusokba. Az lesz az érvelésem, hogy ezen algoritmusok közül sok a faji előítéletek anyagi megnyilvánulása.

Ez az írás elsősorban a vizuális vagy optikai médiumokkal, és különösen a számítógépes látás algoritmusokkal foglalkozik. Azzal fogok érvelni, hogy ebben a kontextusban a feketeség a hipervizibilitás és a láthatatlanság hírhedt dialektikájába esik - a feketeséget túl gyakran sztereotípiákban, olykor vizuálisan karikatúraszzerűen ábrázolják, vagy pedig rendszerszinten láthatatlanként jelenítik meg. Fontos azonban megjegyezni, hogy bár elemzésem kontextusa a túlnyomórészt afroamerikaiakkal szembeni rendszerszintű rasszizmus, ezek a rendszerek globális hatással vannak a színesbőrű emberekre - pontosabban azokra a sötétebb bőrű egyénekre, akik a Fitzpatrick-skála V. vagy VI. típusába tartoznak (vö. Fitzpatrick 1988/1975.). Amikor olyan kifejezéseket használok, mint "fekete", "színesbőrű" vagy "sötét bőrű", ezzel azt akarom jelezni, hogy az általam azonosított problémák globális jelentőségűek, és nagymértékben érintik a Fitzpatrick-skála sötétebb végéhez tartozó embereket szerte a világon, még akkor is, ha elemzésem az Egyesült Államok történelmének és kultúrájának kontextusában történik. Nem tartozik e fejezet hatáskörébe, hogy ezt az elemzést más konjunktúrákra is kiterjesszem, de őszintén remélem, hogy mások segíteni fognak ebben a munkám kiterjesztésével, felülvizsgálatával és megkérdőjelezésével. A fejezet első részében a kritikai fajelméletre támaszkodom, hogy bemutassam, hogyan problematikus ez a dialektika az egalitárius demokrácia szempontjából. A következő részben rövid áttekintést nyújtok a vizuális médiában megjelenő faji előítéletek történetéről az Egyesült Államok kontextusában, amely tökéletesen illusztrálja a feketeségnek ezt a hipervizionális/ láthatatlan dialektikáját. A következő két szakaszban a feketeség hipervizibilitását és láthatatlanságát vizsgálom a kortárs mesterséges intelligencia rendszerekben, és megpróbálom bemutatni, hogy milyen hatalmas tétje van ennek a konjunktúrának. Végezetül néhány előzetes gondolatot fogalmazok meg arról, hogy merre tovább.

A faji különbség hipervisibilitása és láthatatlansága dialektikája

"A fekete az ... és a fekete nem."

Ellison p1989. 9

Michael Omi és Howard Winant (2015) a faji formációk elméletében azt állította, hogy az emberek alapvetően sztereotípiákat használnak a világ értelmezéséhez, még akkor is, ha ezek a sztereotípiák folyamatosan változnak. Az emberek a társadalomban való eligazodáshoz olyan alapvető különbségkategóriákat használnak, mint a faj, a nem, az osztály, az életkor, a nemzetiség és a kultúra, amelyek mind a "másság" bizonyos politikáját feltételezik, amely strukturális marginalizációt, egyenlőtlenséget, kizsákmányolást és elnyomást eredményez. A gépi tanulást produktívan úgy értelmezhetjük, mint egy nagyon hasonló viselkedést, hasonló politikai tettekkel. Ahogy máshol a gépi tanult szemantikai címkékkel összefüggésben érveltem, ezek az algoritmusok a sztereotipikus megkülönböztetések tanulásának iteratív folyamatában vesznek részt, hogy kategorizálják a különböző adatokat, amelyekkel találkoznak (Monea 20192016.). A faj azonban másképp működik, mint ezek a többi sztereotip kategorizálás, mivel, ahogy Omi és Winant elmagyarázza, a faj döntően "testi" és "szemmel látható". Ez alatt azt értik, hogy a faji megkülönböztetések fenotípusos különbségek egy sorát ragadják meg - az Egyesült Államokban a feketeség esetében leginkább a morfológiai különbségeket, mint például a bőrszín, az ajak mérete és a hajszerkezet -, és esszencializálják őket, mintha egy lényegi fajtakülönbség fizikai markerei lennének (2015: 13. o.). Ez tehát másságképzés, mivel létrehozta a határt egy "mi" és egy "ők" között, és újraértelmezi ezt a határt azáltal, hogy alapvető természeti törvényként, tudományos tényként, egy másfajta létezés jelölőjeként jeleníti meg. Ennek a folyamatnak két egyedi aspektusa van, amelyet Omi és Winant "faji megkülönböztetésnek" nevez. Először is, ezeket a fenotípusos különbségeket önkényesen választják ki, nem úgy értelmezik, hogy térben és időben ugyanazokkal a jelölésekkel és konnotációkkal bírnak, és gyakran korábban semmilyen faji besoroláshoz nem kapcsolódtak. Másodszor, ezeket a morfológiai megkülönböztetéseken keresztül írják rá a testre oly módon, hogy a faji különbség pusztán a látvány alapján olvasható.

A faji megkülönböztetésnek ez utóbbi aspektusa a kritikai faji megkülönböztetés egyik központi eleme volt.

Évtizedek óta az elméletben, és talán Frantz Fanon (1967) fogalmazta meg leginkább az "epidermalizáció" koncepciójában. Fanon számára az epidermalizáció egy olyan folyamat, amelynek során a feketék felismerik, hogy a fehérek számára a Másikként való azonosításuk során találkoznak a fehér tekintettel, amely testüket engedély nélkül boncolja és elemzi, hogy osztályozza őket. Fanon írja:

Külről túlhatározott vagyok. Nem annak az "elképzelésnek" vagyok a rabja, amit

mások rólam alkotnak, hanem a külsőmnek. [...] A fehér tekintet, az egyetlen érvényes tekintet, máris elbizonytalanít. *Rögzített* vagyok. Ha egyszer a mikrotomájuk ki van élezve, a fehérek objektíven kivágják a valóságom egyes részeit. (1967: p. 95)

Stuart Hall tömören úgy határozza meg Fanon epidermalizációról alkotott elképzelését, mint "a szó szoros értelmében a faj bőrré történő felírása" (1996: 16. o.). Hortense Spillers hasonlóan ír a "hús hieroglifáiról", amelyben a fekete alanyokat "a vas, az ostor, a lánc, a lánc, a kés, a kutyás őrző, a golyó kiszámított munkája" révén alakítják át hússá (2003: 207. o.). Spillers számára a nyugati humanizmus ezekre a hiero-glifákra épül, mivel a felszabadult Ember definíciószerűen megköveteli, hogy a másikat nem teljesen embernek minősítsék. Ez az örökség generációkon keresztül öröklődik, még azután is, hogy a fekete alattvalók birtokba vehették saját testüket, és továbbra is strukturálja társadalmi életünket. Ahogy Alexander Weheliye leírja, a faji kategóriák "a rabszolgaság és a gyarmatosítás mocsarából faragják ki a modern Ember húsát és csontjait" (2014: 30. o.). Sylvia Wynter (2001) hasonlóan mutatta be, hogyan történik ez a szociogenezisről szóló érveiben, ahol a fenotípusos különbségekre való összpontosítás csak ürügy a faji különbség esszencializálására és a Homo sapiens faj emberekre és nem emberekre való felosztására.

A fajnak ezt a testiségét, szemszerűségét, epidermalizálódását és hús-vérségét alapvetően *vizuális* komponensként értelmezhetjük, amely a faji jellegzetességek - különösen az olyan morfológiai jellemzők, mint a bőrszín, az ajkak mérete és a hajszerkezet - hangsúlyozásával és a faji sztereotípiák konnotálásával teszi a faji marginalizálódást erősítő faji sztereotípiákat áthatóvá. Lisa Nakamura (2007) kimutatta, hogy a fajnak ezt a hipervizibilitását napjainkban a számítás és a digitális vizuális kultúra állandósítja, ami a "digitális faji formációk" előállításához vezet. John Cheney-Lippold (2017) hasonlóképp mutatta be, hogy az algoritmusok hogyan digitalizálták a faji hovatartozást "mérhető típusokká", vagyis a felhasználói adatokon alapuló statisztikai valószínűségekké. Itt szeretném bevezetni a "színes bőrű felhasználók" kifejezést, hogy ebben a digitális kontextusban a "színes bőrű emberek" kifejezést helyettesítsem. A színes bőrű felhasználókat digitálisan rasszosítják a nagy adatok algoritmikus elemzése alapján, ami a faji különbség néhány régi fenotípusos sztereotípiáját újratermeli, miközben újakat hoz létre. Követhetnénk Simone Browne-t (2015) is, és ezt a folyamatot "digitális epidermalizációnak" is nevezhetnénk.

Könnyen beláthatjuk, hogy még az olyan technológiai vállalatok számára is, amelyek etikai szempontból kevésbé elkötelezettek az egyenlőségen alapuló demokrácia iránt, már az a rémálom, hogy a digitális kultúrában a faji sztereotípiák újratermelődését látják, hatalmas elrettentő tényező. Amint azt a cikk során látni fogjuk, az algoritmikus faji megkülönböztetés kritikájára a leggyakoribb válasz a faj láthatatlanná tétele. A feketeség láthatatlanná tétele mindig az érem másik oldala, amely dialektikus feszültségben áll a faji hipervisibilitással. Mindkét lehetőség nem kielégítő, mivel mindkettő menedéket nyújt a rasszizmusnak, bár eltérő módon. A faj láthatatlanná tételének legfrissebb és leglátványosabb példája a "színvak" politika, amelyet az Egyesült Államokban az 1960-as évek óta folytatnak. Ezeket a politikákat a kritikai faji tanulmányok és a kapcsolódó tudományágak tudósai szinte

általánosan elítélték (pl. Bonilla-Silva 2017; Brown et al. 2003; Omi/Winant 2015). A színvakság delegitimálja a pozitív diszkriminációt és a hasonló, faji egyenlőségre törekvő programokat, lehetővé teszi a rasszizmus ellenőrizetlen működését

feltéve, hogy kutyasípokát és más óvatos nyelvezetet használ a faji rosszindulat elfedésére, és elhomályosítja a fontos faji adatok nyomvonalát, amelyek egyébként felhasználhatók lennének a rasszizmus statisztikai tendenciáinak feltárására (pl. a rendfenntartás, a bírósági ítéletek, a jóléti juttatások elosztása stb. terén). Ezenkívül olyannyira tabuvá tette a fajról való beszédet, hogy a "fehér törekenység" kifejezést alkották meg a mai fehér emberek képtelenségének leírására, hogy nyíltan beszéljenek a fajról és a rasszizmusról (Dyson 2018). Így ahelyett, hogy foglalkoznánk a faji megkülönböztetés gyökereivel a gyarmatosításban és a rabszolgaságban, és elvégeznénk a tényleges egalitárius demokrácia felé való elmozdulás nehéz munkáját, a feketeséget felváltva hipervizibilisnek vagy láthatatlannak tüntetik fel, ami mindkettő sok kívánnivalót hagy maga után. Röviden, ahogy Ralph Ellison fogalmazott: "A fekete az ... és a fekete **nem**" (1989, 9. o.).

A faj láthatóságának számbavétele a vizuális médiában

"A fényképezés egy fegyver"

-Oliver Chanarin (PhotoQ 2015)

Az Egyesült Államokban a faji sztereotípiák régóta beépülnek a vizuális médiába és a kommunikációs technológiákba (vö. Dyer hooks 1997, Nakamura 1992, 2007). Ez az örökség az analóg fotográfiától a digitális fotográfiáig terjed, és - mint látni fogjuk - továbbra is hatással van számos számítógépes látásalkalmazásra. Lorna Roth (2009) például részletesen bemutatta, hogy a Kodak a teljes terméksaladját a fehér bőrre optimalizálta. A Kodak hosszú sorozatban gyártotta a "Shirley-kártyákat", amelyeket Shirley Page-ről, a Kodak új termékeivel együtt kiküldött fotók első stúdiómodelljéről neveztek el. Ezeket a Shirley-kártyákat "normál" jelöléssel látták el, és teszt-kártyaként használták a színkiegyenlítő filmalapok és nyomtatók tesztelésére. A fehér bőrre való optimalizálás a színes bőrűek rovására ment, akiknek vonásai egyre inkább eltűntek a bőrük sötétségének arányában. Az egyik legendás eredmény az volt, hogy Jean-Luc Godard francia filmrendező nem volt hajlandó Kodak-filmre forgatni egy mozambiki megbízás során, mert a filmbe keményen be volt kódolva a faji előítélet - a film szó szerint nem működött volna Afrikában. A Kodak csak akkor fejlesztette ki a fekete hús részleteit megőrkítő filmet, amikor a reklámokhoz sötét tárgyakat fényképezni kívánó cégek panaszt tettek, és a Kodaknál elhallgatták ezt a projektet, amelyre a "To Photograph the Details of a Dark Horse in Low Light" (Egy sötét ló részleteinek fényképezése gyenge fényviszonyok között) kódolt kifejezéssel utaltak (Broomberg & Chanarin n.d.).

Még akkor is, amikor a filmet azzal a kifejezett szándékkal készítették, hogy láthatóvá tegyék a fekete bőrt...

ahelyett, hogy a reklámokban láthatóvá tették volna a csokoládészeleteket és a fabútorokat - mélyen gyökerező faji problémák merültek fel. Vegyük például Eric Morgan (2006) történetét a Polaroid ID-2 fényképezőgépről. Az ID-2-t úgy tervezték,

hogy önfejlesztő nyomonként két fényképet készítsen, egy portré- és egy profilképet, egy, az objektívetől 1,2 méterre lévő alanyról. Ez lényegében Alphonse Bertillon antropometriai azonosító rendszerének racionalizálása volt az állami rendőrség számára, amely mindig is bonyolult és összetett módon keresztelte a két fényképezőgépet.

a faji és nemi alapú marginalizáció meglévő gyakorlataival (Browne 2015; Fair 2017; Lyon 2008; Wevers 2018). Az ID-2 esetében ez egy speciális "boost" gombban nyilvánult meg az f lash számára, amely körülbelül százalékkal 40 fényesebbé teszi, ugyanannyi fényt, amennyit a sötétebb bőr elnyel. Önmagában ez a funkció meglehetősen ártalmatlan. Tulajdonképpen előrelépés abban az értelemben, hogy a kamera segítségével a sötét bőrű emberek vonásait tisztábban tudta megörökíteni, bár továbbra is problémássá tette a fekete és a fehér bőr megörökítését ugyanazon a képen. Egy Caroline Hunter nevű 1970-es Polaroid vegyész azonban bizonyítékot talált arra, hogy a Polaroid rengeteg pénzt keresett az ID-2 fényképezőgépek eladásával a dél-afrikai kormánynak, amely ezeket a fényképezőgépeket használta fel a fekete állampolgároknak a fehér területeken való állandóan maguknál hordott igazolványok készítéséhez (lásd Savage 1986).

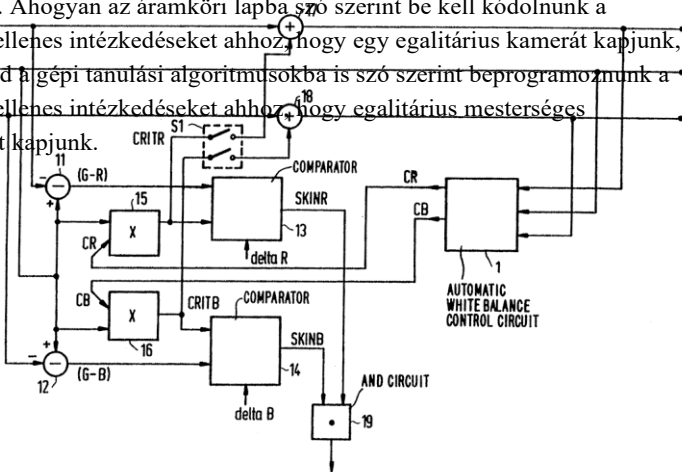
A Kodak "To Photograph the Details of a Dark Horse in Low Light" (Egy sötét ló részleteinek fotózása gyenge fényben) kódolt mondata után címet viselő kiállításukon Adam Broomberg és Alex Chanarin dél-afrikai művészek azt állítják, hogy az ID-2-t az apartheid támogatására tervezték (PhotoQ 2015; Smith 2013).¹ Nem valószínű, hogy ez így van, de az biztos, hogy a Polaroid bizonyos számú vezetője kárhózzátóan sokat tudott az ID-2-esek dél-afrikai kereskedelméről, és túl későn és túl kevés energiával dolgozott azon, hogy beavatkozzon ahhoz, hogy ténylegesen megakadályozza új technológiájuknak az apartheid támogatására való felhasználását. E történetek egyikének sem célja, hogy lekicsinyelje a sötétebb vonások és tárgyak filmre vitelének technikai nehézségeit. Ezek a nehézségek az optikai médiumok velejárói, és valószínűleg függetlenül attól, hogy a fotográfiai technológiák milyen társadalmi-kulturális kontextusban keletkeztek. Amit azonban felfednek, az az, hogy a fekete láthatatlanság és a túlságos láthatóság dialektikája mennyire hiányzott és jelentéktelen volt ezeknek a vállalatoknak, és ez a tény anyagilag is megmutatkozott a kutatási és fejlesztési paradigmáikban, a termékekben, amelyeket piacra vittek, a diskurzusban, amelyet e termékek pozicionálására használtak, és a közönség kritikájára adott válaszaikban. Nem arról van szó, hogy ezek az eszközök önmagukban rasszisták lennének, hanem arról, hogy a fejlesztésük kontextusából származó faji előítéletek befolyásolják, hogy a vállalatok hogyan közelítenek a kutatáshoz és fejlesztéshez, hogyan képzelnek el bizonyos termékeket és hogyan nem másokat, hogyan helyeznek egyeseket magasabbra, mint másokat, hogyan értékelik technológiájuk egyes potenciális rossz alkalmazásait negatív sajtóvisszhangként, amelyet a tervezési folyamat során figyelembe kell venni stb. Mindez anyagilag is megjelenik magában a technológiában, abban, ahogyan azt diszkurzívan pozicionálják, és abban a számtalan felhasználási módban, amelyre a társadalomban használják.

Csak a '90s-ben kezdték az amerikai cégek komolyan venni a fekete vonások jobb filmes megörökítésének igénye. A Kodak "fekete", "ázsiai" és "latin" Shirley-kártyákat adott ki, és elkezdte forgalmazni a Kodak Gold Max filmet, amely "bizonyítottan nagy dinamikai tartományt" kínált - végre "gyenge

fényviszonyok mellett is képes volt lefényképezni a sötét ló részleteit". Ezeket az eredményeket a filmnegatívok kémiai összetételében bevezetett újítások révén érték el, hogy a filmek jobban reagáljanak a sötétebb pigmentekkel és a modifikált negatívokkal.

I Broomberg és Chanarin kiállításáról bővebben lásd: O'toole 2014.

a színkiegyenlítési technikák optimalizálása a különböző bőrtónusokhoz. Mindezek megkövetelték, hogy a színes bőrrück láthatóságára egyre nagyobb hangsúlyt fektessenek, ami a fogyasztói bázistól a Kodakhoz hasonló vállalatok kutatási és fejlesztési csapataiig terjedt. A prioritások megváltozása gyorsan megmutatkozott a digitális fényképezőgép- és videokamera-technológiákban is. Az 1994Egyesült Államokban például a Philips Corporation szabadalmat nyújtott be egy új "színárnyalat-érzékelő áramkörre", amely képes automatikusan beállítani a színárnyalatot a digitális képen, hogy a fehéregyensúlyt a világos és a sötét bőrtónusokhoz egyszerre igazítsa. Ebben az esetben (lásd az alábbi 1 ábrát) szó szerint két alkatrészt van a lapra forrasztva, egy a "SKINR"-hez és egy a "SKINB"-hez, amelyek egy "AND CIRCUIT"-hoz vezetnek, amely kombinálja a kimeneteiket, hogy mindkettőre optimalizálja őket. Bár ez a megoldás nem ideális, mivel a színkiegyenlítés mind a világos, mind a sötét bőrű emberek esetében hibás lesz, ha ugyanabban a felvételben vannak, ez egy lépés a helyes irányba. Külön-külön bármelyik bőrtípus esetében egyformán jól működik, és ha egy képkockán szerepel, akkor az átlag a középérték felé fog közelíteni, így a Fitzpatrick-skála mindkét pólusa egyformán kevésbé lesz látható. Ezt a történelmi ívet tekinthetjük annak a beavatkozási szintnek is emblematikusnak, amely szükséges ahhoz, hogy bármely adott technológiát demokratikusabbá tegyünk a faji marginalizáció tekintetében. Ahogyan az áramköri lapba szó szerint be kell kódolnunk a rasszizmus-ellenes intézkedéseket ahhoz, hogy egy egalitárius kamerát kapjunk, úgy kell majd a gépi tanulási algoritmusokba is szó szerint beprogramoznunk a rasszizmus-ellenes intézkedéseket ahhoz, hogy egalitárius mesterséges intelligenciát kapjunk.



1. ábra: U.S. szabadalomszám (5,428,402/1995)

A faji hiperérzékenység a mesterséges intelligencia rendszerekben

David Sedaris a *This American Life* számára felolvasott "humoros" történetében az Egyesült Államok és Hollandia karácsonya körüli furcsa kulturális különbségeket írja le, és kifinomult nevetés közepette megjegyzi, hogy Hollandiában a Szent

Nicholast hat-nyolc fekete férfi kíséri, akiket az 1950-es évek közepe előtt személyes rabszolgáiként jellemeztek (Sedaris 2001; vö. Sedaris 2004). A holland hagyományban ezek a fekete segítők végül Zwarte Piet, azaz "Fekete Pete" képében stabilizálódtak: a holland férfiak és nők fekete arcfestékbe - fekete arcfestékbe, vörös rúzsba, göndör fekete hajú parókába, arany karikás fülbevalóba és színes spanyolos/mór öltözékbe - öltözve vezették Sinterklaast, azaz Szent Miklóst a felvonuláson, cukorkát és rúgást osztva a jó és rossz gyerekeknek. A hollandok többsége nagyon pozitívan viszonyul Zwarte Piethez, és gyakran lekicsinylik a közte és a fekete arc közötti összefüggéseket, és felváltva megjegyzik, hogy az arca azért ilyen színű, mert kéményseprő, aki most a kéményeken keresztül kúszik, hogy cukorkát vigyen a jó gyerekeknek, akik kint hagyják a cipőjüket, vagy mert mór, akit Sinterklaas adoptált, aki a szezonon kívül Spanyolországban él, és nem az Északi-sarkon, ahogy az amerikai hagyományokban. Ahogy Allison Blakely megjegyzi, mindkét magyarázat meglehetősen kevésbé meggyőző (2001: 47-48. o.).

A mélyen problematikus hagyomány megőrzésére tett kétségbeesett kísérletet csak bonyolítja az a tény, hogy a holland folklórban a Zwarte Piet az ördög neve, akit minden évben elkapnak és láncra vernek az ünnepekre, valamint Hollandia gyarmatosítással, rabszolgakereskedelemmel és rabszolgasággal kapcsolatos történelme - az "apartheid" kifejezés például a hollandoktól származik, és különösen Surinam gyarmati megszállásával összefüggésben alakult ki (Blakely 1980: 27. o.). Bár a hollandok úgy pozicionálják magukat, mint a gyarmati és faji marginalizációtól mentes nemzetet, ez sosem volt így, különösen azóta, hogy az 1950-es évek után számos surinamai áttelepült Hollandiába, és szembesült a faji alapú rendszerszintű marginalizációval (Blakely 1980). Nem csoda, hogy ugyanezek a surinamaiak egyre kevésbé lelkesednek a Zwarte Piet hagyományért, és azt állítják, hogy az sértő, és különösen káros a színes bőrű gyermekek számára, akik később az iskolában zaklatással szembesülnek (Blakely 2001: 48. o.).

David Leonard amelletl érvelt, hogy a blackface esetek esetében a bűnösség meghatározásának megfelelőbb mércéje nem az, hogy a blackface-nek öltözött személy szándékosan sértette-e az embereket, hanem az, hogy az illető kárt okoz-e, akár az egyéneknek, akár a társadalomnak (Desmond-Harris 2014). Ez sokkal okosabb módja a helyzet elemzésének, mivel nemcsak a szándékosság bonyolult kérdésfeltevéseitől szabadít meg, hanem megnyitja annak elemzését is, hogy a fotó- vagy videográfiai média hogyan terjesztheti ki az említett kár hatásainak körét. Például, ha egyelőre eltekintünk a hollandiai blackface kérdésétől, akkor az Egyesült Államok kontextusában minden bizonnyal rendkívül káros. Amint C. Vann Woodward (2001) kimutatta, az afroamerikaiak ellen folytatott, áltudományos módon legitimált rasszizmus, dehumanizáció, szegregáció, jogfosztás és terror kampánya, amelyet "Jim Crow" néven ismerünk, egy blackface minstrel routine-ról kapta a nevét. A blackface így elég jól összefoglalja az amerikai rasszizmus logikáját, mivel

szó szerint jelöli és láthatóvá teszi a fenotípusokat - és az állóképekben is.

ezek elsősorban morfológiai tulajdonságok - ugyanakkor mindig megőrzik a fehér felsőbbrendűséggel kapcsolatos alsóbbrendűségi konnotációjukat.

Ha megnézzük az ImageNet esetét, világosan láthatjuk, hogy ezek a Zwarte Piet-képek hogyan lépnek ki a kontextusukból, és hogyan okoznak társadalmi károkat. Az ImageNet egy nagy, címkézett képekből álló adathalmaz, amelyet először a Princeton Egyetem indított el a Princeton University-n. Az 2009. adathalmaz eredetileg a WordNet szemantikai adatbázisból merített címkéket 80,000 - minden címkét "synset"-nek nevezünk, amely szinonim kifejezések halmaza -, azzal a céllal, hogy minden egyes címkét 500-1000 tiszta és teljes felbontású képpel töltsenek fel (Deng et al. 2009). Az ImageNet 2010-ben indította el az évente megrendezett ImageNet Large Scale Visual Recognition Challenge (ILSVRC) versenyt, amely azóta a számítógépes látó alkalmazások sikerének ipari mércéjeként szolgál (Russakovsky et al. 2015). Ezt a központi jelentőséget Alex Krizhevsky, Ilya Sutskever és Geoffrey Hinton (2012) úttörő sikere cementálta, amikor neurális hálózatok segítségével megnyerték a versenyt, hogy olyan algoritmust állítsanak elő, amely képes megtanulni az ILSVRC képeket osztályozni (vö. Sudmann 2018) 2016., Fontos tudni, hogy az ImageNet nem csak a teljesítmény mércéjeként szolgál szinte minden számítógépes látórendszer számára, hanem mivel ezeket a rendszereket az ImageNet-adatkészleten képzik ki és arra optimalizálják, az ImageNet-adatkészletben lévő bármilyen torzításnak széleskörű következményei vannak, mivel ezek később a gépi tanulás révén a számítógépes látórendszerek nagy részébe keményen be vannak kódolva.



2. ábra: Zwarte Piets és Sinterklaas (Splinter2010²)

- 2 N.b. Nem tudtam megállapítani az ImageNetben található Zwarte Piet képek licencelését, de bármilyen keresés a Flickr-en a kifejezésre olyan képeket ad vissza, mint ez, amelyek szemléltetőek.

Az ImageNet egy olyan szinszethez gyűjt képeket, amely tartalmazza a "Black", "Black per- son", "blackamoor", "Negro" és "Negroid" kifejezéseket, amelyeket a következőképpen határoz meg: "sötét bőrű személy, aki Afrikából származik (vagy akinek az ősei Afrikából származnak)". Ez a szinonimakészlet több okból is érdekes. Először is, már a böngészésből is látszik, hogy az összes kép százalékában sokkal kevesebb képet tartalmaz, mint más szinszettek, amelyek hasznos vizuális részleteket tartalmaznának egy osztályozó felépítéséhez. Ez alatt azt értem, hogy a képek túl nagy része alacsony felbontású, nem mutatja az arc részleteit, a feketék teste távolabb helyezkedik el a kamerától, és túlságosan sok híresség (a teljes adathalmaz mintegy 1százaléka Barack Obama képe) és mém szerepel rajtuk. Bár ez természetesen nem füstölő pisztoly a faji elfogultságra, nemhogy a szándékos faji elfogultságra, azt mutatja, hogy a fekete arcvonások pontos azonosításának képessége alapértelmezés szerint nem élvez prioritást az ImageNet adatain kiképzett számítógépes látásmódú algoritmusoknál. Ami talán közelebb áll a füstölő pisztolyhoz, az az, hogy az ehhez a szinszethez 1,286tartozó, online még mindig elérhető képek közül a legtöbb 79 fekete arcú embereket ábrázol.³ Ezek a képek egy kivételével mind Zwarte Pietnek öltözött emberekről készültek. Ez a furcsa holland jelenség tehát pontosan azokat a következményeket vonja maga után, amelyektől a fekete arc kritikájában tartanak: függetlenül a fekete arcot viselők szándékaitól, amikor a közbeszédbe kerül, a feketeségnek ez a jelzője gyorsan kikerül minden kontextualizálás alól, és ehelyett a rasszizmust újratermeli a fehér felsőbbrendűséggel kapcsolatos konnotációi és a feketeség fenotípusos különbségre redukálható jelzői révén. Ezt szó szerint láthatjuk az ImageNet esetében, ahol a fekete arcú képek kikerültek a kontextusukból, és a teljes adathalmaz alig több mint 6százalékát teszik ki, egy olyan kontinensen kifejlesztett adathalmazban, ahol a túlnyomó többség még csak nem is hallott Zwarte Pietről. Még ha el is fogadjuk azt a rendkívül kétes érvet, hogy ezek a képek Hollandiában ártalmatlanok, egy kevesebb mint egymillió20 ember által kedvelt hagyomány segített a számítógépes látás alapvető infrastruktúrájának elferdítésében.

Vegyük például az ImageNet adathalmaz Google általi használatát. A Google 2014.re-

keresők megnyerték az ILSVRC kihívást az "Inception" algoritmusukkal (más néven "GoogLeNet"), amely egy 22 rétegű konvolúciós neurális hálózat (Szegedy et al. 2015). Ezt a CNN-t az ImageNet adathalmazon képezték ki, és így internalizálta az abban az adathalmazban jelen lévő torzításokat.⁴ A Google I/O alkalmával a Google 2015.bejelentette az új

3 Jelenleg csak az adott szinszethez tartozó képekhez tartozó linkek nyilvánosan elérhető listájából tudok dolgozni. Ez azt jelenti, hogy nem tudok hozzáférni olyan képekhez, amelyeket azóta levettek a világhálóról, ami lehetetlenné teszi, hogy hozzáférjek az adott szinszethez tartozó adatkészletet alkotó képek1,404 teljes halmazához. Többször is kértem, hogy regisztráljak az ImageNetnél, hogy hozzáférhessek a teljes adathalmazhoz, és többször küldtem e-mailt az ImageNet kapcsolattartási címére, de még nem kaptam sem hozzáférést, sem választ.

4 Míg a faji előítéletekről nincsenek összehasonlítható adatok, a nemi előítéletek esetében bebizonyosodott, hogy a neurális hálózatok nemcsak internalizálják az előítéleteket az adatkészleteikben, hanem fel is erősítik azokat (Zhao et al. 2017).

Google Photos szoftver. A Google azzal érvelt, hogy az emberek évente több mint egy trillió képet készítenek, és ilyen ütemben egy második életre lenne szükségük a fotók címkézéséhez, rendszerezéséhez és újbóli megtekintéséhez. A Google Photos volt a megoldás, amire mindannyian vártunk:

A Google Fotók automatikusan rendszerezi emlékeit a fontos emberek, helyek és dolgok szerint. Egyiket sem kell megjelölnie vagy felcímkéznie, és nem kell fáradtságosan albumokat létrehozni. Ha egy adott felvételt szeretne megtalálni, egy egyszerű kereséssel azonnal rátalálhat bármelyik fotóra - legyen szó a kutyájáról, a lánya születésnapjéről vagy a kedvenc Santa Barbara-strandjéről. És ez az automatikus csoportosítás privát, csak az Ön számára. (Google 2015)

Az új Google Photos szoftver elsősorban az Inception/GoogLeNet algoritmusra épült, amelyet az ImageNet adatain képeztek ki, bár az ImageNetből tanult képmintákat kiegészítette egy hatalmas fotóadatbázissal és az általa feltérképezett webhelyek közeli szövegével, valamint néhány egyéb mutatóval, például a keresés során a felhasználó hely- és időbélyegzőjét és a metaadatokon keresztül a képeket is figyelembe vette (Brewster 2015). Az ImageNet központi szerepe nem csoda, hiszen nemcsak a számítógépes látás algoritmusainak viszonyítási alapjául szolgál, és az iparág nagy részénél szabványos, hanem a Google Photos fejlesztése Bradley Horowitz, a Google streamerként, fotóként és megosztásért felelős alelnöke irányítása alatt történt, aki korábban már meglátta a Flickr Creative Commons licencű képeinek értékét, amikor a Yahoo vezetőjeként segített megvásárolni a vállalatot (Levy 2015).

Mindebből az a tanulság, hogy a Google Fotókat olyan emberek tervezték, akik az ImageNet-et megkérdőjelezhetetlen ipari szabványnak tekintették, és akik erősen bíztak a Flickr-képek hasznosságában. Így nem voltak abban a helyzetben, hogy előre lássák a vizuális adatokban rejlő faji torzításokat, amelyeket a gépi tanulási algoritmusai az osztályozók kifejlesztéséhez használtak. Ez a probléma alig egy hónappal a Google Photos megjelenése után került előtérbe, amikor 2015 júniusában egy Jacky Alcine nevű fekete szoftvermérnök a Twitteren közzétett egy, a Google fotómegjelölő szoftverén keresztül futtatott képsorozatot, amelyben a róla és barátjának készült képeket "gorillákról készült" fotóknak jelölték (Alcine 2015).⁵ Az eset egyértelműen ideget ütött, mivel napokon belül több tucatnyi cikk jelent meg, amelyekben rasszistának nevezték a Google algoritmusait, és azóta ez az egyik leggyakrabban idézett példa az algoritmikus elfogultságra a technológiai újságírásban. Ez az összehasonlítás csak bizonyos fenotípusos különbségek önkényes meghatározása alapján lehetséges, mint az esszencializált különbség egyedüli markerei.

5 Érdemes megjegyezni, hogy bár a Google a legjelentősebb példa erre, korántsem ez az egyetlen eset. A Flickr alig egy hónappal az Alcine incidens előtt azzal került a hírekbe, hogy egy

fekete férfit (és egy fehér nőt) tévesen "állatnak" és "majomnak" címkézett, valamint a dachau koncentrációs táborról készült fotókat "dzsungel tornateremnek", "sportnak" és "lugasnak" címkézte (Hern 2015).

ence. Röviden, amikor a bőrszínre, az ajkakra és a hajra úgy gondolunk, mint a faji különbség sarokkövére, akkor megnyílik ez a csúszás az osztályozók között. Továbbá, amikor ezeket a különbségeket esszencializáljuk és rasszista sztereotípiákhoz kapcsoljuk, ennek az elcsúszásnak a kon- jegyzálása elviselhetetlenné válik. Ez nemcsak az emberek intelligenciáját, hanem magát az emberi mivoltukat is megkérdőjelezi. Ez a probléma minden színes bőrű felhasználó számára a legnagyobb tétet jelenti.

Amint az a Google válaszából kiderül, a vállalat hasonlóan komoly problémaként értelmezte ezt az esetet, amely veszélyeztetheti a számítógépes látás platformjainak jövőjét. Néhány órán belül egy Yonatan Zunger nevű Google-mérnök válaszolt Alcine tweetjére, amelyben engedélyt kért, hogy megvizsgálhassa a Google Photos-fiókját, hogy kiderítse, mi romlott el. Másnap Zunger azt tweetelte, hogy a Google egyáltalán nem ismerte fel az arcot Alcine és barátja képein, és megjegyezte: "Dolgozunk hosszabb távú javításokon is mind a nyelvészet (az emberekről készült fotókon szereplő szavakra kell vigyázni [lang-függő]), mind pedig "magának a képfelismerésnek a területén". (pl. a sötét bőrű arcok jobb felismerése) (Zunger 2015). Zunger megígérte, hogy a Google folytatja a munkát ezeken a kérdéseken, ami magában foglalja olyan rendszerek fejlesztését, amelyek jobban fel tudják dolgozni a különböző bőrszínek és megvilágítások különböző kon- trastjait. Néhány nappal később a Google szóvivője a következőket mondta a BBC-nek: "Megdöbbenve és őszintén sajnáljuk, hogy ez történt. Azonnali intézkedéseket teszünk annak érdekében, hogy megakadályozzuk az ilyen típusú eredmények megjelenését" (BBC 2015). Mégis, ahogy a következő részben látni fogjuk, a Google még nem talált megoldást erre a problémára. Az algoritmusait képző adatkészletek miatt a faji hovatarozást vagy a fenotípusok hiperszempon- tív hangúlyozásaként kell megjeleníteni, ami egy erősen kurátori új adatkészlet nélkül továbbra is csúsztató fog produkálni a színes bőrű felhasználók és a majmok között, vagy láthatatlanként.

A faji hovatarozás láthatatlansága a mesterséges intelligencia rendszerekben

Bár a Google kezdeti válaszában gyorsasága és őszintesége ígéretesnek tűnt, több mint két év elteltével a *WIRED* arról számolt be, hogy a Google mindössze annyit tudott tenni, hogy eltávolította a potenciálisan sértő kifejezések automatikus címkézését a Fotók szoftveréből (Simonite 2018). Ugyanebben a jelentésben a *WIRED* megjegyezte a Google Fotókkal végzett kísérletsorozatának eredményeit. Először is lefuttattak a rendszeren egy 40 000 képből álló, állatokkal jól megrakott gyűjteményt, és megállapították, hogy a rendszer nem talált találatokat a "gorilla", "csimpánz", "csimpánz" és "majom" kifejezésekre. Egy második kísérletben megpróbálták kizárólag csimpánzokról és gorillákról készült képeket feltölteni, és azt találták, hogy a rendszer még mindig nem ismerte fel a kifogásolt kifejezéseket. Egy harmadik, még elmarasztalóbb teszt során a *WIRED* feltöltött egy több mint 10 000 képből álló gyűjteményt, amelyet az arcfelismerési

kutatáshoz használnak. A "fekete férfi", "fekete nő" vagy "fekete személy" automatikus címkékre történő keresés csak olyan fekete-fehér fotókat adott ki, amelyek megfeleltek a megadott nemnek, de *nem* válogatott.

emberek faj szerint. Röviden, erre a PR-katasztrófára válaszul a feketék láthatatlanná váltak a Google Fotókban.⁶ Ez a színvakság kiterjed a Google Open Images adathalmazára is, amely "30,1M képszintű címkét tartalmaz 19,8k fogalomhoz, 15,4M bounding boxot 600 tárgyszóhoz, és 375k vizuális kapcsolat anno- tációt az osztályok⁵⁷ bevonásával" (Kuznetsova et al. 2018). Ezek közül a több millió la- bels, több ezer fogalom vagy több száz osztály közül - amennyire az adatbázis vizsgálata után meg tudom ítélni - egyik sem címkézi kifejezetten a versenyt. A színes bőrű felhasználók csak a hiányukkal azonosíthatók.

A fekete láthatatlanság problémája régóta jelen van olyan rendszerekben, amelyek vizuális adatokat dolgoznak fel olyan alkalmazásokhoz, mint az arcfelismerés és a mozgásérzékelés. Vegyük például azt a 2009-es vírusos YouTube-videót, amely egy HP laptopról készült, amelyet arra terveztek, hogy az arcfelismerés segítségével kövesse a felhasználók arcát, és a webkamerával kövesse őket mozgásuk során, és amely egyáltalán nem regisztrálta a "fekete Desi" mozgását, annak ellenére, hogy fehér munkatársa mozgását könnyedén követte (wzamen01 2012). A *Consumer Reports* (2009) azzal próbálta megcáfolni azt az érvet, hogy ez faji előítélet lenne, azzal érvelve, hogy ez inkább a fényviszonyok tényezője, és bár úgy mutatják be eredményeiket, mintha a rendszer világosabb és sötétebb bőrszín esetén ugyanolyan megvilágításban ugyanúgy működne, a videójukból egyértelműen kiderül, hogy ez nem így van, mivel a színes bőrű felhasználóik arcát jobban meg kell világítaniuk, mielőtt a rendszer elkezdi követni őt. Egy másik példa az Xbox Kinect, amelyről 2010-ben azt jelentették, hogy gondot okoz a színes bőrű felhasználók arcának felismerése (Ionescu 2010). Ez elsősorban az avatárjuk automatikus bejelentkezési képességére volt hatással, mivel a Kinect játékmenet nagyrészt a csontvázmozgások alapján működik. Más szóval, a Kinect képes a fekete testeket látni, de a fekete arcokat nem, és megkönnyítheti a játékmenetüket mindaddig, amíg nem kell felismernie az arcukat, amit egyes játékok megtesznek. A *Consumer Reports* (2010) hasonlóképpen azzal érvelt, hogy ez csupán a világítással kapcsolatos probléma, és azt állította, hogy "megcáfolta" azt az elképzelést, hogy a Kinect "rasszista".

Az biztos, hogy ezek a gépek maguk nem szándékosak. előítéletes gondolkodásban részt vevő ügynökök, hanem csupán "megcáfoltnak" tekintik a rasszizmusra vonatkozó állításokat, miután bebizonyosodott, hogy a színesbőrűek által történő megfelelő működés hiánya kellemetlen mértékű önteltséget és az empátia majdnem teljes hiányát igényli. Valami nyilvánvalóan történik itt, és ez érezhető hatással van a rasszizált testekre, ami olyan egyértelmű azokon a videókon, amelyekben az arcfelismerő és mozgásérzékelő technológiák felhasználását megkísérlő és kudarcot valló emberek láthatóak. Ezek az esetek legalábbis a kutatási és fejlesztési fázisban az előrelátás hiánya, valamint az aggodalom hiánya miatt bekövetkezett hiányosságok valamilyen kombinációjának tárgyi megtestesülései.

- 6 Érdekes módon ez a láthatatlanság másképp érvényesül a Google nyilvánosan kevésbé látható számítógépes látás projektjeiben. A Google Cloud Vision API 2016-ban indult útjára, mint új eszköz, amely különösen a fejlesztők számára teszi elérhetővé a számítógépes látás algoritmusait (Google 2016). A Cloud Vision API még mindig olyan címkéket használ, mint a "csimpánz" és a "gorilla", bár eddig nem találtam vizsgálatokat arra vonatkozóan, hogy a Google Photos ugyanez a faji hipervisibilitás fennáll-e a Google Cloud Vision API-nál.

olyan termékkel piacra lépni, amely nem működne a polgárok egy nagyszámú és védett kisebbségi csoportja számára. Továbbá, az általános érvek, miszerint minden technológia kudarcot vall, hogy ezek a technikai korlátok elkerülhetetlenek, és hogy a józan ész szerint a termékeket a többségi piacra kell tervezni (és ez a legtöbb esetben a "fehér" embereket jelenti), a legjobb esetben is elégtelenek, a legrosszabb esetben pedig mélységesen sértőek.

Vegyünk egy másik példát: a nyilvános mosdókban gyakran előfordul, hogy a mozgásérzékelős eszközök, például a szappanadagolók nem ismerik fel a színes bőrű felhasználókat (pl. Fussell; Plenke 2015). Bharat Vasan, a Basis Science COO-ja a *CNET*-nek kifejtette, hogy vannak olyan rendszerek, amelyek képesek elkerülni ezt a problémát azáltal, hogy érzékelik az érzékelők alatt lévő objektumok sötétségét, és a reflektort ennek megfelelően állítják be, de ezek a rendszerek túl drágák sok ilyen alsó kategóriás mozgásérzékelős eszközhöz (Profis 2014). Itt válik a legnyilvánvalóbbá a színhasználók figyelembevételének krónikus hiánya. Ezeket a rendszereket *nyilvános* használatra tervezték, és így alapértelmezés szerint figyelembe kell venniük a színes bőrű felhasználókat. Továbbá a fényviszonyok számos aspektusa előre megjósolható (pl. fluoreszcens felülvilágítás, gyakran úgy, hogy magának a mozgásérzékelő tárgynak az árnyéka esik az érzékelendő tárgyra). Ez a kép csak még bonyolultabbá válik, ha figyelembe vesszük, hogy az automatizált mosdóinnováció kutatásának és fejlesztésének fő helyszíne a börtönökben van, amelyekben az Egyesült Államokban aránytalanul nagy számban élnek színes bőrűek (Edwards 2015).

Ez sokkal jelentősebb probléma, mint az, hogy az automatizált közművek nem működnek a színes bőrű felhasználók számára. Ugyanezek a problémák az orvosi technológiákra is kiterjednek, és korlátozzák a mindennapi orvosi megfigyelésben áttörő technológiaként hirdetett új viselhető technológiák hatékonyságát. Például a pulzoximetria, amely optikai úton méri az artériás hemoglobin oxigéntelítettségét, bizonyíthatóan kevésbé hatékony a sötétebb bőrszínű embereknél (Bickler/Feiner/Severinghaus 2005). Bár az FDA megköveteli, hogy ezek az eszközök megfeleljenek bizonyos pontossági küszöbértékeknek, mielőtt forgalomba kerülhetnek, nem írja elő, hogy a Fitzpatrick-skálán hova kell sorolniuk a vizsgálati alanyokat. Így anyagi ösztönzéssel jár a világosabb bőrű tesztlányok használata, mivel a sötétebb bőrtípusra való tervezés és felhasználás lassítja a piacra jutást. Az új, nem invazív neuroimaging technikákat, mint például a funkcionális közeli infravörös spektroszkópia (fNIRS), olyan orvosi problémák tanulmányozására és potenciális kezelésére használják, mint az Alzheimer-kór, a Parkinson-kór, az epilepszia, a traumatisztikus agysérülés, a skizofrénia, a hangulatzavarok és a szorongásos zavarok (Irani et al. 2007). Az fNIRS a vér oxigénellátottságán és térfogatán keresztül méri az agyi aktivitást a prefrontális cortexban, és hasonlóan kevésbé hatékony a sötétebb pigmentáció és a sötétebb, vastagabb haj esetén (Saikia/Besio/Mankodiya 2019). A fekete láthatatlanság ugyanezek a problémák kiterjednek a számos kortárs viselhető technológiába, például a FitBitbe és az Apple Watch-ba beépített optikai

szívkövetőkre is (Kim 2017; Profis2 014).

A kol- vagy bár a felhasználók algoritmikus láthatatlanságának talán a legjobb példája a lidar-alapú, legmodernebb tárgyfelismerő rendszerek, például az autonóm járművekben vagy az "önvezető autókban" használt rendszerek. Az autonóm járműveknek szörnyű költség-haszon számításokat kell végezniük a baleseti forgatókönyvekben, hogy meghatározzák, hogyan lehet a lehető legkevesebb embert megölni vagy megsebesíteni (pl. Roff 2018). Új kutatások jelentek meg, amelyek azt mutatják, hogy az ilyen rendszerek statisztikailag kisebb valószínűséggel azonosítják a sötét bőrű gyalogosokat emberként, hogy ezeket a számításokat figyelembe vegyék (Wilson/Hoffman/Morgenstern 2019). A kutatók megállapították, hogy ez még akkor is igaz maradt, ha figyelembe vesszük a napszakot (azaz a megvilágítást, ami a legmegfelelőbb kifogás minden olyan technológiai hibára, amely a fekete arcok és testek felismerésére irányuló technológiai hibát jelenti) és a vizuális okklúziót. Ez előre látható volt, mivel az általuk elemzett rendszerhez használt képzési adatok közel kétszer 3.5 annyi képet tartalmaztak világos bőrű emberekről, mint sötét bőrűekről (ibid.: 1. o.). A fekete láthatatlanság tehát nem pusztán identitáspolitikai kérdés, hanem szó szerint élet-halál kérdése lehet a színes bőrű felhasználók számára az egyre inkább mesterséges intelligencia által vezérelt jövőnkben.

Következtetés

A túlságosan láthatóság és a láthatatlanság dialektikája történelmileg a rasszizmus legrosszabb formáinak nyújtott menedéket az Egyesült Államokban és külföldön. A vizuális médiának a kezdetektől fogva ez a jelenség volt a jellemzője, és jelenleg bebetonozódik a mesterséges intelligencia paradigmáiba, amelyekről egyre inkább azt hisszük, hogy az emberi civilizáció következő szakaszát fogják bevezetni. A mesterséges intelligencia tervezésének e dialektikához való hozzáállása tehát központi csatátér kellene, hogy legyen mindazok számára, akik a mesterséges intelligencia demokratizálásán dolgoznak. Nem lehet egalitárius vagy demokratikus technológiánk, ha az AI-rendszereinkbe a marginalizáció már létező rendszereit kódoljuk bele. Bár e fejezet keretein túlmutat, hogy a mesterséges intelligencia-rendszerekben a hipervizibilitás/láthatatlanság dialektikájának teljes alternatíváját fogalmazzuk meg, úgy gondolom, hogy a Szilícium-völgyben még mindig nem indult el néhány ismétlődő refrén, amelyekkel legalább azt a teret megnyithatjuk, ahol egy jobb alternatívát elképzelhetünk. Megan Carcia (2017) a *WIRED* "How to Keep Your AI from Turning into a Racist Monster" című cikkében néhány józan ésszel használható megközelítést ajánl, amelyeket nagy hatásokkal lehetne támogatni és széles körben bevezetni. Ezeknek a rendszereknek jobban kellene képessé tenniük a felhasználókat az AI-rendszerek problémás összetevőinek elemzésére, hibakeresésére és elhárítására. Ez a folyamat nem csak a munka tömegesítése, ami

sokkal vonzóbbá teszi a részvényeseiknek elkötelezett technológiai vállalatok számára, hanem a folyamat során a felhasználókat is oktatja e munka fontosságáról, valamint az általuk elemezni segített rendszerek technológiai működéséről. Másodsor, a Silicon Valley-nek azonnal és további kifogások nélkül több sokszínű számítógépes programozót kell felvennie. A sokszínűbb fejlesztőcsapatok sokkal nagyobb valószínűséggel ismerik fel a keményen kódolt marginalizációt a kiadás előtt, és sta-

tisztikailag valószínűsíthetően nagyobb nyereséget termel. Harmadszor, adjunk több harmadik félnek felhatalmazást arra, hogy részt vegyen a mesterséges intelligencia rendszerek csúcstechnológiai auditálásában. Még ha az üzleti titoktartás és a spamektől való félelem meg is akadályozza az összes kód nyílt közzétételét, bizonyos megbízható harmadik feleknek, akik jobban elkötelezettek a demokrácia előmozdítása és a marginalizáció felszámolása iránt, hozzáférést kellene biztosítani ezekhez a rendszerekhez, hogy segítsenek a keményen kódolt elfogultságok ellenőrzésében. Végül pedig tovább kellene fejlesztenünk az egalitárius mesterséges intelligenciát követelő közbeszédet, és egy olyan vállalatközi eszköz- és szabványkészletet kellene létrehoznunk, amely segít a vállalatokat jobban motiválni és felelősségre vonni. Bár ezek Carcia ötletei, egyben a miénk is, mivel valahogyan az internetes diskurzusban is visszhangzanak, anélkül, hogy valaha is megvalósulnának a Szilícium-völgyben. Lehet, hogy ezt ki kell egészítenünk hagyományosabb alulról szerveződő és aktivista taktikákkal, hogy ezt az udvarias kérést követeléssé alakítsuk. A demokráciához gyakran forradalmárookra van szükség.

Hivatkozások

- Alcine, Jacky (2015); Retrieved from h29515909156865https://twitter.com/jackyalcine/status/6153 (accessed July 2019).
- BBC (2015): "Google Apologises for Photo App's Racist Blunder". In: BBC News. Retrieved from https://www.bbc.com/news/technology-33347866 (accessed July 9, 2019).
- Bickler, Philip E./Feiner, John R./Severinghaus, John W. (2005): "A bőr pigmentációjának hatása a pulzoximéter pontosságára alacsony telítettségénél". In: *Anesthesiology* 4/102, pp. 715-719.
- Blakely, Allison (2001): *Feketék a holland világban: Bloomington: The Evolution of Racial Imagery in a Modern Society*: Indiana University Press: Bloomington: Indiana University Press.
- Blakely, Allison (1980): "A Mikulás fekete segítje: A Glimpse of Race Relations in Holland". In: *New Directions* 7/2, pp. 27-29.
- Bonilla-Silva, Eduardo (2017): *Rasszizmus rasszisták nélkül: Color-Blind Racism and the Persistence of Racial Inequality in America (Fifth Edition)*, Lanham: Rowman & Littlefield.
- Brewster, Signe (2015): "How Google's New Photos App Can Tell Cats From Dogs". In: *Backchannel*. Retrieved from https://medium.com/backchannel/how-google-s-new-photos-app-can-tell-cats-from-dogs-ffd651dfcd80 (hozzáférés: 2019. július 9.).
- Broomberg, Adam/Chanarin, Oliver (n.d.): "To Photograph a Dark Horse in Low Light". Retrieved from http://www.broombergchanarin.com/to-photograph-a-dark-horse-in-low-light-1-1/ (elérés: 2019., július).
- Brown, Michael K./Carnoy, Martin/Carrie, Elliott/Duster, Troy/Oppenheimer,

David B./Schultz, Marjorie M./Wellman, David (2003): Berkeley és Los Angeles: Whitewashing Race: The Myth of a Color-Blind Society, Berkeley és Los Angeles: University of California Press.

- Browne, Simone (2015): *Sötét ügyek: Durham: On the Surveillance of Blackness*, Durham: Duke University Press.
- Consumer Reports (2010): Kinect: "Consumer Reports Debunches the 'Racist' Kinect". In: Consumer Reports. Letöltve: <https://www.consumerreports.org/cro/news/2010/11/consumer-reports-debunks-the-racist-kinect/index.htm> (július 9, 2019).
- Consumer Reports (2009): "Tényleg rasszista a HP webkamerája? Consumer Reports Weighs In". Retrieved from <https://youtu.be/NphmOV0lrBg> (July 2019,).
- Deng, Jia/Dong, Wei/Socher, Richard/Li, Li-Jia/Li, Kai/Li, Fei-Fei (2009): "ImageNet: A Large-Scale Hierarchical Image Database". In: Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition pp1., 248-255.
- Desmond-Harris, Jenée (2014): "Nem érted, mi a baj a fekete arccal? Here's Why It's So Offensive". In: Vox. Retrieved from <https://www.vox.com/2014/10/29/7089591/why-is-blackface-offensive-halloween-costume> (2019. július 9.).
- Dyer, Richard (1997): *White: New York: Essays on Race and Culture: Routledge*.
- Dyson, Michael Eric (2018): *White Fragility: Why It's So Hard for White People to Beszéljünk a rasszizmusról*, Boston: Beacon Press.
- Edwards, Philip (2015): "(És miért hibásodnak meg olyan gyakran)". In: Vox. Retrieved from <https://www.vox.com/2015/9/3/9255805/bathroom-motion-sensors> (accessed July 9, 2019).
- Ellison, Ralph (1989): *Ellison Ellison, Ellison Ellison*, New York: Láthatatlan ember, New York: Vintage.
- Fanon, Frantz (1967): *Markmann*, New York: *Fekete bőr, fehér maszkok* (Charles Lam Markmann, ford.): Grove Press.
- Fair, Freda L. (2017): "Surveilling Social Difference: Black Women's 'Alley Work' in Industrializing Minneapolis". In: *Surveillance & Society* 15/5, 655-675.
- Fitzpatrick, Thomas B. (1975): "Soleil et Peau. In *Journal de Médecine Esthétique* 33-34,.
- Fitzpatrick, Thomas B. (1988): "A napreaktív bőrtípusok érvényessége és gyakorlati hasznossága I-től VI-ig. In *Archives of Dermatology* 124/6, 869-871.
- Fussell, Sidney (2017): "Miért nem tudja ez a szappanadagoló azonosítani a bőrt?". In: Gizmo- do. Retrieved from <https://gizmodo.com/why-cant-this-soap-dispenser-iden-tify-dark-skin-179797931773> (accessed July 9, 2019).
- Google (2016): "Google Cloud és az Autodesk 10x-es javulást tesz lehetővé a médiarendelés hatékonyságában". In: Google Cloud Platform Blog. Retrieved from <https://cloudplatform.googleblog.com/2016/04/Google-Cloud-and-Autodesk-enable-10x-improvement-in-media-rendering-efficiency.html> (hozzáférés: 2019, július).
- Google (2015): "Picture This: A Fresh Approach to Photos". In: Google Blog. Retrieved from <https://googleblog.blogspot.com/2015/05/picture-this-fresh-approach-to-photos.html> (accessed July 9, 2019).

Hall, Stuart (1996): "Frantz Fanon utóélete: Miért Fanon? Why Now? Miért fekete bőr, fehér maszk?". In: Read, Alan (szerk.), A feketeség ténye: Frantz

- Fanon and Visual Representation, London: Institute of Contemporary Arts, pp. 12-37.
- Hern, Alex (2015): "Flickr Faces Panaszok a 'Offensive' Auto-Tagging for Photos". In: The Guardian. Retrieved from <https://www.theguardian.com/technology/2015/may/20/flickr-complaints-offensive-auto-tagging-photos> (accessed July 9, 2019).
- hooks, bell (1992): *Black Looks: Boston: Race and Representation: South End Press.*
- Ionescu, Daniel (2010): "Is Microsoft's Kinect Racist?". In: PCWorld. Letöltve a https://www.pcworld.com/article/209708/Is_Microsoft_Kinect_Racist oldalról. html (hozzáférés: 2019., július).
- Irani, Farzin/Platek, Steven M./Bunce, Scott/Ruocco, Anthony C./Chute, Douglas (2007): "Funkcionális közeli infravörös spektroszkópia (fNIRS): Agyi rendelkezések tanulmányozásának fontos alkalmazásai". In: *The Clinical Neuropsychologist* 21/1, pp. 9-37.
- Kim, Meeri (2017): "Wearables Still Haven't Solved the Problems of Skin Science, But New Ideas are Coming". In: *Wearable News*. Retrieved from <https://www.wearable.com/health-and-wellbeing/skin-science-complex-wearables-4441> (hozzáférés: 2019. július 9.).
- Krizhevsky, Alex/Sutskever, Ilya/Hinton, Geoffrey E. (2012): "ImageNet Classification with Deep Convolutional Neural Networks". In: *Advances in Neural Information Processing Systems* 25, pp. 1097-1105.
- Kuznetsova, Alina/Rom, Hassan/Alldrin, Neil/Uijlings, Jasper/Krasin, Ivan/Pont-Tuset, Jordi/Kamali, Shahab/Popov, Stefan/Malloy, Matteo/Duerig, Tom/Ferrari, Vittorio (2018): "The Open Images Dataset V4: Unified Image Classification, Object Detection, and Visual Relationship Detection at Scale". In: *arXiv*. Retrieved from <https://arxiv.org/abs/1811.00982> (hozzáférés: 2019., július).
- Levy, Steven (2015): "Bradley Horowitz szerint a Google Photos a képek Gmailje. And That Google Plus Is Not Dead". In: *Backchannel*. Retrieved from <https://medium.com/backchannel/bradley-horowitz-says-that-google-photos-is-gmail-for-your-images-and-that-google-plus-is-not-dead-54bed641526> (accessed July 2019).9,
- Lyon, David (2008): *Lyon Lyon: "Biometria, azonosítás és felügyelet"*. In: *Bioetika* 22/9, pp. 499-508.
- Monea, Alexander (2019): "From Aristotle to Computational Topoi". In: Sundvall, Scott (szerk.), *Rhetorical Speculations: The Future of Rhetoric, Writing, and Technology*, Logan: Utah State University Press, pp. 203-225.
- Monea, Alexander (2016): "Monea Monea: The Graphing of Difference: Google's Knowledge Graph": Numerikus mediáció és a Google Knowledge Graph esete. In: *Cultural Studies ↔ Critical Methodologies* 16/5, pp. 452-461.
- Morgan, Eric (2006): "Polaroid és Dél-Afrika". In: *Enterprise & Society* 7/3, pp. 520-549.

- Nakamura, Lisa (2007): *Minneapolis: Digitalizing Race: Visual Cultures of the Internet*: University of Minnesota Press.
- Omi, Michael/Winant, Howard (2015): *Omi: Racial Formation in the United States (Third Edition)*, New York: Routledge.
- O'toole, Sean (2014): "Making, Refusing, Remaking: Adam Broomberg és Oliver Chanarin legújabb fotográfiája." In: *Safundi* 15/ 2-3, pp. 1-14.
- PhotoQ (2015): "Broomberg & Chanarin: Chomarinberg: Low Light: Low Light". Retrieved from <https://vimeo.com/123396189> (accessed July 9, 2019).
- Plenke, Max (2015): "The Reason This 'Racist Soap Dispenser' Doesn't Work on Black Skin". In: *Mic*. Retrieved from <https://mic.com/articles/124899/the-reason-this-racist-soap-dispenser-doesn-t-work-on-black-skin#.1hOLrb9JR> (accessed July 9, 2019).
- Profis, Sharon (2014): "Do Wristband Heart Tracker Actually Work? A Checkup". In: *CNET*. Retrieved from <https://www.cnet.com/news/how-accurate-are-wristband-heart-rate-monitors/> (accessed July 9, 2019).
- Roth, Lorna (2009): "Shirley-t, a végső normát nézve: Színegyensúly, im-age technológiák és kognitív méltányosság." In: *Canadian Journal of Communication* 34/1, pp. 111-136.
- Russakovsky, Olga/Deng, Jia/Su, Hao/Karuse, Jonathan/Satheesh, Sanjeev/Ma, Sean/ Huang, Zhigeng/Karpathy, Andrej/Khosla, Aditya/Bernstein, Michael/Berg, Alexander C./Fei-Fei, Li (2015): "ImageNet Large Scale Visual Recognition Challenge". In: *International Journal of Computer Vision* 115/3, pp. 211-252.
- Savage, Michael (1986): Savage: "The Imposition of Pass Laws on the African Population in South Africa, 1916-1984". In: *African Affairs* 85/339, pp. 181-205.
- Sedaris, David (2001): Sedaris: "Don't They Know It's Christmas After All". In: *This American Life* 201. Retrieved from <https://www.thisamericanlife.org/201/them> (accessed July 9, 2019).
- Sedaris, David (2004): Sedaris: "Hat-nyolc fekete ember". In: *Dress Your Family in Corduroy and Denim*, New York: Dress Your Family in Corduroy and Denim, New York: Little Brown and Co., pp. 157-165.
- Simonite, Tom (2018): "When it Comes to Gorillas, Google Photos remains Blind". In: *WIRED*. Retrieved from <https://www.wired.com/story/when-it-comes-to-gorillas-google-photos-remains-blind/> (accessed July 9, 2019).
- Smith, David (2013): David Smith: "'Racism' of Early Colour Photography Explored in Art Exhibition." In: *The Guardian*. Retrieved from <https://www.theguardian.com/artanddesign/2013/jan/25/racism-colour-photography-exhibition> (accessed July 9, 2019).
- Spillers, Hortense (2003): "Mama babája, Papa talán': An American Grammar Book". In: *Black, White, and in Color: Essays on American Literature and Culture*, Chicago: University of Chicago Press, pp. 203-229.
- Splinter, Hans (2010): "Sinterklaas 2010". In: *Flickr*. Retrieved from <https://www.flickr.com/photos/archeon/5214550043/> (hozzáférés: 2019. július

- Szegedy, Christian/Liu, Wei/Jia, Yangqing/Sermanet, Pierre/Reed, Scott/Anguelov, Dragomir/Erhan, Dumitru/Vanhoucke, Vincent/Rabinovich, Andrew (2015): "Going Deeper with Convolutions". In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1-9.
- U.S. Patent No. 5,428,402 (1995) Retrieved from <https://patents.google.com/patent/US5428402A/en> (accessed July 2019).
- Weheliye, Alexander (2014): *Habeas Viscus: Durham: Racializing Assemblages, Biopolitics, and Black Feminist Theories of the Human*: Duke University Press.
- Wevers, Rosa (2018): "A biometriai előítéletek leleplezése: Facing Gender, Race, Class and Ability in Biometric Data". In: *Journal for Media History* 21/2, pp. 89-105.
- Wilson, Benjamin/Hoffman, Judy/Morgenstern, Jamie (2019): "Predictive Inequity in Object Detection". In: arXiv. Retrieved from <https://arxiv.org/abs/1902.11097> (accessed July 9, 2019).
- Woodard, C. Vann (2001): *New York: The Strange Career of Jim Crow*, New York: The Strange Career of Jim Crow: Oxford University Press.
- Wynter, Sylvia (2001): "A szociogén elv felé: Fanon, identitás, a tudatos tapasztalat rejtélye és milyen érzés "feketének" lenni". In: Durán-Cogan, Mercedes F./Gómez-Moriana, Antonio (szerk.), *National Identities and Sociopolitical Changes in Latin America*, New York: Routledge, pp. 30-66.
- Wzamen (2012): "HP számítógépek rasszisták". Letöltve: <https://youtube/t4DT3tQqgRM> (hozzáférés: 2019. július).
- Zhao, Jieyu/Wang, Tianlu/Yatskar, Mark/Ordóñez, Vincente/Chang, Kai-Wei (2017): "A férfiak is szeretnek vásárolni: Reducing Gender Bias Amplification Using Corpus-Level Constraints". In: arXiv. Retrieved from <https://arxiv.org/abs/1707.09457> (hozzáférés: 2019. július).
- Zunger, Yonatan (2015); Retrieved from <https://twitter.com/yonatanzunger/status/615585375487045632> (accessed July 9, 2019).

A mesterséges intelligencia demokratizálódásának feltérképezése a GitHubon

Egy első megközelítés

Marcus Burkhardt

Az elmúlt évek80 során a digitális számítógép gyökeresen megváltoztatta a világot, amelyben élünk, és azt, ahogyan hozzá és egymáshoz viszonyulunk. Hasonlóképpen a számítástechnika is gyökeresen megváltozott. Eleinte a gyakorlati számítástechnikai gépek egyedi neveket ¹viseltek, de az 1950-es évek elején a tulajdonneveket gyorsan átkerültek a sorozatjelölések, amelyek a nagyszámítógépek és az időmegosztó rendszerek korszakát jelképezik. Az 1970-es és 1980-as években megjelentek a mikro-, otthoni és személyi számítógépek, valamint a grafikus felhasználói felületek, amelyek az 1990-es években váltak uralkodóvá. Ebben az évtizedben a World Wide Web (WWW) megjelenésével a hálózatba kapcsolt számítástechnika és a hálózatépítés ismét megváltoztatta a számítástechnika arculatát. A dot-com buborék kipukkadásától függetlenül a web a 2000-es évek elején a web és a 2.0közösségi web néven átkeresztelve virágzott. Ebben az időszakban a számítástechnikai eszközök egyre mobilabbá váltak, és az asztali számítógépeket notebookok, okostelefonok és táblagépek váltották fel, a szoftverek pedig fokozatosan olyan szolgáltatásokká és alkalmazásokká alakultak, amelyek az elosztott feldolgozás és tárolás érdekében felhőinfrastruktúrákra támaszkodnak.

2017 júniusában Sundar Pichai, a Google Inc. vezérigazgatója újabb paradigmikus változást jelentett be a számítástechnika történetében. Az innováció nem úgy kell hajtani, hogy a problémákat elsősorban nem digitálisan vagy mobilként közelítjük meg, hanem a gépi tanulás területén elért legújabb eredmények által táplált AI first megközelítéssel: "Úgy gondoljuk, hogy az okostelefonoknak okosabbnak kell lenniük; tanulniuk kell Öntől, és alkalmazkodniuk kell Önhöz. Az olyan technológiák, mint az eszközön belüli gépi tanulás, képesek megtanulni a használati szokásaidat, és automatikusan előre jelezni a következő lépésedet, amivel időt takaríthatsz meg" (Pichai 2018). Ez a kijelentés tükrözi a gépi tanulási alkalmazások egyik központi ígéretét, nevezetesen azt, hogy képesek alkalmazkodni az előre nem látott jövőkhöz anélkül, hogy egy adott eseményt előzetesen beprogramoznának: olyan konkrét tárgyak vagy személyek vizuális felismerése, amelyeket a program korábban nem

"látott", és nem is képezték ki rá, önvezető autók, amelyek képesek biztonságosan kezelni az új helyzeteket, vagy chatbotok, amelyek beszélgetést folytatnak-

I A *gyakorlati számítógépek* fogalmáról lásd Turing (1992).

emberekkel való kapcsolattartás megnyerő módon. Ezzel szemben minél inkább beépülnek az ilyen technológiák a mindennapi életbe, annál több aggály merül fel a lehetséges kockázatokkal kapcsolatban, például a képzési adathalmazokban rejlő torzítások és egyenlőtlenségek miatt. Ennek eredményeképpen a gépi tanulási modellek gyakran (társadalmi) struktúrákat hoznak létre ahelyett, hogy alkalmazkodnának hozzájuk. A kritikai algoritmustanulmányok vitáira támaszkodva ez a tanulmány azt a kérdést teszi fel, hogy a gépi tanulás és a mesterséges intelligencia mint a technológiai fejlesztés és innováció területei hogyan strukturálódnak. A gépi tanulás és a mesterséges intelligencia GitHubon található *kódolási kultúráinak* első feltérképezésével a tanulmány amellel érvel, hogy a gépi tanulási technológiák társadalmi és kulturális következményeinek kritikai perspektíváinak kialakításához a kódkönyvtárak és programozási keretrendszerek eddig nagyrészt elhanyagolt infrastrukturális rétegeit is jobban figyelembe kell venni.

A mesterséges intelligencia demokratizálása

Az elmúlt években a mesterséges intelligencia (AI) és különösen a gépi tanulás (ML) iránti érdeklődés ismét az egekbe szökött. Ezt a folyamatos fejlődést bizonyos mértékig a vezető technológiai vállalatok, például a Google és anyavállalata, az Alphabet, az Amazon Web Services (AWS), a Facebook, az IBM, a Microsoft stb. hajtják. Ez egyrészt az e vállalatok által felhalmozott adatok tömeges felhalmozásán, másrészt a nagyméretű felhőinfrastruktúrák és infrastrukturális szolgáltatások létrehozásán alapul. Ezek a vállalatok azonban nem egyszerűen csak hozzájárulnak a mesterséges intelligencia és az ML gyors technológiai fejlődéséhez, hanem részt vesznek az intelligens, intelligens és autonóm technológiákról alkotott elképzelések kialakításában is, amelyek a technológiai fejlődés sarokkövei és a társadalmi fejlődés, valamint a gazdasági jólét alapkövei lesznek az elkövetkező években.

Ebben központi szerepet játszik a mesterséges intelligencia demokratizálására irányuló legújabb törekvés. A Google (IANS 2017), az IBM (Moore 2018), az Apple (Simonite 2017) és a Microsoft (n.d.) egyaránt mozgósítja a demokratikus mesterséges intelligencia fogalmát, hogy elősegítse az ML által vezérelt technológiai innováció felé való elmozdulást. Ebben az összefüggésben a demokratizálás "olyan cselekvésként/fejlesztésként értelmezhető, amely valamit mindenki, a "köznép" számára hozzáférhetővé tesz" (Schmarzo 2017). A Microsoft számára ez azt jelenti, hogy "minden ember és minden szervezet" (n.d.) részesülhet az AI várható előnyeiből, amelynek hatásai állítólag olyan messzemenőek lesznek, mint a nyomtatásé:

Az 1400-as években a nyomtatás megjelenésével az információ robbanásszerű terjedése következett be - ez volt az első demokratizáló esemény a hozzáférés

körül, amely lehetővé tette, hogy az emberek mindenhol elkezdjenek tanulni. Az információhoz való hozzáférés onnantól kezdve csak terjedt. [...] A kérdés az, hogy hogyan használhatjuk ki mindazt, amink van a számítástechnika terén.

hatalom, hogy megoldja ezt az alapvető kényszert? Hogy jobban megértsük a világot? Ez a mesterséges intelligencia lényege. Nem arról van szó, hogy a mesterséges intelligencia legyőzi az embert a játékokban, hanem arról, hogy mindenki többet érjen el - emberek és gépek együtt dolgoznak azon, hogy a világot jobb helyé tegyék. (Ibid.)

Tekintettel a mesterséges intelligenciával kapcsolatos eltúlzott várakozások hosszú történetére, indokolt némi szkepticizmus a forradalmi cezúra állításával kapcsolatban. Ennél sokkal fontosabb azonban az a kérdés, hogy a mesterséges intelligenciát hogyan teszik ténylegesen hozzáférhetővé, azaz hogyan valósul meg a demokratizálódás. Ismét a Microsoft esete szolgálhat paradigmátikus példaként. A vállalat négyes stratégiát követ:

(1) a mesterséges intelligencia felhasználása a környezeti számítástechnológiákkal való interakció új módozatainak kifejlesztésére; (2) intelligencia beépítése minden alkalmazásba; (3) a fejlesztők számára az "intelligens képességek" felhasználásának lehetővé tétele; (4) a számítástechnikai infrastruktúra szolgáltatásként történő rendelkezésre bocsátása (ibid.).

A demokratizálás ígérete a technológia fejlesztőinek és felhasználóinak egyaránt szól. A fejlesztők számára ez a demokratizálás azt jelenti, hogy saját termékeikben felhasználhatják a mesterséges intelligenciát, és részt vehetnek a mesterséges intelligencia jövőjének alakításában azáltal, hogy nyílt vagy fizetős hozzáférést kapnak az erőforrásokhoz és szolgáltatásokhoz, például a szoftverkönyvtárakhoz, az előre betanított gépi tanulási modellekhez, a keretrendszerekhez, a platformokhoz és az infrastruktúrához. Másrészt a felhasználók a mesterséges intelligencia és a gépi tanulás által "átítatott" (ibid.) technológiák kedvezményezettjeként vesznek részt a mesterséges intelligencia demokratizálásában. Az ilyen technológiákban az intelligencia jellemzően korlátozott hatókörű képességek és funkciók széles skálájaként jelenik meg: sakkozó vagy go-játékot játszó szoftveralkalmazások, arcfelismerő és mosolygós képeket készítő kamerák, hangszórók, amelyek képesek felismerni, értelmezni és végrehajtani a hangalapú parancsokat, autonóm módon közlekedő autók, chatbotok, amelyek szórakoztató, hasznos vagy informatív beszélgetéseket folytatnak az emberekkel stb.

A gépi tanulás gyakorlata

Az ML-technológiák kapacitásait úgy tervezték és úgy állították be, hogy meglepjék a felhasználókat. A tudomány- és technológiatanulmányok, valamint a médiatudományok számos kutatója közül Kate Crawford és Ryan Calo amellet érvelt, hogy "fel kell mérni a technológiák társadalmi, kulturális és politikai környezetükre gyakorolt hatását" (2016: 311). Fontos tehát, hogy kritikusan megértsük a gépi tanulási technológiákat általában, és különösen a demokratikus mesterséges intelligencia felé irányuló jelenlegi törekvéseket. Ez a kritikai

megértés annál is inkább jelentős, mivel a modern kori mély neurális hálózatok, valamint más ML-megközelítések elvileg elvileg kikerülnek az emberi megértést.

A vállalatok által az algoritmikus rendszerekre rótt titkolózás és a gépi tanulási rendszerek belső átláthatatlansága ellenére (vö. Burrell 2016), az ML-technológiákról sok mindent lehet tudni, ahogyan Adrian Mackenzie ékesszólóan érvelt:

A gépi tanulás ma már aligha számít homályos vagy titkos tudásnak. Ezeket a technikákat tankönyvekben [...], hogyan kell használni könyvekben [...], valamint számos videón és weboldalon bemutatott oktatóanyagban, előadásban és demonstrációban [...] részletesen dokumentálják. Többé-kevésbé olvashatunk, sőt játszhatunk is a szoftveres megvalósításokkal [...]. (2015: 43 lp.)

Az ilyen sokszínű forrásokra támaszkodva Mackenzie maga is kritikus tanulója, gyakorlója és kutatója lett a többszörösen szituált, hibrid gépi tanulási gyakorlatoknak. A *Machine Learners* című könyvében Mackenzie gyakorlatias vizsgálatot mutat be "a gépi tanulásról mint a tudás előállításának formájáról és a hatalom stratégiájáról" (ibid.: 9). Foucault archeológia fogalmát követve Mackenzie hat, az ML-ben központi szerepet játszó működési formáció archeológiáját bontja ki:

vektorizáció, optimalizálás, valószínűsítés, mintafelismerés, regularizáció és szaporítás. Ezek az általános műveletek a hardver- és szoftverarchitektúrákat, az adatok és adathalmazok szervezését, a modellek tervezésének és tesztelésének gyakorlatát, a tudományos és mérnöki tudományágak, valamint a szakmai és népszerű pedagógia metszéspontjait átívelő gépi tanulási diagramban metszik egymást. (ibid.: 18)

Az a megközelítés, hogy a gépi tanulást nem távolról, hanem "a gépi tanulás megtanulása" (ibid.: 18) jól rezonál Wendy Chun állításával, miszerint "a szoftvereket csak *a médiában* lehet megérteni" (Chun 2008: 323). Az ML-t ma alkotó dolgok közepette Mackenzie nemcsak tankönyvekkel, oktatóanyagokkal, matematikai képletekkel, algoritmusokkal és adatkészletekkel, hanem számos szoftverkönyvtárral is foglalkozik. Bár a kódkönyvtárakat és a programozási keretrendszereket gyakran a mai szoftverkulturúra döntő infrastrukturális elemeként emlegetik, a kritikai kutatások alig vizsgálják őket közelebbről (vö. Berry 2011; Marino 2014). A *Machine Learners* is gyakran elismeri a kódkönyvtárak fontosságát, de nem kutatja őket részletesen. Mackenzie azonban mellékesen nyújt néhány értékes betekintést abba, hogy a kódkönyvtárak és programozási keretrendszerek hogyan alakítják a gépi tanulás gyakorlatát azáltal, hogy "a szabványos műveletek, minták és funkciók repertoárját kristályosítják ki az adatok átformálására és olyan modellek felépítésére, amelyek a dolgok, emberek, folyamatok stb. közötti eseményeket és asszociációkat osztályozzák és jósolják meg" (Mackenzie 2017: 23). Az architektúrájuk "osztályozza és rendezi" (ibid.: 77) a gépi tanulást, mint egymással összefüggő gyakorlatok területét éppúgy, mint

a tudáspro- dukció területét. Ezek alkotják a "kódolási és kapcsolódó adatgyakorlatok felhalmozódó üledékét [...], amelyben a gépi tanulás gyökeret ereszt" (ibid.: 23). Mackenzie számára a

a kódkönyvtárak megvalósítása és széles körű használata tehát a kortárs *kódolási kultúra* megnyilvánulása.

A kódkönyvtárak olyan erőforrásokat biztosítanak, amelyekből a fejlesztők meríthetnek és amelyekre építhetnek. Azáltal, hogy előre meghatározott funkciókat és funkcionalitásokat kínálnak, a fejlesztőket mentesítik a szoftverek alapról történő felépítése alól. Ugyanakkor a kódkönyvtárak funkcionális logikájukat és gyakorlati lehetőségeiket a fejlesztőkre kényszerítik. Ebben a regardában a kódkönyvtárak a "kooperatív cselekvés" médiumainak tekinthetők, amelyek olyan felhalmozott erőforrásokból állnak, amelyeket működőképes szoftverre lehet *laminálni* (Goodwin 2018: 129). A kódkönyvtárak azonban az együttműködés helyszínei is. Ezeket a kódmegeosztó platformok, például a GitHub "rekurzív nyilvánosságában" (Kelty 2008: 28) hozzák létre, járulnak hozzá, karbantartják, frissítik és törlik őket.²

A "kódolás Facebookjaként" (Wulf 2017) a GitHub ma már több mint 96 millió tárolónak ad otthont (GitHub 2019), amelyek többsége kódokat vagy kódolással kapcsolatos újraforrásokat tartalmaz. A mai kódolási kultúra számára a GitHub a nyílt és zárt forráskódú szoftverprojektek befogadásának, valamint a kódforrások, köztük könyvtárak és keretrendszerek megtalálásának, közreműködésének és megvitatásának súlypontjaként szolgál (Mackenzie 2018). Ez különösen érvényes az ML jelenlegi fejlesztéseire. Mégis, pontosan hogyan alakul ki az ML mint gyakorlat és terület a GitHubon? Hogyan keringenek a GitHubon az ML algoritmikus technikái, az adathalmazok, a gépi tanult modellek és más erőforrások? Mely szereplők vesznek részt az együttműködés, a csere és az újratárgyalás e terének kialakításában, és milyen hatalmi stratégiákat alkalmaznak? Vagy másképp fogalmazva: Hogyan valósul meg a mesterséges intelligencia demokratizálódása a kód és a kódmegeosztási gyakorlatok infrastrukturális szintjén?

A mesterséges intelligencia demokratizálódásának feltérképezése a kódban

A GitHubon a gépi tanulás és a mesterséges intelligencia számos heterogén artikulációjának és megnyilvánulásának feltérképezéséhez a kutatók igénybe vehetik a platform által biztosított alkalmazásprogramozási interfészt (API), amely lehetővé teszi a tároló metaadatainak lekérdezését, amelyek megfelelnek bizonyos keresési kritériumoknak, például a tárolók leírásában szereplő konkrét kulcsszavaknak vagy a létrehozó által a tárolóhoz rendelt témáknak.³ A GitHub kereső API azonban bizonyos korlátozásokkal jár: lekérdezésenként csak korlátozott számú kulcsszót engedélyez, és minden egyes lekérdezésre csak maximális számú eredményt 1.000 ad vissza. Bár ez elegendő lehet ahhoz, hogy egy adott kulcsszóra vonatkozóan feltárjuk a GitHubon leglátványosabb projekteket a következő szempontok szerint

- 2 A verziókezelő rendszerek és a GitHub jelentőségéről a kortárs kódolási kultúrák szempontjából lásd Burkhardt (2019) és Mackenzie (2017, 2018).
- 3 A kódolási és megosztási kultúrák hasonló feltérképezését végezte el Kollanyi (2016) a botok tekintetében.

a forkok vagy csillagok száma, amelyeket egy tároló kapott, ez a korlátozás figyelmen kívül hagyja az ML-hez és a mesterséges intelligenciához kapcsolódó kulcsszavak sokféleségét, valamint a szoftverfejlesztés és a kódmegosztás gyakorlatának hosszú farkát az ML területén. A gépi tanulás megvalósításainak és materializációinak átfogóbb feltérképezése érdekében olyan keresési taktikát kell alkalmazni, amely az ML-hez és a mesterséges intelligenciához kapcsolódó tárolók terét iteratív módon térképezi fel a keresési kifejezések széles listájának azonosításával és feldolgozásával, valamint a keresési paraméterek (programozási nyelvek, elágazások száma, kapott csillagok száma és mérete) lépésről lépésre történő korlátozásával:

keresési feltételek

keresőszavak + programozási nyelv

keresési kifejezések + programozási nyelv + villaszám

keresőszavak + programozási nyelv + villaszám + csillagszám

keresési kifejezések + programozási nyelv + villaszám + csillagszám + méret

Ez a keresési taktika egy dinamikusan változó környezetről nyújt pillanatfelvételt. Az eredmények hiányosak maradnak, de átfogóbbak, mint a lekérdezésenkénti 1000 találatból álló de- fault korlát.⁴ Ami kimarad, és elvileg láthatatlan marad, az a GitHubon tárolt összes privát tároló.

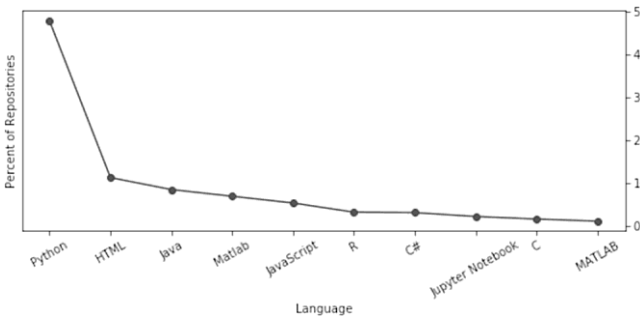
Az adatkészlet összeállítása csak az első lépés a gépi tanuláshoz a kortárs kódolási kultúrában való megjelenésének feltérképezésében. Az adatkészlet 211.802 egyedi tárolót tartalmaz.⁵ Ezek közül 41.818-ban szerepel kulcsszóként a mesterséges intelligencia⁶, 103.344-ben pedig a gépi tanulás⁷. Figyelemre méltó, hogy mindössze 3 064 tároló hivatkozik mindkét fogalmi területre, noha a közbeszédben a mesterséges intelligencia jelenlegi nyara nagyrészt a gépi tanulás területén végbement fejlesztéseknek tulajdonítják.

4 Az eredményhalmazok bizonyos hézagai pontosan meghatározhatók: Az eredményhatár olyan programozási nyelvek esetében, mint a Python, amelyek nem forkoltak és nem kaptak csillagot, és viszonylag kicsik (a Python esetében ez 0, 1, 1, 2, 2, 3 ..., 11 KB). Más esetekben azonban a GitHub API által visszaküldött eredmények feltételezett száma és a ténylegesen lekérdezett eredmények száma közötti ellentmondások formájában hiányosságok jelennek meg.

5 Keresési feltételek: Mesterséges intelligencia, mesterséges intelligencia, mesterséges intelligencia, mesterséges intelligencia, gépi tanulás, gépi tanulás, gépi tanulás, gépi tanulás, mélytanulás, mélytanulás, mélytanulás, mélytanulás, neurális háló, neurális háló, neurális háló, neuralnet, neuralnet, neuralnet, neural-net, neural networks, neuralnetworks, neural-networks, neural network, neuralnetwork, neural-network, neural, bigdl, caffe, caffe2, cntk, coreml, deeplearning4j, keras, lasagne, mlb, mlpack, moa, mocha.jl, mxnet, neon, Paddle Paddle, pylearn, pylearn2, pytorch, scikit-learn, shotgun, singa, tensorflow, tflearn, theano.

6 Változatok: mesterséges intelligencia, mesterséges intelligencia, mesterséges intelligencia, ai.

7 Változatok: machine learning, machine-learning, machine-learning, ml.

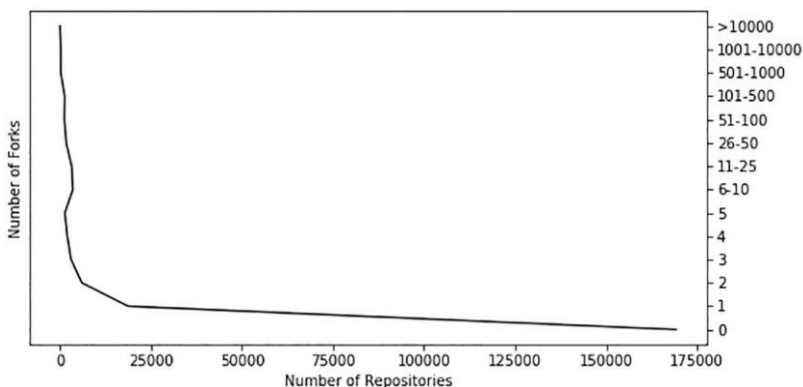


Nyelv	%
Python	47.74
HTML	11.25
Java	8.46
Matlab	6.90
JavaScript	5.33
R	3.19
C#	3.09
Jupyter Note- könyv	2.18
C	1.59
MATLAB	1.10
Összesen	90,83

1. ábra: A lekérdezett adattárakban használt legfontosabb programozási 10 nyelvek eloszlása.

A gépi tanulás és a mesterséges intelligencia számos programozási nyelvben materializálódik. A *Python* (és a Python-alapú *Jupyter Notebooks*) messze a leggyakrabban használt programozási nyelv az ML számára. Figyelemre méltó módon ezt követik tizenegy százalékban a HTML-tárhelyek. Amint azt később tárgyalni fogjuk, ennek oka, hogy a GitHubon nem minden tároló tartalmaz kódot mint erőforrást. Sokan információs és oktatási erőforrásokat tartalmaznak, például oktatóanyagokat, könyvkéziratokat, kutatási tanulmányokat, tananyagokat vagy linkgyűjteményeket.

A tárolók népszerűségére vagy relevanciájára a kapott elágazások és csillagok számából lehet következtetni. A GitHub és általában a Git verziókezelő rendszer kontextusában a forkok a tárolók másolatai. A forkolás a Git-alapú együttműködés egyik központi gyakorlatát jelenti: "Leggyakrabban a forkokat arra használják, hogy változtatásokat javasoljanak valaki más projektjéhez, vagy hogy valaki más projektjét használják kiindulópontként a saját ötletükhöz" (GitHub a). A csillagok ezzel szemben platformspecifikus jelzők, amelyek a felhasználók valamilyen érdeklődését jelzik egy adattár iránt: "Csillagokat adhatsz tárolókra és témákra, hogy nyomon követhesd az általad érdekesnek talált projekteket, és felfedezd a kapcsolódó tartalmakat a hírfolyamodban" (GitHub b).



2. ábra: Az elágazások eloszlása tárolók szerint

A csillagok és elágazások eloszlása a tárolók között a hatványtörvényt követi. Csak kevés adattár kap nagy láthatóságot, míg a legtöbb adattárnak nincs elágazása, és alig vagy egyáltalán nem kapott csillagot. A gépi tanulás hosszú farka a kipróbálás, a kísérletezés és a tanulás útján történő tanulás terepe. Ezt a teret azonban kurzusfeladatok, folyamatban lévő eredeti kutatások, a gépi tanulással kapcsolatos személyes forrásgyűjtemények és új, feltörekvő, sikertelen vagy elhagyott kódkönyvtárak is benépesítik. A cikkhez visszakeresett több mint egyedi 200.000 tárolók közül kevesebb mint 0,8 százaléknak van több mint 100 vagy annál több elágazása, és csak egy százalék 1.8kapott vagy 100 több csillagot.

	Teljes szám	≥100 villa	≥100 csillag
Mesterséges intelligencia adattárak	41.818	284	593
Gépi tanulási adattárak	103.344	799	1.677
Mélytanulási adattárak	33.749	857	1.842
Neurális hálózati adattárak	46.681	558	1.365
Minden tároló	211.802	1.751	3.951

A programozási nyelvek használatára vagy a népszerűség eloszlására vonatkozó magas szintű megfigyelések első pillantást engednek arra, hogy a mesterséges intelligencia és a mesterséges tanulás hogyan artikulálódik a GitHubon. A 200 leggyakrabban forkolt és csillagozott tároló részletesebb elemzése azonban feltárja a fejlesztett, közzétett, együttműködő, karbantartott, vitatott, frissített és letöltött erőforrások heterogenitását és sokféleségét. A 200 legnépszerűbb adattár között 42 olyan található, amely a kódkönyvtárak vagy programozási keretrendszerek kategóriájába sorolható. Ugyanakkor több mint

ötven százalékuk információs és oktatási forrásokat tartalmaz. A táruk mintegy 15 százaléka algoritmusok referencia implementációit, gépi tanulással tanult modelleket kínál konkrét alkalmazási területekhez vagy gépi tanuláson alapuló szoftveralkalmazásokat. És csak kevés adattár nyújt infrastrukturális erőforrásokat, programozási nyelveket, kísérleteket vagy adathalmazokat az ML-hez.

Tipus	#
Információs/oktatási forrás	115
Könyvtár/Keretrendszer	
Algoritmusok/Modellek/Alkalmazások	42
sok	31
Infrastruktúra/Optimalizálás	8
Kísérletek	2
Nyelvek	1
Adatkészletek	1

Az információs és oktatási források különböző célközönségeket szólítanak meg, mint például a gépi tanulással kezdők, a keretrendszerek tanulói, a gépi tanulással foglalkozó szakemberek és a kutatók. Ennek eredményeképpen egyes tárházak a gépi tanulás alapjairól szóló re- forrásokat, online tanfolyamokhoz való anyagokat, különböző kódkönyvtárakhoz tartozó oktatóanyagokat vagy bevezető könyvkiadványokat kínálnak. Mások *gyűjtemények, átfogó listák vagy kurátori listák* a gépi tanulásról általánosságban, a legfrissebb kutatási eredményekről vagy specifikusabb témákról, például infrastruktúrákról. Ezeket a forrásokat részben egyéni fejlesztők, újrakeresők vagy szerzők hozzák létre, állítják össze és bocsátják rendelkezésre. Sok információs és oktatási tárat azonban vállalatok hoznak létre és tartanak fenn, hogy a felhasználókat *megismertessék* és toborozzák a termékeikhez. Az *Amazon Web Services Labs* által kiadott *amazon-sagemaker-examples*⁸ repozitórium például alapvető információkat és oktatási anyagokat tartalmaz a vállalat *SageMaker* platformjának a gépi tanulási modellek képzésére, optimalizálására és telepítésére való használatáról. Itt válnak láthatóvá a gépi tanulási gyakorlatok infrastrukturális összetettségei. A gépi tanulás az adatbázis-alapú programozás módjaként értelmezve nagy feldolgozási kapacitásokra támaszkodik, amelyeket ma felhőalapú számítástechnikai szolgáltatások formájában nyújtanak - egy olyan globális technológiai vállalatok által uralt piacon, mint az Amazon, a Microsoft, a Google, az IBM stb. (Flexera 2019). Mi több, az oktatási és információs források gazdag sokféleségével való foglalkozás feltárja a programozási könyvtárak központi szerepét a kódolási kultúrákban. A gépi tanulás megtanulása mélyen összefonódik a kódok használatának megtanulásával

8 <https://aws.amazon.com/de/sagemaker/>

könyvtárak és keretrendszerek. Míg vannak olyan tárházak, amelyek a "gépi tanulás a semmiből" feladatának szenteltek⁹, addig az oktatási források inkább arról ígérnek információt, hogy hogyan lehet gépi tanulást végezni x könyvtárral vagy y keretrendszerrel.

Az információs forrásokkal rendelkező tárolók sokféleségét tükrözi a kódtárolók het- erogenitása. Napjainkban a ma- chine tanulás és a mesterséges intelligencia területén messze a legnépszerűbb tárház a TensorFlow keretrendszer.¹⁰ Az eredetileg a Google Brain által belső használatra kifejlesztett keretrendszert a Google 2015nyílt forráskódúként adta ki, és azóta is a vállalat tartja karban, de számos külső közreműködőt is vonzott. A TensorFlow a mélytanulás keretrendszere, amely elsősorban az új- rális hálózatok képzését és telepítését támogatja, és így a gépi tanulás egy sajátos, az elmúlt évtizedben dominánssá vált paradigmájának szenteli magát. Más népszerű könyvtárak, mint például a *scikit-learn*¹¹, a gépi tanulás más területeire összpontosítanak, és ennek következtében a megközelítések szélesebb körét támogatják.¹² Bár a scikit-learn segítségével alapvető neurális hálózatok is megvalósíthatók, ebben a könyvtárban kissé marginális szerepet játszanak. Ezt a mély és megerősítő tanulás, valamint a GPU-hardverek (scikit-learn és 2019,5 7) használatának hiánya is alátámasztja. A TensorFlow és a scikit-learn tehát a gépi tanulás különböző artikulációi és materializációi. Abban azonban hasonlítanak egymáshoz, hogy mindkettő általános célú keretrendszereknek tekinthető, azaz nem kifejezetten konkrét alkalmazási feladatokra összpontosítanak.

Számos speciális célú gépi tanulási könyvtár is viszonylag nagy népszerűsége telt szert. Ezek közé tartozik a *Unity ML-Agents Toolkit*¹³, az *OpenCV*¹⁴, valamint a *ChatterBot*¹⁵ és a *Rasa*¹⁶ keretrendszerek. Az ML-Agents Toolkit lehetővé teszi az úgynevezett gépi tanuló ágensek implementálását a Unity játékmotorba. Az eszközkészlet egy játékkészítő környezet pluginjaként természetesen ezt az alkalmazási területet célozza, de lehetővé teszi a ro- botikai algoritmusok fejlesztését vagy az önvezető autók virtuális környezetben történő betanítását is (vö. Unity n.d.). A Chatterbot és a Rasa AI-alapú chatbotok megvalósítására szolgáló keretrendszerek. A gépi tanulás itt úgy artikulálódik, mint az előzetesen betanított modellek használata a természetes nyelvi megértéshez és a párbeszédkezeléshez, valamint a folyamatos

9 <https://aws.amazon.com/eriklindernoren/ML-From-Scratch>

10 <https://github.com/tensorflow/tensorflow>

11 <https://github.com/scikit-learn/scikit-learn>

12 A gépi tanulási paradigmák és megközelítések átfogó áttekintését lásd Domingos (2015).

13 <https://github.com/Unity-Technologies/ml-agents>

14 <https://github.com/opencv/opencv>

15 <https://github.com/gunthercox/ChatterBot>

16 <https://github.com/RasaHQ/rasa>

az ilyen modellek javítása a bétatesztek során és a valós világbeli telepítés során gyűjtött képzési adatok alapján. Az ilyen keretrendszerek nem általában a gépi tanuláshoz nyújtanak kódforrásokat a szakterület-specifikus szoftverek fejlesztéséhez, hanem valamilyen gépi tanulási elemet építenek be. Ez felveti azt a kérdést, hogy a speciális célú keretrendszerek hogyan strukturálják a gépi tanulási gyakorlatokat sajátos módon, valamint hogyan írják elő a velük létrehozott alkalmazások működési logikáját és teljesítményét. A chatbotok esetében ez azt jelenti, hogy megkérdezzük, hogyan írják be a *társalgási* és *kommunikációs* képességet az ilyen keretrendszerek.

Az általános és a speciális célú könyvtárak mellett legalább egy harmadik típus is megkülönböztethető, amelyet meta-könyvtáraknak nevezhetünk. Az ilyen típusú könyvtárak egyik népszerű példája a *Keras*¹⁷. Ez a TensorFlow, a Theano¹⁸ és a CNTK¹⁹ mélytanulási keretrendszerek tetejére épül, és egységes felületet biztosít a különböző kódkönyvtárak sokaságához. A Keras tehát egy absztrakciós réteget ad hozzá, amely frontendként szolgál több mélytanulási könyvtár backendjéhez. E könyvtárak közül kettőt globális vállalatok fejlesztenek - a TensorFlow esetében a Google, a CNTK (Cognitive Toolkit) esetében pedig a Microsoft. A harmadik keretrendszert főként tudományos környezetben fejlesztették ki, és a Montreal Institute for Learning Algorithms tartja fenn. Aktív fejlesztésének végét 2017-ben jelentették be, többek között arra hivatkozva, hogy "erős ipari szereplők támogatnak különböző szoftver stackeket egy ösztönző versenyben" (Bengio 2017). Valóban, nagy globális digitális technológiai vállalatok versenyeznek a nyílt forráskódú gépi tanulási könyvtárak terén a megfelelő keretrendszerekkel. A *PyTorch*²⁰ (Facebook), a *Neo-AI-DLR*²¹ (Amazon), az *Aerosolve*²² (AirBnB), valamint a már említett keretrendszerek TensorFlow (Google) és CNTK (Microsoft) csak néhány példa. A Google, a Facebook, az Amazon, a Microsoft stb. által nyílt forráskódú gépi tanulási könyvtárak kiadása a mesterséges intelligencia demokratizálására tett erőfeszítésként értelmezhető. Ugyanakkor a nyílt forráskódú kódolási kultúra a vállalati beavatkozás és verseny terepévé vált. Hogy ez hogyan befolyásolja a gépi tanulási technológiák jövőjét, valamint azt, hogy a gépi tanulás hogyan épül be technológiai világunk szövetébe, még megválaszolatlan kérdések. A kritikai kutatásnak tehát még jobban oda kell figyelnie a kódkönyvtárak és programozási keretrendszerek logikájára és politikájára.

17 <https://github.com/keras-team/keras>

18 <https://github.com/Theano/Theano>

19 <https://github.com/microsoft/CNTK>

20 <https://github.com/pytorch/pytorch>

21 <https://github.com/neo-ai/neo-ai-dlr>

22 <https://github.com/airbnb/aerosolve>

Irodalom

- Bengio, Yoshua (2017): "MILA and the Future of Theano." 2017. <https://groups.google.com/forum/#!topic/theano-users/7Poq8BZutbY>.
- Berry, David M (2011): *Berry Ber: The Philosophy of Software: The Philosophy of Software: Basingstoke: Code and Mediation in the Digital Age: Palgrave Macmillan.*
- Burkhardt, Marcus (2019): "Version Control. Zur Softwarebasierten Koordination von Ko-Laboration". In: Gießmann, Sebastian/ Röhl, Tobias/Trischler, Ronja (szerk.), *Materialität Der Kooperation*, Wiesbaden: Springer VS, pp. 91-117.
- Burrell, J. (2016): "Hogyan gondolkodik a gép: Understanding Opacity in Machine Learning Algorithms (A gépi tanulási algoritmusok átláthatatlanságának megértése)". In: *Big Data & Society* 3/1. <https://doi.org/10.1177/2053951715622512>.
- Chun, Wendy Hui Kyong (2008): "A forrásszerzésről", avagy a kód mint fétis. In: *Configurations* 16/3, pp. 299-324. <https://doi.org/10.1353/con.0.0064>.
- Crawford, Kate/ Calo, Ryan (2016): "Van egy vakfolt a mesterséges intelligencia kutatásában". In: *Nature News* 538/7625, pp. 311-13. <https://doi.org/10.1038/538311a>.
- Domingos, Pedro (2015): *A mesteralgoritmus: How the Quest for the Ultimate Learning Machine Will Remake Our World*, New York: Basic Books.
- Flexera (2019): "RightScale State of the Cloud Report 2019". Hozzáférés: 2019. június 9. <https://info.flexerasoftware.com/SLO-WP-State-of-the-Cloud-2019-DE>.
- GitHub (2019): "Az Octoverse 2018-as állapota." Hozzáférés június11,2019. <https://octoverse.github.com/>.
- (a): "A csillagokról." Hozzáférés: 2019. június 11. <https://help.github.com/en/articles/about-stars>.
- (b): "Fork a Repo." Hozzáférés: 2019. június 6., <https://help.github.com/en/articles/fork-a-repo>.
- Goodwin, Charles (2018): *Co-Operative Action. Learning in Doing: Social, Cognitive and Computational Perspectives*, New York: Cambridge University Press.
- IANS (2017): "Demokratizálni akarjuk a mesterséges intelligenciát: Google." *YourStory. Com.* Accessed June 9, 2019. <https://yourstory.com/2017/08/democratise-artificial-intelligence-google/>.
- Kelty, Christopher M. (2008): *Two Bits: The Cultural Significance of Free Software. Experimental Futures*, Durham: Duke University Press.
- Kollányi Bence (2016): "Honnan jönnek a botok? A GitHubon megosztott botkódok elemzése". In: *International Journal of Communication* 10/0, pp. 4932-4951.
- Mackenzie, Adrian (2015): "A jóslatok előállítására: What Does Machine Learning Want?" In: *European Journal of Cultural Studies* 18/4-5, pp. 429-45.

<https://doi.org/10.1177/1367549415577384>.

— (2017): *Machine Learners: Cambridge: Archaeology of a Data Practice*: MIT Press.

- (2017): "Infrastrukturák csak névlegesen? Identifying Effects of Depth and Scale". In: Harvey, Penny/Jensen, Casper Brunn/Morita, Atsuro (szerk.), *Infrastructures and Social Complexity: A Companion*, London: Routledge.
- (2018): "48 millió konfiguráció és egyre több: Platform Numbers and Their Capitalization." In: *Journal of Cultural Economy* 11/1, pp. 36-53. <https://doi.org/10.1080/17530350.2017.1393443>.
- Marino, Mark C (2014): In: "Field Report for Critical Code Studies": *Computational Culture: A Journal of Software Studies* 4. <http://computationalculture.net/field-report-for-critical-code-studies-2014%e2%80%a8/>.
- Microsoft (n.d.): "A mesterséges intelligencia demokratizálása". *Történetek* (blog). Hozzáférés 2018., szeptember. <https://news.microsoft.com/features/democratizing-ai/>.
- Moore, Todd (2018): "Think 2018: A mesterséges intelligencia demokratizálódása." *IBM Code* (blog). 2018. március 20. <https://developer.ibm.com/code/2018/03/20/think-2018-democratization-artificial-intelligence/>.
- Pichai, Sundar (2017): Google I/O'17 Keynote. <https://www.youtube.com/watch?v=Y2VF8tmLFHw>.
- (2018): Google I/O'18 Keynote. <https://www.youtube.com/watch?v=QzbpXC0oxL0>.
- Schmarzo, William (2017): "A mesterséges intelligencia, a mélytanulás és a gépi tanulás demokratizálása a Dell EMC Ready megoldásaival". In: *InFocus Blog | Dell EMC Services*. Hozzáférés június https://infocus.dellemc.com/william_schmarzo/democratizing-artificial-intelligence-deep-learning-machine-learning-with-dell-emc-ready-solutions/.
- scikit-learn (2019): "Scikit-Learn felhasználói kézikönyv (0.22.Dev0 kiadás)". Hozzáférés június 7., 2019. https://scikit-learn.org/dev/_downloads/scikit-learn-docs.pdf.
- Simonite, Tom (2017): "Apple Just Joined Tech's Great Race to Democratize AI". *WIRED*. <https://www.wired.com/2017/06/apple-siri-ai/>.
- Turing, Alan M. (1992): Turing Turing, Turing: "Intelligent Machinery": "Intelligent Machinery". In: *Mechanical Intelligence*, szerkesztette Darrel C. Ince. *Collected Works of A.M. Turing*, Amsterdam: *Collected Works of A.M. Turing*: North-Holland, pp. 107-27.
- Egység (n.d.): "Machine Learning." *Unity*. Hozzáférés: 2019. június 7. <https://unity3d.com/machine-learning>.
- Wulf, Josh (2017): "9 forrás a JavaScript kódolás megkezdéséhez". *OpenSource.Com*. Accessed June 7, 2019. <https://opensource.com/article/17/6/get-started-coding-javascript>.

A mesterséges intelligencia médiapolitikai dimenziójáról

A mélytanulás mint fekete doboz és az OpenAI¹

Andreas Sudmann

Sem a gépek, sem a programok nem fekete dobozok; ezek olyan műtárgyak, amelyeket megterveztek, mind a hardvert, mind a szoftvert, és mi kinyithatjuk őket, és belenézhetünk. (Allen Newell/ Herbert A. Simon [1997|1976], 82.)

Mit jelent a modern mesterséges intelligencia (AI) technológia médiapolitikai dimenziójának kritikus vizsgálata? Ahelyett, hogy konkrét, mesterséges intelligencia által vezérelt alkalmazások, például kép- vagy beszédfelismerő rendszerek politikai vonatkozásait vizsgálnánk, e dolgozat fő hangsúlyt magának a mesterséges intelligencia technológiai infrastruktúrájának politikai következményeire helyezi, különös tekintettel a gépi tanulás megközelítésére, amelyet azóta Deep Learningnek (röviden: DL, más néven neurális hálózatok szimulációja vagy mesterséges neurális hálózatok - ANN) 2006-neveznek. Ez az esszé először azt tárgyalja, hogy az ANN/DL-t alapvetően átláthatatlan vagy "fekete doboz" technológiaként kell-e felfogni, amely talán hozzáférhetetlen vagy csak részben hozzáférhető az emberi megértés számára. Másodsor, és az első kérdéssel kapcsolatban, a cél az, hogy kritikusan szemügyre vegyük az OpenAI nevű kutatócég napirendjét és tevékenységét, amely szándékosan támogatja a mesterséges intelligencia demokratizálását, és megpróbálja a DL-hez hasonló technológiákat hozzáférhetőbbé és átláthatóbbá tenni. Nyilvánvaló, hogy az ilyen idealista állításokat nem szabad egyszerűen természetesnek venni, különösen, ha figyelembe vesszük az OpenAI-hoz hasonló cégekbe befektetett nagy mennyiségű tőkét. Az olyan stratégiák, mint például a nyílt forráskihelyezésű AI, inkább azt a célt szolgálják, hogy demonstrálják e vállalatok technológiai potenciálját, hogy egymást felülmúlják, és/vagy hogy ritka tehetségeket vonjanak be. De talán még annál is fontosabb, minthogy egyszerűen megkérdőjelezzük az ilyen állítások hitelességét vagy ideológiai vonatkozásait, itt sokkal alapvetőbb problémákkal kell foglalkoznunk: Hogyan lehet hozzájárulni az átláthatósághoz és az ak-tivitáshoz?

¹ Ez egy olyan esszé kissé átdolgozott és kibővített változata, amely ugyanezen a címen jelent

meg a következő kiadványban: Ramón Reichert, Mathias Fuchs (szerk.), *Rethinking AI. Neural Networks, Biopolitics and the New Artificial Intelligence*, *Digital Culture & Society* 4/1, pp2018., 181-200.

egy olyan fekete doboz hozzáférhetősége, amelyet - talán - egyáltalán nem lehet kinyitni? És lehetséges-e a mesterséges intelligencia demokratizálása az adatok általános demokratizálása nélkül?

1. Az úgynevezett "AI forradalom"

Mielőtt ezekkel a kérdésekkel foglalkoznánk, fontos összefoglalni, hogyan sikerült a DL-nek a közelmúltban a mesterséges intelligencia domináns paradigmájává válnia. Egy e tekintetben jelentős esemény 2012-ben történt. Akkor a Torontói Egyetem három tudósa, Alex Krizhevsky, Ilya Sutskever és Geoffrey Hinton először képzett ki hatékonyan egy úgynevezett Convolutional Neural Networköt, amely a hagyományos mesterséges neurális hálózat (ANN) egy speciális, kép- és tárgyfelismerésre optimalizált változata, az immár híres ImageNet adatbázis alapján, valamint gyors, párhuzamosan szervezett GPU processzorokon (vö. Huang 2016). A GPU-k jelentős implementálása jelentette a különbséget: A torontói csapat több mint a felére tudta csökkenteni a korábbi képfelismerési megközelítések hibaarányát.² Bár ez a növekedés nem hangzik túl lenyűgözőnek, mégis elég nagy volt ahhoz, hogy felkeltsse az olyan vezető informatikai vállalatok, mint a Google és a Facebook figyelmét, amelyek gyorsan felvettek olyan vezető tudósokat, mint Yann LeCun és Geoffrey Hinton, és olyan AI startup cégeket is felvásároltak, mint a DNNResearch és a DeepMind. Ezeknek a vállalatoknak a DL-technológia iránti erős érdeklődése nem volt meglepő, mivel már eddig is nagy mennyiségű adatot tároltak, és most már hozzáférhettek egy olyan nagy teljesítményű technológiához, amely lehetővé teszi ezek intelligens feldolgozását és hasznosítását (vö. Reichert 2014). A DL-nek köszönhetően például lehetővé válik a felhasználók által a közösségi médiaplatformokra feltöltött képek automatikus címkézése, vagy a fogyasztói magatartás elemzése, hogy személyre szabott hirdetéseket generáljanak, vagy személyre szabott ajánlásokat tegyenek. Természetesen számos más alkalmazási terület van, ahol a DL/ANN technológiát jelenleg is használják: például az önvezető autók érzékszervi adatainak feldolgozására, a tőzsdei előrejelzésekhez szükséges adatok elemzésére, vagy optimalizált gépi fordításokra stb. Általánosságban elmondható, hogy a DL-algoritmusok univerzális eszközt jelentenek a mintafelismerési és előrejelzési feladatokhoz, hatékony eszközt az adattartalom bizonytalanságának és homályosságának kezelésére (Goodfellow/Bengio/Courville 2016).

Azonban eltartott egy darabig, amíg a modern gépi tanulási algoritmusok képesek voltak

hogy kibontakoztassák a bennük rejlő lehetőségeket. A DL néhány technológiai alapelemét már az 1940-es és 1950-es években kifejlesztették (vö. Sudmann 2016). Ennek a mesterséges intelligencia paradigmának az alap gondolata már akkor is az volt, hogy olyan számítógépes rendszert kell kifejleszteni, amely megfigyelés és tapasztalat útján képes tanulni, hogy konkrét problémákat oldjon meg vagy bizonyos tanulási feladatokat teljesítsen, anélkül, hogy a folyamatot konkrét

szabályok vagy elméletek irányítanak...

-
- 2 Mindössze hét év alatt az adathalmazban lévő objektumok osztályozásának pontossága 71,8%-ról 97,3%-ra nőtt. Nem utolsósorban a magas értéke miatt volt 2017 az utolsó éve ennek a híres versenynek. A részleteket lásd Gershgorn (2017) cikkében.

(Mitchell 1997). Ez az alapvető megközelítése minden létező gépi tanulási rendszernek, szemben a mesterséges intelligencia rendszerek úgynevezett szimbolikus, szabályalapú formáival, amelyek in- telligens viselkedését jellemzően többé-kevésbé kézzel kódolják előre (Boden 2014). Bár számos ML- megközelítés létezik, a közelmúltban kiderült, hogy a DL-módszerek a leghatékonyabbak, legalábbis a mesterséges intelligencia kutatás olyan kulcsfontosságú területein, mint a természetes nyelvfeldolgozás és a számítógépes látás.

Nagyon tágra szólva, a DL egyik legfontosabb jellemzője, hogy a technikák egy olyan osztálya, amelyet lazán inspirál a biológiai neurális hálózatok szerkezete és tanulási folyamatai (Alpaydin 2016). Más gépi tanulási feladatokhoz hasonlóan a DL-algoritmusok is több ezer, ha nem millió gyakorló adat elemzésével tanulnak több ezer vagy akár milliós iteráció alapján egészen addig a pillanatig, amíg a rendszer képes nem látott adatokat helyesen megjósolni. Ami azonban megkülönbözteti a DL-t más gépi tanulási megközelítésektől, az a tanulási folyamat *hierarchikus elosztása*. A DL technológia olyan hálózatokat szimulál, amelyek jellemzően több milliós mesterséges neuronból állnak, amelyek különböző rétegekre szerveződnek - egy bemenetre, egy kimenetre és egy flexibilis számú köztes rejtett rétegre (Trask et al. 2015). Ha egy hálózatot mélynek nevezünk, akkor legalább két vagy több köztes réteggel rendelkezik, amelyek a hálózaton keresztül feldolgozzák az információkat. A rétegek legalsó szintjén a hálózat a bemenet nagyon egyszerű formáit elemzi (például vonalak és élek, vizuális adatok esetén), és ezt az információt továbbítja a következő rétegszintre, amely bonyolultabb formákat dolgoz fel (például a tárgy olyan részeit, mint az arc vagy a lábak), és ismét továbbítja ezt az információt a következő legmagasabb szintre, egészen a végső réteggig, a kimeneti réteggig, amely aztán meg tudja jósolni, hogy egy bizonyos ismeretlen bemenet helyesen illeszkedik-e egy bizonyos kimenethez (a képen egy bizonyos tárgy látható vagy sem?).

2. A mélytanulás médiapolitikája

Az sem meglepő, hogy a jelenlegi mesterséges intelligencia boom hamarosan a bölcsészeti- és társadalomtudományok figyelmét is felkeltette, míg 2016 előtt a kemény tudományokon kívül számos tudományág többé-kevésbé figyelmen kívül hagyta a DL/ANN-hez hasonló gépi tanulási technológiákat. Természetesen hosszú hagyománya van a mesterséges intelligencia lehetőségeiről és korlátairól szóló interdiszciplináris és transzdiszciplináris vitáknak (vö. Weizenbaum 1976, Searle 1980), ám ezek a viták jellemzően nem foglalkoztak részletesen az ANN technológiájával. E tekintetben néhány fontos kivételt kell kiemelni, különösen az elmefilozófia, valamint a kognitív pszichológia diskurzusait, amelyek nagyon korán érdeklődést mutattak a mesterséges intelligencia szimbolikus és konnektivista formái iránt egyaránt (vö. Horgan/Tienson Haugeland 1996, 1997).

Sőt, még a médiatudományok területén is találhatunk néhány olyan esetet, amikor a tudósok

az ANN-technológia megvitatása.³ Ilyen például a *Computer als Medium* (1994) című antológia bevezetője, amelyet Friedrich Kittler, Georg Christoph Tholen és Norbert Bolz közösen szerkesztett. Érdekes módon a szöveg (amelyet csak Bolz írt) valahogyan megragadja a mesterséges intelligencia konnekcionista paradigmájának episztemikus relevanciáját, de mindezt anélkül teszi, hogy részletesen feltárná annak médiaelméleti vagy -történeti következményeit.

Az elmúlt körülbelül két évben (többé-kevésbé a DeepMind Alpha-Go sikere után) a helyzet jelentősen megváltozott. Egyre több könyv és cikk jelenik meg, például a humán tudományok területén, amelyek általában a mesterséges intelligenciával és különösen a DL-technológiával foglalkoznak (Sudmann Pasquinelli 2016, Finn 2017, McKenzie 2017, Engemann/Sudmann 2017, 2018). Pasquinelli (2017) például nemrég írt egy rövid esszét az ANN-ről történeti és filozófiai szempontból, és (Eco és Peirce hivatkozásaival) amellett érvel, hogy a technológia csak az induktív reaszoningot tudja kezelni, míg arra, amit Peirce abduktív érvelésnek nevez, képtelen. Továbbá vannak olyan szerzők, mint Nick Bostrom (2014), Ed Finn (2017) vagy Luciano Floridi (2017), akik már nagyon is részt vesznek a jelenlegi mesterséges intelligencia-technológiák politikai és etikai vitájában. Például Nick Bostrom *Superintelligence* című könyve: *Paths, Dangers, Strategies* (2014) című könyve nagy közfigyelmet keltett, részben azért, mert riogató tézise szerint a szupergépi intelligencia technológiai fejlődése az emberiség legnagyobb veszélye, amit később Elon Musk Twitter-bejegyzése is visszhangzott. Mégsem osztja mindenki, aki a mesterséges intelligencia politikai és etikai aspektusaival foglalkozik, ezeket az apokaliptikus nézeteket. Luciano Floridi például meg van győződve arról, hogy az emberiség képes kezelni az AI által vezérelt társadalmat, amennyiben "az emberek által felügyelt tervezési, ellenőrzési, átláthatósági és elszámoltathatósági rendszert" (2017: online) instanciáljuk.

Mégis, ami még mindig hiányzik a szellemi vitából, az a következő kérdések megvitatása

AI/DL egy határozottan médiapolitikai szempontból. De mit jelent ez a fókusz, és miért van rá egyáltalán szükségünk? Először is, természetesen sokféleképpen lehet gondolkodni a mesterséges intelligencia általános és különösen a DL médiapolitikai dimenziójáról. Az egyik lehetséges megközelítés például az lenne, ha azt állítanánk, hogy a "médiapolitika" mint elemzési napirend a politika médiával és/vagy a média és a politika történelmi kapcsolatával foglalkozik (vö. Dahlberg/Phelan 2011). Egy ilyen beszámoló alapján feltehetnénk például azt a kérdést, hogy a mesterséges intelligencia/DL technológia hogyan írja be magát a média és a politika viszonyaiba, vagy hogyan vesz részt a politika közvetítésében. Mindkét esetben feltételezhetjük, hogy a) a média/közvetítés és a politika alapvetően különböző fogalmak, és hogy b) a lehetséges elemzési perspektívákat nagymértékben alakítja és irányítja a fogalmakról alkotott alapértelmezésünk (beleértve azt is, hogy magát az AI-technológiát médiumként fogjuk fel).

3 Természetesen a médiatudományok szemszögéből nézve van még néhány publikáció, amelyek

például a mesterséges intelligencia technológiával foglalkoznak általában: Dotzler 2006.

Egy másik lehetséges megközelítés az lenne, ha azt állítanánk, hogy a médiának eredendően politikai dimenziója van (és hasonlóképpen azt is állíthatnánk, hogy semmi politikai nem létezik egy médiumon vagy bizonyos médiumokon kívül). Mégis kérdés marad, hogy ez igaz-e a média vagy médium *minden* fogalmára, vagy csak bizonyos fogalmakra. Ez azonban meglehetősen teoretikus vita, mivel a legtöbb médiapolitikai koncepció többé-kevésbé a média hagyományos értelmezésén alapul, amely a médiát tömeg- vagy populáris médiaként értelmezi (vö. Zaller 1999, Dahlberg/Phelan 2011).⁴

Jelen esszé kontextusában azonban a médiapolitikai beszámolót tágabban és bizonyos értelemben alapvetőbb módon értelmezem. Egy ilyen elméleti szemlélet, ahogyan én szeretném megfogalmazni, nem annyira a mesterséges intelligencia megjelenési formáival vagy látható interfészeivel foglalkozik, hanem sokkal inkább a technológiát létrehozó és alakító mediális infrastruktúra és entitások politikai következményei és hatásai érdeklik (függetlenül a konkrét "felhasználási esetektől" is).⁵ Más szóval, engem a modern AI-technológia *infra-mediális feltételei* és azok politikai dimenziója érdekel, amit én úgy hívok, hogy a modern AI-technológia *infra-mediális feltételei*.⁶ Számomra a mesterséges intelligencia-technológia létrehozásának és formálásának folyamatában részt vevő *minden* entitás általában mediátorként fogható fel (vö. Latour 2005). És általában véve a technológia minden közvetítője politikai szempontból is számít. Azonban nem minden közvetítő fogalmazható meg egyformán médiumként, legalábbis akkor nem, ha a fogalom szűkebb értelmezését alkalmazzuk, például a médiát olyan entitásoknak vagy diszpozitívoknak tekintjük, amelyek lehetővé teszik a kommunikációt, vagy amelyek információt tárolnak, feldolgoznak vagy továbbítanak.⁷ Éppen ezért általában érdemes különbséget tenni a közvetítő(k) és a médium/közeg fogalma között.

Bár amellet érvelek, hogy szükségünk van egy ilyen megkülönböztetésre, mégis eléggé szkeptikus vagyok a "médium" vagy "média" kifejezés stabil fogalmának használatával kapcsolatban (még akkor is, ha ez megkönnyítené a két fogalom megkülönböztetésének feladatát). Az én szememben a

4 Ennek megfelelően az egyik lehetséges megközelítés az lenne, ha megvizsgálnánk a különböző mesterséges intelligencia-technológiák reprezentációjának politikáját a népszerű médiában, például a filmben és a televízióban.

5 Hasonló, a "médiainfrastruktúrák" kifejezést kritikai fogalomként használó beszámolót találunk Parks/ Stasielski (2015) írásában.

6 Egy ilyen perspektíva nem foglalkozik közvetlenül egy konkrét elméleti kerettel. Általában ez a fókusz számos elemző megközelítéssel, például a médiaarcheológiával, a történeti episztemológiával vagy a szereplőhálózat-elmélettel összeegyeztethető. Az "infra" előtag kiemeli, hogy a médiát (működését) egy olyan szinten vizsgáljuk, amely általában láthatatlan (például a kód szintjén).

7 Ez egy másfajta beszámoló arról, hogyan lehet a médiát Latour "közvetítők" és "közvetítők" közötti különbségtételére hivatkozva konceptualizálni. Latour számára a "közvetítő [...]" az, ami transzformációk nélkül szállít jelentést vagy erőt", szemben a "közvetítő(k)kkel", amelyek "átalakítják, lefordítják, eltorzítják és módosítják a jelentést vagy azokat az elemeket, amelyeknek a

hordozásával szemben állnak" (2005: 39). A közvetítők bizonyos értelemben fekete dobozként működnek, mivel bemenetük lehetővé teszi a megfelelő kimenet előrejelzését (anélkül, hogy ismernénk az objektum belső műveleteit). Ezzel szemben a mediátorok esetében egy adott bemenet sajátosságai ellenére sosem könnyű megjósolni a vonatkozó kimenetet (ibid., vö. még Thielmann 2013).

empirikus világunk entitásainak (beleértve gondolataink immateriális világát is) értelmezése érdekében a média és/vagy médium kifejezések analitikus szempontból termékenyebbek, ha nem rögzített kategóriáknak, hanem flexibilis episztemológiai-heurisztikus kategóriáknak tekintjük őket.⁸ Ennek megfelelően a médiaelmélet, ahogy én képviselem, úgy értelmezhető, mint az az általános feladat, hogy feltárjuk, milyen különböző módokon érzékelhető a világ mint médium vagy mint médium (bizonyos jellemzőkkel, funkciókkal, feliratokkal), ahelyett, hogy egyszerűen elismerjük, hogy minden, ami a világban van, valamilyen ontológiailag stabil értelemben (mint az entítások előfeltétele, hogy láthatóak, érzékelhetőek, vagy hogy bizonyos formájuk van stb.) médiumtól vagy médiumtól függ. Ezért, bár a médiapolitika olyan koncepciója mellett döntök, amely a *közvetítők* (és a konkrét médiumok) *konstitutív szerepét* vizsgálja, mégis egy meglehetősen nyitott elemzési fókuszot képviselek, amely teret enged a legkülönbözőbb perspektíváknak.

Ez utóbbi álláspont a médiapolitika politikai dimenziója tekintetében is tanulságos megközelítésnek tűnik. Például könnyen állíthatjuk, hogy a mesterséges intelligenciával kapcsolatban szinte minden politikai, különösen, ha úgy véljük, hogy a mesterséges intelligencia/technológia társadalmi és kulturális létünk minden aspektusára hatással van. Ugyanakkor a politikai kihívások, amelyeket az AI és a DL technológia tartogat számunkra, nagyon különböző természetűek (az AI által vezérelt autonóm fegyverrendszerek egzisztenciális fenyegetése, az AI hatása a munka jövőjére stb.), ezért nem hivatkozhatunk egyszerűen egy minden egyes problémára alkalmas politikai elméleti főszámításra. Hogy egy másik példát hozzak: Az autonóm fegyverrendszerek tervezésekor nyilvánvalóan erős politikai érdek fűződik ahhoz, hogy a technológiák előállításának, valamint működésének közvetítői rejtve maradjanak. Másfelől, például annak a kérdésnek a tekintetében, hogy a mesterséges intelligencia hogyan befolyásolja a munka jövőjét, fontosabb lehet, hogy az ebben a folyamatban részt vevő összes közvetítő a lehető legátláthatóbb és legelérhetőbb legyen.

Ez a kérdés nem jelenti azt, hogy "bármit megtehetünk" a mesterséges intelligencia és a szellemi fogyatékoság politikájának kezelését illetően. Ehelyett amellet érvelek, hogy először is meg kell próbálnunk feltárni, hogy a kortárs mesterséges intelligencia technológiák hogyan jelennek meg politikai jelenségeként (mielőtt egy bizonyos politikai elméletet alkalmaznánk a mesterséges intelligenciára). Ez a fókusz számos releváns szempontot foglal magában, beleértve annak elemzését, hogy maguk az informatikusok hogyan konceptualizálják a mesterséges intelligencia technológiát politikai szubjektumként.

Ebben az összefüggésben azt is szem előtt kell tartani, hogy a gépi tanulás témaköre általában és az ANN/DL témaköre különösen a legtöbb humán- vagy társadalomtudományi területen dolgozó tudós számára még mindig ismeretlen terület, még akkor is, ha már tanultak mesterséges intelligenciát. Ez

alapvetően azt jelenti, hogy eltarthat egy ideig, amíg az olyan tudományágak, mint a média- vagy a kultúrakutatás, valóban képesek lesznek értékelni a releváns

8 A médiagondolkodás ilyen koncepciójának igazolására talán sokféle megközelítés létezik. Nyilvánvalóan ismét hivatkozhatunk Latour "közvetítő" kategóriájára. Hasonló elméleti hivatkozás ebben az összefüggésben Vogl "médiává válás" (2008) fogalma is.

a DL technológiáinak és infrastruktúráinak politikai és/vagy etikai vonatkozásai. Nyilvánvaló, hogy ez a probléma az AI/DL fekete dobozként való megvitatásának és az OpenAI-hoz hasonló projektek értékelésének is központi tényezője. A terület számos kutatója számára a médiatudományok egyik központi feladata, hogy a tudástranzlációs vagy transzformációs folyamatokra összpontosítson, és egyfajta metaperspektívából elemezze, hogyan használja, adaptálja és alakítja át egy másik diszciplína az egyik diszciplína tudását (vö. például Bergermann/Hanke 2017: 127). De hogyan nyújthat a médiatudomány releváns betekintést az AI/DL fekete dobozának problémájába, ha még az informatikusoknak vagy az adattudósoknak is mélyreható gondot okoz ezzel foglalkozni? Nyilvánvaló, hogy a médiatudományok technikai értelemben semmit vagy csak keveset tudnak hozzájárulni e fekete doboz kinyitásához, mégis talán rávilágíthatnak különböző aspektusokra: például a probléma szociokulturális feltételek, implikációk és hatások komplex hálózatának feltárására. Továbbá a médiatudományok - természetesen - kritikusán vizsgálhatják, hogy az adattudósok hogyan kezelik a mesterséges intelligenciát és annak fekete doboz problémáját politikai kérdésként. Ehhez azonban foglaljuk össze, mit jelent - vagy jobban mondva mit jelenthet -, hogy empirikus világunk bizonyos entitásait fekete dobozként érzékeljük.

3. Mélytanulás: Fekete doboz, amelyet nem lehet kinyitni?

A fekete doboz kifejezés pontos eredetéről viták folynak. Philipp von Hilgers feltárta a fogalom történetét, és a második világháború történetéig, pontosabban a magnetron technológiáig vezette vissza (Hilgers 2009). Azóta a fogalmat nagyon különböző kontextusokban, ellentétes jelentéssel alkalmazták és specifikálták. Egyrészt utalhat a repülőgépek vagy autók adatmegfigyelő rendszereire; másrészt olyan rendszereket foglal magában, amelyek belső működése átláthatatlan vagy hozzáférhetetlen, és így csak a bemeneteik és kimeneteik alapján figyelhetők meg (vö. Pasquale 2015: 3). A fekete doboz kifejezés egyik korai definícióját Norbert Wiener adta meg híres könyvének, a *Kibernetika* (1948) kiadásához 1961 fűzött előszó lábjegyzetében: "Fekete doboz alatt egy olyan készüléket értek [...], amely egy meghatározott műveletet hajt végre [...], de amelyről nem feltétlenül rendelkezünk információval arról a szerkezetről, amellyel ezt a műveletet végzi" (xi. o.).⁹ Végül, de nem utolsósorban, ahogy Latour kifejti, figyelembe kell vennünk, hogy a tudomány és a technológia műveletei mindig fekete dobozként hatnak: "Amikor egy gép hatékonyan működik, amikor egy ténykérdés el van intézve, csak a bemenetekre és a kimenetekre kell összpontosítanunk, nem pedig a belső komplexitására. Így paradox módon minél sikeresebb a tudomány és a technológia, annál átláthatatlanabbá és homályosabbá válik" (Latour 1999: 99). Első látásra ez is igaznak tűnik.

- 9 E definícióhoz lásd W. Ross Ashby 1956 *An Introduction to Cybernetics (Bevezetés a kibernetikába)* című könyvét.

a DL-rendszerek esetében. És mégis, a technológia más formáival ellentétben a DL-technológia esete másnak tűnik.

Jellemzően, függetlenül az imént említett fekete doboz-hatástól, a gyakorlatban használt technológia számos, ha nem a legtöbb művelete így vagy úgy, de hozzáférhető az emberi megértés és magyarázat számára. Ezzel szemben a DL-algoritmusok fekete doboznak tűnnek, amelyet nem lehet kinyitni. Legalábbis több szakértő jelenleg ezt tekinti a mesterséges intelligencia egyik legnagyobb problémájának. De valóban igaz-e, hogy a DL alapvetően átláthatatlan technológia, és ha igen, milyen mértékben? És még ha ez így is van, nem fogadhatjuk el egyszerűen, hogy az ANN átláthatatlan technológia, amíg zökkenőmentesen működik? Az utóbbi kérdés megválaszolása talán kevésbé tűnik bonyolultnak, mint az első. Valójában úgy tűnik, hogy már most is nagy konszenzus van számos tudós, politikus és vezető informatikai vállalat között a felelős vagy etikus mesterséges intelligencia kifejlesztésére, és a technológia elszámoltathatóbbá tétele ennek a törekvésnek az egyik lényeges része.

Ez a széles körű konszenzus természetesen nem meglepő. Az egy dolog, ha egy AI/DL rendszer kiábrándító filmajánlattal áll elő, de ha intelligenciagépeket használunk komolyabb, élet- és halálkérdéseket érintő ügyekben, a történet teljesen más. Ahogy Tommi Jaakkola, az MIT informatikusa nemrég rámutatott: "Legyen szó befektetési döntésről, orvosi döntésről vagy esetleg katonai döntésről, nem akarunk csak úgy egy "fekete dobozos" módszerre hagyatkozni" (Knight 2017).

Emiatt nem biztos, hogy elegendő tudni, hogy a mesterséges intelligencia/DL-rendszer előrejelzései kellően pontosak. Továbbá meg akarja érteni, hogy a rendszer miért jut egy bizonyos előrejelzésre. Mindkét szempont rendkívül fontosnak tűnik ahhoz, hogy biztosítsuk az AI által vezérelt döntésekbe vetett bizalmat. Az AI előrejelzési modelljeinek értelmét megragadni azonban meglehetősen nagy kihívásnak tűnik. Ez utóbbit illusztrálандó: A Mount Sinai Icahn School of Medicine kutatói nemrégiben kifejlesztettek egy "Deep Patient" nevű mesterséges intelligenciaprogramot. A rendszer elképesztően jól képes azonosítani a betegségek különböző formáit, sőt az olyan pszichiátriai rendellenességek korai jeleit is, mint a skizofrénia, de még mindig nem tudják, hogyan lehetséges ez. Természetesen a Deep Patient nagy segítség lehet az orvosok számára, de szükségük van arra, hogy a rendszer megindokolja jóslatait, hogy biztos támpontot kapjanak a pácienseik orvosi kezeléséhez. "Meg tudjuk építeni ezeket a modelleket" - magyarázza Joel Dudley, az Icahn School of Medicine orvosi biológiai informatikai igazgatója - "de nem tudjuk, hogyan működnek" (Knight 2017).

A következőkben azt vitatom meg, hogy ez a feltételezés, amely rendszeresen megjelenik a jelenlegi mesterséges intelligenciáról szóló diskurzusokban, mennyiben félrevezető és mennyiben szorul tisztázásra. Először is szem előtt kell tartani, hogy a jelenlegi DL-technológia mögött álló matematika meglehetősen

egyszerű (vö. Goodfellow/Bengio/Courville 2016). Végül soron ez egy mat- ter a statisztikáról szól, bár a statisztika egy fejlett formájáról. Ezt a szempontot fontos kiemelni, mivel megfigyelhető a DL misztifikálására irányuló általános tendencia, amely kontraproduktív, és amelyet meg kell fékezni. Másodszor, sok szakértő hangsúlyozza, hogy

Az ANN *valójában* egy hozzáférhető technológia, különösen, ha összehasonlítjuk a bi- ologikus neurális hálózatokkal. Roland Memisevic, a toronto-berlini székhelyű TwentyBN DL vállalat vezető kutatója például rámutat, hogy "a DL algoritmusok legalábbis sokkal könnyebben hozzáférhetőek, mint az emberi agy, ahol a neuronális aktivitás mintái, valamint a tanulás által végrehajtott átalakulások még ma is nagyon átláthatatlanok. Ezzel szemben, ha egy ANN-modellt nézünk, akkor mindent rögzíthetünk, megfigyelhetünk, mérhetünk, a legapróbb részletekig. Könnyen kideríthető például, hogy mely tulajdonságok eredményezték azt, hogy egy kutyát tévesen macskának bélyegeztek, mert bizonyos fülformák újra és újra bizonyos téves besorolásokhoz vezethetnek" (Memisevic 2018, *saját fordításom*). Amit azonban valóban nehéz megérteni, az a mesterséges neuronok kölcsönhatása, ahogy Memisevic is egyetért: "mivel ilyen nagyszámú, párhuzamosan működő neuronról van szó, az ember emergens jelenségekkel szembeül, amelyek során az egész többet foglal magában, mint részeinek összege" (ibid.).

Így, bár kétségtelenül igaz, hogy az informatikusoknak foglalkozniuk kell azzal, amit általában a DL értelmezhetőségi problémájának neveznek, ez nem olyan alapvető, mint ahogyan azt a jelenlegi diskurzusban gyakran leírják (vö. Knight 2017). És nem meglepő módon a számítástechnikai iparon belüli és kívüli informatikusok jelenleg nagyon is azzal vannak elfoglalva, hogy megpróbáljanak megbirkózni ezzel az értelmezhetőségi problémával. Valójában az újrakutatók már számos megközelítést dolgoztak ki a DL előrejelzési modelljeinek jobb megértésére és visszafejtésére.

4. A megmagyarázható mesterséges intelligencia stratégiái (XAI)

Az egyik példa arra, hogy nemcsak az ANN, hanem általában a gépi tanulási technológiák hozzáférhetőbbé váljanak, a Marco Tulio Riberio, Sameer Singh és Carlos Guestrin által kifejlesztett Local Interpretable Model-Agnostic Explanations (LIME) program. A szerzők leírása szerint ez "egy olyan technika, amellyel *bármilyen* gépi tanulási osztályozó előrejelzései megmagyarázhatók, és amelyek hasznosságát különböző, a bizalomhoz kapcsolódó feladatokban lehet értékelni" (Riberio/Singh/Guestrin 2016). A LIME alap gondolata az, hogy a mesterséges intelligencia rendszer számára a bemenetek (pl. szövegek vagy képek) különböző formáit úgy változtatjuk meg, hogy megfigyelhetjük, hogy a bemenet ezen variációi hatással vannak-e és hogyan hatnak a kimenetre. A *Science* című folyóiratban nemrégiben megjelent cikk a LIME gyakorlati működését egy filmkritikákkal táplált ANN-re hivatkozva magyarázza el:

[Egy neurális hálózat] beemeli a filmkritikák szavait, és megjelöli azokat, amelyek pozitívak. Ribeiro programja, a Local Interpretable Model-Agnostic Explanations (LIME), a pozitívnak jelölt kritikákat veszi, és a szavak de- letálásával vagy kicserélésével finom variációkat hoz létre. Ezeket a változatokat ezután lefuttatná

a fekete dobozon, hogy lássa, továbbra is pozitívnak tekinti-e őket. A több ezer

tesztek segítségével a LIME képes azonosítani a szavakat - vagy egy kép vagy molekuláris struktúra részeit, vagy bármilyen más adatot -, amelyek a legfontosabbak a mesterséges intelligencia eredeti ítéletében. A tesztek közül kiderülhet, hogy a "szörnyű" szó létfontosságú volt egy pánikláshoz, vagy hogy a "Daniel Day Le-wis" szó vezetett egy pozitív értékeléshez. (Voosen 2017)

Amint az már ebből a rövid leírásból is kiderül, túlzásnak tűnik azt állítani, hogy ez a modell valóban mélyreható magyarázatot ad. Alapvetően egy "kísérleti rendszerről" van szó, amely egyszerűen kiemeli azokat az elemeket, amelyek fontos szerepet játszanak a rendszer döntéshozatali folyamatában, anélkül, hogy ténylegesen feltárná a predikciós modellben rejlő érvelést. Ez utóbbihoz való hozzáférés nélkül azonban az XAI-modellek nehezen tudják teljesíteni a demokratikus mesterséges intelligencia nagy ígéreteit: teljes hozzáférhetőséget, átláthatóságot és ellenőrzést biztosítani.

Egy másik érdekes eszköz, amely - bizonyos értelemben - segített láthatóvá tenni az ANN-ek működését, a Google mérnökei és tudósai által 2015-ben bemutatott, ma már híres "DeepDream" nevű program. A DeepDream egy speciális DL-alapú képfelismerő algoritmus, mégis kicsiképp működik, mint egy tipikus CNN. Először is, az algoritmust több millió olyan képpel képzik ki, amelyeken egy adott tárgy (például egy macska) látható, így az NN egy bizonyos ponton képes megjósolni vagy osztályozni azokat a tárgyakat (például macskákat) olyan képeken, amelyekre nem volt betanítva. A kezdeti betanítás után a hálózat fordítva is működhet. Ahelyett, hogy a hálózat súlyait módosítanánk, mint ahogyan az a back prop algoritmusnál szokásos eljárás lenne, a súlyok változatlanok maradnak, és csak a bemenetet (a macska eredeti képét) módosítjuk minimálisan. Ez a technika nagyon érdekes eredményeket hoz, ha olyan képekre alkalmazzuk, amelyek nem tartalmazznak macskákat, de úgy vannak címkézve, mintha tartalmaznának. Ebben az esetben a szoftver elkezd módosítani és javítani a képek bizonyos mintáit, hogy azok egyre inkább úgy nézzenek ki, mint egy macska, de ne úgy, mint egy konkrét, a mi em-pirikus világunkban létező macska, hanem úgy, ahogyan egy neurális hálózat megtanulta érzékelni, ha nem is mondhatnánk: megálmódni. E folyamat eredményeként a rendszer szürreális és groteszk képeket produkál: például egy pizzáról készült fotó sok kis kutyaarcot vonhat maga után, vagy a Mona Lisát is átváltoztathatjuk LSD-szerű hallucinációs rémálommá.¹⁰ A generált képek legalább két érdekes szempontot tárnak fel: Egyrészt azt mutatják, hogy a DL nem teljesen titokzatos technológia, amennyiben az algoritmus feljavitja az ismerős vizuális jellemzőket. Másrészt a képek illusztrálják, hogy az algoritmus mennyire másképp működik az emberi észleléshez képest, más szóval előtérbe helyezve, hogy olyan aspektusaira is fókuszálhat egy képnek, amelyekre emberként általában nem figyelünk (vö. Knight 2017). De vajon a DeepDream is hozzájárul a mesterséges intelligencia demokratizálásához? Minden bizonnyal vannak

-
- 10 A DeepDream kritikus szemszögéből, médiaelméleti és pszichoanalitikai szempontból lásd Apprich (2017). A következő weboldalon a Deep- Dream algoritmus használatára láthatunk néhány példát: <https://deepdreamgenerator.com/#gallery>.

jó okunk van szkeptikusnak lenni, de legalább a DeepDream esztétikai megközelítést kínál a mesterséges neurális hálózatok működésének szemléltetésére, és ezáltal lehetővé teszi, hogy az emberek elgondolkodjanak például a DL-technológiák átláthatatlanságán. A DeepDream talán számos olyan művészt is megihletett, akik már jó ideje használják a DL-technikákat munkájukban. Az "Obvious" párizsi művészkollekció például kifejezetten kijelentette, hogy művészetével demokratizálni kívánja a mesterséges intelligenciát. A csoport nemrégiben azzal került a címlapokra, hogy a Christies aukciósháznál jelentős összegért eladtak egy portrét, amelyet egy úgynevezett GAN algoritmus segítségével készítettek.¹¹

Egy harmadik lehetséges megközelítés egy DL-rendszer működési mechanizmusainak feltárására az úgynevezett "Pointing and Justification (PJ-X)" modell, amelyet a Berkeley-i Kaliforniai Egyetemen és a Max Planck Informatikai Intézetben fejlesztettek ki (lásd Park et al. [2016]). A modell egy természetes nyelvi magyarázattal kombinált figyelemmechanizmus segítségével képes igazolni a predikciós vagy osztályozási feladatait azáltal, hogy kiemeli és dokumentálja az algoritmikus döntést alátámasztó bizonyítékokat. A rendszer kulcsfontosságú eleme, hogy két különböző adatkészlettel van betanítva. Az elsőknek az a célja, hogy meghatározza, mit mutat egy kép, míg a másodíknak az a funkciója, hogy feltárja, miért jelenik meg valami (pl. egy bizonyos emberi tevékenység vagy tárgy) egy adott képen. Az ötlet tehát az, hogy a tárgyakat vagy emberi tevékenységeket ábrázoló képeket ne csak a leírásukkal (címkézésükkel), hanem a hozzájuk tartozó magyarázattal is összefüggésbe hozzuk. Ez utóbbi célból a képzési adatok minden egyes képéhez három kérdés, valamint tíz válasz tartozik. Ennek alapján a rendszer olyan kérdésekre tud válaszolni, mint például: "Hazudik-e a képen látható személy f?". A válasz pedig lehet a következő: "Nem, mert a személy lába még mindig a földön áll" (vö. Gershgorn 2016). Ismétlem, ez a modell - mint a fentiek mindegyike - még messze van attól, hogy képes legyen megmagyarázni saját belső működését vagy más gép (vagy ha úgy tetszik, egy másik ANN) működését. Talán ehhez a sajátos képességhez az kellene, hogy a gépek valamiféle öntudatot, vagy akár meta-tudatot fejlesszenek ki. Mielőtt ez megtörténne (ha ez valaha is megtörténik), a DL technológiának meg kell értenie az érvelést, a tervezést, az oksági kapcsolatokat és így tovább. Jelenleg a DL vagy az ANN technológiája csak korrelációkat szolgáltat, de mélyreható ok-okozati magyarázatokat nem. Ez utóbbira azonban szükség lenne a mesterséges intelligencia demokratizálására vonatkozó állítások fényében, ha figyelembe vesszük például, hogy a törvények és más szabályozási és szervezési formák mindenképpen szükségessé teszik az oksági magyarázatokat. Amíg azonban a DL-modellek nem képesek megfelelően kezelni az ok-okozati problémákat, addig félrevezető vagy egyszerűen túlzás az XAI egy formájaként beszélni ezekről a megközelítésekről.

- I I Ugyanakkor az Obvious-t kritizálták a projektjükért, mert nem tettek mást, minthogy egy külföldi algoritmus, hogy ezt és más portréikat létrehozzák.

5. Az OpenAI politikája

Amint korábban jeleztem, a megmagyarázható és - általánosabban - felelős mesterséges intelligencia modelljeinek megalkotása nyilvánvaló motivációkkal bír. Mindenekelőtt azoknak, akik jelenleg DL-rendszereket fejlesztenek, erős gazdasági érdekük, hogy ellensúlyozzák a mélységesen átláthatatlan AI-technológiával kapcsolatos társadalmi félelmeket és szkepticizmust. Mindazonáltal számos tudós és ipari szereplő hangsúlyozza a megmagyarázható mesterséges intelligencia kifejlesztésének politikai és etikai jelentőségét, a fent leírt értelmezhetőségi problémához kapcsolódó kereskedelmi szempontokon túlmenően. Az egyik leglátványosabb és legerősebb szereplő az OpenAI, amely "nonprofit kutatócégként" (önleírás) indult,¹² és szintén a DL-technológiára specializálódott. A cég így vázolta fel küldetésének célját, nem sokkal a 2015 októberében történt megalapítása után:

Célunk, hogy a digitális intelligenciát olyan módon fejlesszük, amely a legnagyobb valószínűséggel az emberiség egészének javát szolgálja, anélkül, hogy a pénzügyi megtérülés szükségessége korlátozná. Mivel kutatásaink mentesülnek a pénzügyi kötelezettségektől, jobban tudunk a pozitív emberi hatásokra összpontosítani.

Úgy gondoljuk, hogy a mesterséges intelligenciának az egyéni emberi akarat kiterjesztésének kell lennie, és a szabadság szellemében a lehető legszélesebb körben és egyenletesen kell elterjednie. Ennek a kísérletnek a kimenetele bizonytalan, és a munka nehéz, de hiszünk abban, hogy a cél és a struktúra helyes. Reméljük, hogy ez az, ami a szakma legjobbjai számára a legfontosabb. ("Az OpenAI bemutatása")

Annak érdekében, hogy az OpenAI-t "ne korlátozza a pénzügyi visszafordítás szükségessége", a vállalat alapítói, köztük a legjelentősebbek, Elon Musk és Sam Altman, több mint 1 milliárd dollárt fektettek be ebbe a vállalkozásba. Érdekes módon ez a kezdeti induló bejegyzés nem utal kifejezetten vagy közvetlenül arra, amit Elon Musk az OpenAI-ba történő kezdeti befektetésének egyik fő motivációjának nevezett, nevezetesen, hogy az (általános) mesterséges intelligenciát az emberiség legnagyobb egzisztenciális fenyegetésének tartja.¹³ Ez az apokaliptikus nézet az AI-kutatás kezdete óta létezik, sőt, már korábban is létezett. Valójában, ahogy Bernhard Dotzler médiatudós már az 1980-as évek végén rámutatott, a mesterséges intelligencia jövőjére vonatkozó legmegalapozottabb előrejelzések már Alan Turing munkásságában megtalálhatók (vö. Dotzler 1989). És mégis, mivel nagyon

12 Érdemes megjegyezni, és egyben sokatmondó, hogy az OpenAI ma már egy "capped-profit" vállalat (vö. Coldewey 2019: online).

- 13 Februárban Musk 2018, bejelentette, hogy a Teslánál végzett (jövőbeli) munkájával kapcsolatos potenciális összeférhetlenség miatt kilép az OpenAI igazgatótanácsából (Vincent 2018).

a közelmúltban a mesterséges intelligencia fejlődése nem sok okot adott arra, hogy azt várjuk, hogy egy disztópikus "Terminátor" valóság van a sarkunkban.

Az emberiség történetében először a jelenlegi helyzet valóban más lehet, mint a jelenlegi DL-technológia tagadhatatlanul gyors fejlődésével szemben. Legalábbis Muskon túl sok szakértő ezt gondolja. Az OpenAI napirendje elismeri ezt az új helyzetet, de árnyaltabb, kevésbé drámai módon:

A mesterséges intelligencia rendszerek ma lenyűgöző, de szűkös képességekkel rendelkeznek. Úgy tűnik, hogy folyamatosan faragni fogunk a korlátaikon, és szélsőséges esetben gyakorlatilag minden szellemi feladatban elérik az emberi teljesítményt. Nehéz elképzelni, hogy az ember-szintű mesterséges intelligencia mennyire hasznos lehet a társadalom számára, és ugyanilyen nehéz elképzelni, hogy mennyire károsíthatja a társadalmat, ha rosszul építik vagy használják. ("Az OpenAI-ról")

Valójában senki sem képes előre látni a mesterséges intelligencia jövőjét, és senki sem tudja megítélni, hogy az inkább pozitív vagy negatív hatással lesz-e a társadalomra és a kultúrára. Azt is mondhatnánk magunknak, hogy a technológia sosem eredendően jó vagy rossz, ezért csak a konkrét felhasználás számít. Ez az érv azonban mindig is meglehetősen problematikus volt, hiszen valójában nem mindegy, hogy nukleáris technológiáról vagy mondjuk szélenergiáról van szó. Ráadásul, még ha eléggé közhelyes is, hogy a jövő bizonytalan, azt sem szabad elfelejtenünk, hogy soha nem lehetünk biztosak abban, hogy melyik konkrét történelmi ponton lépünk rossz útra a mesterséges intelligencia káros alkalmazásai felé. Különösen ez utóbbi érv tűnik egybeesőnek azzal, ahogyan az OpenAI összekapcsolja jelenlegi programját a beláthatatlan jövő problémájával:

A mesterséges intelligencia meglepő története miatt nehéz megjósolni, hogy mikor kerülhet emberi szintű mesterséges intelligencia a közelünkbe. Amikor ez bekövetkezik, fontos lesz, hogy legyen egy olyan vezető kutatóintézet, amely a saját érdekeivel szemben a mindenki számára jó eredményt helyezi előtérbe. ("Az OpenAI-ról")

Ami érdekes ebben a szövegben, az az a hallgatóságos feltételezés, hogy a mesterséges intelligencia drasztikus negatív vagy pozitív hatásaira vonatkozó kérdés még mindig meglehetősen spekulatív, és nem annyira a technológia jelenlegi állását érinti ("mikor kerülhet elérhető közelségbe az emberi szintű mesterséges intelligencia"). Bár az OpenAI-nak igaza van abban, hogy kerülni kell a spekulatív vitákat, fontosnak tűnik felismerni, hogy a *DL-nek már most is vannak* pozitív és problematikus következményei egyaránt. A technológia sok meglepően jó dolgot tehet, hiszen *már most is* egy nagyon erős és egyben veszélyes megfigyelési technológiává *vált*, amely nemcsak a világ (félautomatikus) megfigyelésének lehetőségeit bővíti (miután erre kiképezték),

hanem annak értelmezésére is képes.

Nemrégiben kiderült, hogy az ANN/DL nem csak tárgyak, emberek, tájak és állatok azonosítására képes (ismét csak a betanítás után), hanem elkezdte megérteni a meglehetősen összetett cselekvéseket és gesztusokat is. Más szavakkal: A DL rendszerek elkezdtek megérteni azt, amit a világról való common-érzéki tudás egy alapvető formájának nevezhetünk. Ennek érdekében az ANN-t nem fotókkal, hanem több százezer rövid videoklippel (amelyek bizonyos tevékenységeket és kézmozdulatokat mutatnak) képezték ki. Ezért a média sajátossága, azaz itt az állóképek és mozgóképek közötti különbség, mint a DL-technológia közvetítője, alapvető tényező a mesterséges intelligencia fejlett formáinak kifejlesztéséhez.

Ed Finn nemrégiben azt állította, hogy a mai algoritmikus kultúrát minden eddiginél jobban hajtja "a világ hatékony kiszámíthatóságának vágya" (2017: 26). Anélkül, hogy külön megkülönböztetné őket a tanulási algoritmusoktól, az algoritmusokat általában "kulturális gépeknek" (54) tekinti, amelyek működését nagyon is meghatározza az "egyetemes számítás ideológiája" (23). Valóban, úgy vélhetnénk, hogy különösen a modern DL-technológia az instrumentális ész egy fantasztikus változatát táplálja, éppen azért, mert újra felébreszti a régi, Leibniz-i álmodást egy olyan *mathesis universalis*ről, amely képes világunk minden aspektusát tökéletesen megérteni. De ennél is több, a DL nagy ígérete nemcsak az, hogy a gépek megértik a világot, hanem az is, hogy kiszámíthatóvá teszik azt egyre több szempontból: hogyan alakul a tőzsde, mit akarnak az emberek vásárolni, meghal-e valaki vagy sem, és így tovább. A DL-t már most, a történelemnek ebben a pillanatában is tekinthetjük úgy, mint azt a technológiát, amely képes olyan komplexitások elemzésére, amelyeket az emberek kognitívan nem képesek feldolgozni. A DL algoritmikus ereje abban rejlik, hogy a múltból tanulva képes mintákat azonosítani a jelen értékeléséhez, hogy uralni tudja a bizonytalan jövőt. És mindez egyre gyorsabb módon történik. A DeepMind például nemrég mutatta be Go-programjának új változatát, az "AlphaGo Zero"-t, amely mindössze három nap alatt képes volt megtanulni az ősi társasjátékot a semmiből (anélkül, hogy bármilyen szabályt bevezetett volna, hogyan működik a játék, vagy hogyan lehet sikeresen játszani), és a régebbi, 2015/16-os rendszerrel szemben (amely az emberi világbajnok Lee Sedolt verte meg) a győzelemre is képes volt (1000Perez 2017).

Az innováció gyorsasága a DL területén arra is emlékeztet bennünket, hogy óvatosnak kell lennünk azzal kapcsolatban, hogy gyorsan levonjuk a következtetéseket arról, hogy a mesterséges intelligencia technológia mit tud vagy mit nem tud elérni. Ezért nemcsak a mesterséges intelligencia távoli jövőjéről szóló spekulációkkal kell felhagynunk, hanem óvatosnak kell lennünk a szkeptikus nézeteinkkel is azzal kapcsolatban, hogy mire képesek (vagy nem képesek) a mesterséges intelligencia rendszerek. Általánosságban véve el kell ismernünk, hogy még sok munka vár ránk, ha megpróbálunk megbirkózni az AI és a gépi tanulási technológia jelenlegi fejlődésével. Talán az OpenAI-hoz hasonló vállalatoknak sikerül majd elérhetőbbé tenniük az AI-technológiát. De pontosan hogyan igazolják központi állításukat, miszerint demokratizálják az AI-

t? Ha még egyszer megnézzük a vállalat hivatalos weboldalát, rájövünk, hogy az nagyon kevés információt nyújt: "A vezető gépi tanulási konferenciákon publikálunk, nyílt forráskódú szoftvereszközökkel gyorsítjuk az AI-t.

kutatásokat, és blogbejegyzéseket teszünk közzé a kutatásaink kommunikálása érdekében" ("Az OpenAI-ról"). A cégnek lényegében ennyi a mondanivalója az AI demokratizálásának menetrendjéről, legalábbis ha csak a weboldal hivatalos küldetésnyilatkozatát vesszük figyelembe. Ami nagyon figyelemre méltó ebben a passzusban, az az, hogy semmi különös nincs benne. A Facebook, a Microsoft és sok más informatikai vállalat alapvetően ugyanezt az agenda-t (vö. Boyd 2017).

Persze lehet azzal érvelni, hogy az OpenAI legalábbis elindította a felelős és biztonságos mesterséges intelligencia fejlesztésének jelenlegi hullámát. A fontosabb kérdés azonban az: Az OpenAI lényegében azt csinálja, amit az AI-játék többi nagy szereplője - vagyis a meglévő technológia fejlesztését és/vagy a mesterséges általános intelligencia (AGI) kifejlesztéséhez vezető helyes út megtalálását -, és hogyan tudja legitimálni márkaképét és korábbi státuszát, mint nonprofit (és most már profitorientált) kutatócég? Ami ezt a kérdést illeti, és nagyon hasonlóan a DeepMindnél kialakult helyzethez, az OpenAI kutatásai a szimulációkkal (például játékokkal) kapcsolatos megerősített tanulási stratégiákra összpontosítanak ahelyett, hogy a felügyelt tanulás általános megközelítését alkalmaznák, amely az empirikus világból származó, helyesen címkézett adatoktól függ (vö. Rodriguez 2017). A megközelítésük sajátos korlátain belül mind az OpenAI, mind a DeepMind agenda megglehetősen sikeresnek bizonyult. Mégis, a szimulációk egyelőre még mindig nem alkalmasak az empirikus tanulási adatok helyettesítésére. Ha ez tartós problémának bizonyul, annak óriási következményei lesznek arra nézve, hogy miként képzeljük el a szimulációk episztemológiai státuszát (a digitális és vizuális kultúra számos elméletében és történetében), de ezt még meg kell várni. Az ok, amiért ezt a pontot kiemeltem, más: Mint az imént láttuk, a DL fekete doboz problémájának sokféle aspektusa van. Nem célo, hogy minden részletébe belemerüljek annak, hogy a vezető informatikai cégek jelenleg hogyan próbálnak nagy hatékonyságú AGI-rendszereket kifejleszteni. Ehelyett, amit megtanulhatunk, ha közelebbről megvizsgáljuk ezeket a különböző kutatási programokat, az az egyszerű tény, hogy a DL nem egy homogén megközelítés, hanem egy gyűjtőfogalom az AI kutatás és tervezés nagyon különböző megközelítéseisehez.

Továbbá a DL heterogenitására való hivatkozás nem csak az alábbi esetekben fontos a mesterséges intelligencia "fekete doboz" problémájának kezelése szempontjából, hanem az intelligens gépek kritikus szemléletének kialakítása szempontjából is. Hogy csak egy példát mondjak: Néhány évvel ezelőtt Alexander Galloway írt egy nagyon érdekes cikket, amelyben politizálta a fekete dobozt azzal az érveléssel, hogy az már nem olyan rejtjelező, mint a második világháború idején a magnetron technológia, hanem olyan funkcióvá vált, amelyet többé-kevésbé teljesen meghatároznak a bemenetei és kimenetei (vö. Galloway 2011: 273). A kifejezés használatával nem kizárólag technikai eszközökre gondol, hanem az emberek, tárgyak stb. minden olyan hálózatára és infrastruktúrájára utal, amelyek egymással kölcsönhatásba léphetnek, ám ezáltal csak külső funkcióikat artikulálják. Nyilvánvaló, hogy Galloway fekete doboz-fogalma bizonyos hasonlóságokat mutat azzal, ahogyan a fogalmat a szereplő-hálózat elméletben

használják, bár van egy fontos különbség: Galloway szerint a hálózat azon elemei, amelyek egy fekete dobozt alkotnak, már nem

képesek bármit is felfedni magukról. Más szóval: Úgy véli, hogy ezek a hálózatok fekete dobozzá váltak, amelyet *nem lehet* kinyitni (Hilgers is így definiálja a fekete dobozt - mint olyan rendszert, amelynek belső folyamatai folyamatosan hozzáférhetetlenek maradnak; vö. Hilgers 2009). Ezzel szemben például Michel Callon azt állította, hogy minden olyan fekete dobozt, amelynek szereplő-hálózati műveletei nem megfelelően modellezik egy rendszer működését, nemcsak fel lehet, hanem fel is kell törni, és ezáltal "új szereplők raját" kell létrehozni (Callon 1986). Első pillantásra úgy tűnik, hogy Galloway fekete doboz fogalma hasznos lehet a modern DL/ANN algoritmusok által közvetített infrastruktúrák és technológiai hálózatok leírására. Ez azonban nem olyan egyszerű, mint amilyennek elsőre tűnik. Galloway modelleje adott bemenetek és kimenetek meglétén alapul. Az ANN-technológia azonban nem mindig úgy működik, hogy mind a bemenetek, mind a kimenetek rendelkezésre állnak. Az úgynevezett felügyelet nélküli gépi tanulás esetében például az algoritmust adott kimenetek nélkül képzik ki. Ezért, ahogy ez az egyszerű példa is mutatja, ha egy DL/ML infrastruktúra árnyalatait fekete dobozként akarjuk megérteni, Galloway beavatkozása csak korlátozottan lehet hasznos. Ugyanakkor - és ez az a szempont, ahol a szereplő-hálózati elmélet ismét a képbe kerül - nem feltételezhetjük egyszerűen, hogy a fekete doboz problémája mint politikai (vagy etikai) kérdés csak magát az algoritmust érinti. Ehelyett a kérdés számos különböző közvetítőt és médiumot foglal magában: jogi szempontokat, intézményi eljárásokat, környezetvédelmi kérdéseket, meglévő politikai és jogi szabályozásokat stb.¹⁴

Ezeket a szempontokat akkor is fontos figyelembe venni, ha arról gondolkodunk, hogy a DL pro- a nagymamák faji vagy nemi előítéleteket mutatnak. Nagy volt a felzúdulás, amikor a Microsoft "Tay" nevű chatbotját a Twitter-felhasználók rasszista, szexista, valamint antiszemita kijelentések megtanulására képezték ki (Vincent 2016). Ez a botrány nagyon tanulságos volt, mivel megmutatta, hogy a tanulási algoritmusok műveletei közül valójában mennyi függ az adatoktól és - ami még fontosabb - az adatokat címkéző emberektől, legalábbis a felügyelt tanulási feladatok esetében. Más szavakkal: Nem vagy legalábbis nem annyira az algoritmusok termelnek előítéleteket vagy politikai problémás kimenetelt, hanem valójában az emberi szereplők, akik a tanulási adatokat tervezik és generálják, köztük az olyan platformokon keresztül felvett és szervezett több száz vagy ezer tömegmunkás, mint az Amazon Mechanical Turk vagy a CrowdFlower. Ha tehát a mesterséges intelligencia előítéletes problémájáról akarunk beszélni, akkor az előítéletek és ideológiák általános struktúráival is foglalkoznunk kell, amelyek még mindig meghatározzák társadalmunkat, és így a szakértőket és a mesterséges intelligencia rendszereket tervező munkásokat is. Továbbá ez a példa világosan mutatja

14 A közvetítők széles spektrumát érintő átláthatatlansági problémák megvilágításához hasznos

lehet Bur- rell (2016) fekete bokszeról szóló beszámolója, amelyet három szintre bont: "(1) az átláthatatlanságot mint szándékos vállalati vagy állami titkolózást, (2) az átláthatatlanságot mint technikai analfabétizmust, és (3) a gépi tanulási algoritmusok jellemzőiből és a hasznos alkalmazásukhoz szükséges léptékből adódó átláthatatlanságot."

miért fontos közelebből megvizsgálni, hogy a média bizonyos formái hogyan működnek a modern mesterséges intelligencia technológia kulcsfontosságú közvetítőiként.

Kétségtelenül eléggé rövidlátó, hogy az AI-ról, mint fekete dobozról szóló vita eddig szinte kizárólag a technológiai szempontokra összpontosított a narrátorok értelmében. Ez a "demokratikus mesterséges intelligencia" kritikáját is érinti. Nick Bostrom filozófus például nemrégiben megkérdőjelezte az AI átláthatóbbá tételének egész logikáját: "Ha van egy gombod, amely rossz dolgokat tehet a világgal, nem akarsz mindenkinek odaadni" (idézi: Metz 2016). Első látásra ez az érv meggyőzőnek tűnhet, ugyanakkor kissé furcsának tűnik. Ha például az atomfegyverekre gondolunk, könnyen megfigyelhetjük, mennyire bonyolult egy esetlegesen "veszélyes gombot" csak magunknak tartani. (Itt is utalhatunk a közelmúltbeli vitákra, amelyek az amerikai *etikajogáról* szóltak, hogy eldönthesse, használ-e atomfegyvert első csapásként vagy sem). Nem akarom azt állítani, hogy az elrettentés egyensúlyának koncepciója a hidegháború idején valóban békét biztosító hatású volt, és nem akarom a nukleáris fegyverek konkrét technológiáját a mesterséges intelligenciával egy szintre helyezni. Csak azt akarom illusztrálni, hogy miért lehet bonyolultabb az egész gyakorlat és diskurzus a re- szponzív vagy átlátható mesterséges intelligenciáról, mint azt Bostrom kijelentése sugallja. Az sem igaz, hogy az átlátható mesterséges intelligencia gondolatának egyetlen alternatívája az lenne, ha az összes releváns tudást titokban tartanánk a mesterséges intelligenciáról. Legalábbis ez utóbbi stratégia nem jöhet szóba az OpenAI számára, mivel ez a vállalat identitását rombolná le.

Továbbá fontos kiemelni, hogy a fekete dobozos próba ugyanúgy, mint a fekete dobozos prob-

A mesterséges intelligencia demokratizálására tett kísérletek nem csak magát a technológiát érintik, hanem azt is el kell ismernünk, hogy a mesterséges intelligencia demokratizálására tett kísérletek nem korlátozódhatnak pusztán a technológiát körülvevő eszközök és tudás nyílt forráskódúvá tételére (lásd a mesterséges intelligenciával mint fekete dobozzal kapcsolatos további kritikus nézeteket az átláthatóság és az elszámoltathatóság kérdésein túlmenően): Matzner 2017). Nem disztópikus álláspont, ha azt állítjuk, hogy már most is egy olyan posztpriváci korban élünk, amelyben az embereknek nagyon kevés kontrolljuk van a személyes életükkel és aktivitásaikkal kapcsolatos adatgyűjtési, tárolási, feldolgozási és továbbítási folyamatok felett. Edward Snowden leleplezései máris igaznak bizonyították a megfigyeléssel kapcsolatos legrosszabb konspirációs elméleteket (Sprenger 2015). A probléma itt nem csak az, hogy a vállalatok vagy a titkosszolgálatok, vagy általában a kormányok akaratunk ellenére gyűjtik és elemzik a magánjellegű adatokat. Túl gyakran egyszerűen sokan nem törődnek eléggé azokkal az adatokkal, amelyeket online jelenlétük vagy ennek vagy annak az alkalmazásnak a használata közben generálnak és terjesztenek. És még ha egyénileg meg is próbálják védeni magánjellegű adataikat, nincs garancia arra, hogy barátai, családtagjaik vagy kollégáik is így tesznek.¹⁵ Ezek a szempontok már jóval a jelenlegi mesterséges intelligencia boom előtt is kulturális kritikák

tárgyát képezték. Ezért nem szabad egyszerűen arról beszélnünk, hogy

15 Például, ahogyan az a Cambridge Analytica botrány során kiderült, a vállalat 2012-2015 között egy hátsó ajtót használt a Facebook API-ban, ahol a Facebook-barátok választása, hogy

a mesterséges intelligencia demokratizálása érdekében, hanem továbbra is arra törekszünk, hogy általában véve demokratikus normákat biztosítsunk az adatvezérelt világunk számára. E cél elérése érdekében a mesterséges intelligencia politikai elemzésének összekapcsolása az adatközpontúságról szóló szélesebb körű vitával csupán az első, de vitathatatlanul nagyon fontos lépés.

Jelenleg nehéz elképzelni olyan intézményt vagy törvényt, globális vagy helyi szinten, amely megvédhetne minket a mesterséges intelligencia veszélyeitől és a nagy adatokkal való visszaéléstől. És nincs semmi alapos okunk azt hinni, hogy kifejezetten az olyan cégek, mint az Open-AI, a Facebook vagy a Google el fogják érni ezt a célt. Ugyanakkor talán rövidlátó lenne ezeket a technológiai vállalatokat a demokratikus digitális kultúra ellenségeinek tekinteni, csak azért, mert ők azok a hegemon erők, amelyek mind az adatokat, mind az azok értelmezésére alkalmas intelligens algoritmusokat ellenőrzik. Nyilvánvaló, hogy vannak az AI-nak olyan veszélyei, amelyek sokkal sürgetőbbek, például ha nem demokratikus államok az AI-t vagy a DL-technológiát a politikai ellenzék elnyomására használják, vagy ha terroristák kibertámadásokra használják az AI-t. Ez a veszély nem a big data vagy az AI paranoia példája: Amint azt a szakértők nemrégiben demonstrálták, csupán egy úgynevezett API-hoz való hozzáféréssel közel 100%-os pontossággal képesek a machine learning algoritmusok visszafejtésére. A hackerek képesek ellopni az olyan informatikai vállalatoktól, mint az IBM vagy a Microsoft, az AI-technológiát bármilyen konkrét céljuk érdekében (a technikai részleteket lásd Claburn 2016). Természetesen egy valóban nyílt mesterséges intelligencia már eleve megoldhatja ezt a problémát. De akkor viszont fel kell tennünk a kérdést, hogyan tudjuk biztosítani, hogy egy nyílt mesterséges intelligenciát ne használjanak fel káros célokra.

Jelenleg úgy tűnik, hogy az OpenAI-t kevésbé érdekli a konkrét politikai elképzelés a

A mesterséges intelligencia és a mesterséges általános intelligencia kifejlesztéséért folytatott versenyben való részvételre. Ezért eléggé csábító azt hinni, hogy az OpenAI politikai vagy etikai programja alapvetően csak egy PR-trükk és semmi más. De ahelyett, hogy egyszerűen csak azt kérdőjeleznénk meg, hogy az OpenAI konkrét gyakorlatai megfelelnek-e a napirendjünknek vagy sem, talán termékenyebb lenne egy médiapolitikai beszámoló számára, ha az átlátható vagy felelős mesterséges intelligencia politikai következményeit és hatásait egy tágabb kontextusban vitatnánk meg: hogyan alakítja át az algoritmustanulás technológiája az instrumentális racionalitás feltételeit, amely oly mélyen kapcsolódik digitális kultúránk és társadalmunk minden aspektusához. És ez a fontos projekt még csak most kezdődött.

Köszönetnyilvánítás

Szeretnék köszönetet mondani Alexander Moneának, aki nagyon hasznos

visszajelzéseket és észrevételeket tett a szöveg eredeti változatával kapcsolatban. Szeretném továbbá köszönetemet kifejezni a bochumi Center of Advanced Internet Studies-nak a lehetőségért, hogy vezető ösztöndíjasként kutathattam e dolgozat témájában. 2017.

egy alkalmazás engedélyeibe, az adott alkalmazás az összes személyes adatát (beleértve a privát üzenetek tartalmát is) megadhatja.

Hivatkozások

- "Az OpenAI-ról." *OpenAI* Website. <https://openai.com/about/#mission> [Utolsó hozzáférés: 2018/03/06].
- Alpaydin, Ethem (2016): *Gépi tanulás. Az új mesterséges intelligencia*. Cambridge, MA: MIT P. Aprich, Clemens (2017): "Daten, Wahn, Sinn". *Zeitschrift für Medienwissenschaft* 17, 54-62.
- Ashby, W. Ross (1956): *Bevezetés a kibernetikába*. London: London: Chapman & Hall. Bergermann, Ulrike és Christine Hanke (2017): "Boundary Objects, Boundary Média. Von Grenzobjekten und Medien bei Susan Leigh Star und James R. Griesemer". In: *Grenzobjekte und Medienforschung*. Eds. Sebastian Gießmann és Nadine Taha. Bielefeld: transcript, 117-130.
- Burrell, Jenna (2016): "Hogyan "gondolkodik" a gép: Understanding Opacity in Machine Learning Algorithms (A gépi tanulási algoritmusok átláthatatlanságának megértése)". *Big Data & Society*, 3/1., pp. 1-12. <https://doi.org/10.1177/2053951715622512>.
- Bolz, Norbert (1994): Bolz: "Computer als Medium - Einleitung." In: *Computer als Medium*. Szerk. Norbert Bolz, Friedrich Kittler és Christoph Tholen. München: Fink, 9-16.
- Boden, Margaret A. (2014): "GOFAI." In: *The Cambridge Handbook of Artificial Intelligence*. Eds. Keith Frankish és William M. Ramsey. Cambridge, UK: Cambridge UP, 89-107.
- Bostrom, Nick (2014): *Superintelligence. Utak, veszélyek, stratégiák*. Oxford: Oxford UP.
- Boyd, Eric (2017): "Microsoft és a Facebook nyílt ökoszisztémát hoz létre a mesterséges intelligenciamodellek interoperabilitására." *Microsoft.com*. Szeptember Online7.: <https://www.microsoft.com/en-us/cognitive-toolkit/blog/2017/09/microsoft-facebook-create-open-ecosystem-ai-model-interoperability/> [Utolsó elérés: 2018/03/06].
- Callon, Michel (1986): Callon: "Egy szereplő-hálózat szociológiája: Az elektromos jármű esete." In: *Mapping the Dynamics of Science and Technology: Sociology of Science in the Real World*. Eds. Michel Callon, John Law, John és Arie Rip. Sheridan House Inc. 29-30.
- Claburn, Thomas (2016): "How to steal the mind of an AI: Machine-learning models vulnerable to reverse engineering". *The Register*. Online: https://www.theregister.co.uk/2016/10/01/steal_this_brain/ [Utolsó elérés: 2018/03/06].
- Coldewey, Devin (2019): "Az OpenAI nonprofitról "capped-profit"-ra vált, hogy tőkét vonzzon". *Techcrunch.com*. (<https://techcrunch.com/2019/03/11/openai-shifts-from-nonprofit-to-capped-profit-to-attract-capital/>).
- Dahlberg, Lincoln és Sean Phelan, szerk. (2011): *Diskurzuselmélet és kritikus média Politika*. Basingstoke: Palgrave Macmillan.

- Dotzler, Bernhard (1989): Dotzler, Bernhard: "Tudás/élmény: Versuch über die Verortung der Künstlichen Intelligenz" *MaschinenMenschen. Katalog zur Ausstellung des Neuen Berliner Kunstvereins*, 17.-23.07. Berlin: NBK. 127-132.
- (2006): *Diskurs und Medium. Zur Archäologie der Computerkultur*. Bd. 1. München: Fink.
- Engemann, Christoph és Andreas Sudmann, szerk. (2018): *Gépi tanulás. Médien, Infrastrukturen und Technologien der Künstlichen Intelligenz*. Bielefeld: transcript.
- Finn, Ed (2017): *What Algorithms Want. Képzlet a számítástechnika korában*. Cambridge, MA. MIT P.
- Floridi, Luciano (2017): "Az algoritmusok térhódítása nem feltétlenül rossz hír az embereknek". *Financial Times*, május Online4.: <https://www.ft.com/content/ac9e10ce-30b2-11e7-9555-23ef563ecf9a> [Utolsó hozzáférés: 2018/03/06].
- Galloway, Alexander R. (2004): *Jegyzőkönyv. Hogyan létezik az ellenőrzés a decentralizáció után*. Cambridge, MA: MIT P.
- (2011): "Black Box, Schwarzer Block. *Die technologische Bedingung*. Ed. Hörl, Erich. Frankfurt/M.: Suhrkamp, 267-280.
- Gershgorn, Dave (2016): "Nem értjük, hogyan hozza meg az AI a legtöbb döntést, ezért most az algoritmusok önmagukat magyarázzák." *Quartz*. December 20. Online: <https://qz.com/865357/we-dont-understand-how-ai-make-most-decisions-so-now-algorithms-are-explaining-themselves/> [Utolsó elérés: 2018/03/06].
- (2017): "Az adatok, amelyek átalakították a mesterséges intelligencia kutatását - és valószínűleg a világot". *Quartz*. Július 26. Online: http://www.notey.com/@qz_unofficial/external/17246232/the-data-that-transformed-ai-research%E2%80%9494and-possibly-the-world.html [Utolsó hozzáférés: 2018/03/06].
- Goodfellow, Ian, Yoshua Bengio és Aaron Courville (2016): *Deep Learning*. Cambridge, MA; London: MIT P.
- Hilgers, Philip von (2009): "Ursprünge der Black Box." *Rekursionen. Von Faltungen des Wissens*. Eds. Ana Ofak és Philipp von Hilgers. München: Fink, 281-324.
- Horgan, Terence und John Tienson (1996): Tienken: *Connectionism and the Philosophy of Psychology*. Cambridge, MA: MIT Press.
- Huang, Jensen (2016): "A mesterséges intelligencia gyorsítása GPU-kkal: A New Computing Model." *Nvidia*. Online: <https://blogs.nvidia.com/blog/2016/01/12/accelerating-ai-artificial-intelligence-gpus/>.
- "Az OpenAI bemutatása." *OpenAI Blog*. Online: <https://blog.openai.com/introducing-openai/> [Utolsó hozzáférés: 2018/03/06].
- Knight, Will (2017): "The Dark Secret at the Heart of AI". *MIT Technology Review*. 4. April 2017. <https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/> [Utolsó elérés: 2018/03/06].

Latour, Bruno (1999): Latour: *Pandora reménye. Esszék a tudománytudományok valóságáról*. Cambridge, MA: Harvard UP.

- (2005): *Reassembling the Social. Bevezetés az aktor-hálózat-elméletbe*. Oxford: Oxford UP.
- Matzner, Tobias (2017). "A fekete dobozok megnyitása nem elég - adatalapú felügyelet a fegyelmzésben és büntetésben és napjainkban". *Foucault Studies* 23: 27-45.
- Memisevic, Roland (2018): Mevemisic: "Wunderwerke der Parallelisierung." In: Sudmann/En- gemann, a.a.O. , o.S. (Kiadás előtti változat).
- Metz, Cade (2016): "Elon Musk vad terve a mesterséges intelligencia felszabadítására". *Wired*. April Online28.: <https://www.wired.com/2016/04/openai-elon-musk-sam-altman-plan-to-set-artificial-intelligence-free/> [Utolsó hozzáférés: 2018/03/06].
- Mitchell, Thomas (1997): *Gépi tanulás*. New York: McGraw-Hill.
- Newell, Allan és Herbert A. Simon (1997 [1976]): "Computer Science as Empirical Inquiry. Symbols and Search." *Mind Design II: Filozófia, pszichológia, mesterséges intelligencia*. Ed. John Haugeland, Cambridge, MA: MIT P, 81-110.
- Nott, George (2017b): "A Google kutatási vezetője megkérdőjelezi a "megmagyarázható mesterséges intelligencia" értékét" *Computerworld*. 23. Juni 2017. Online: <https://www.computerworld.com.au/article/621059/google-research-chief-questions-value-explainable-ai/> [Last access: 2018/06/03].
- Park, Dong Huk et al. (2016): "Figyelmes magyarázatok: Justifying Decisions and Pointing to the Evidence." (Döntések igazolása és a bizonyítékokra való rámutatás). 14. December. Online: <https://16/arxiv.org/pdf/12.04757v1.pdf> [Utolsó hozzáférés: 2018/03/06].
- Parks, Lisa és Nicole Starosielski, szerk. (2015): *Signal Traffic. Critical Studies of Media Infrastructures*. Chicago: U of Illinois P.
- Pasquale, Frank (2015): *A Black Box Society. A pénzt és az információt irányító titkos algoritmusok*. Cambridge, MA: Harvard UP.
- Pasquinelli, Matteo (2017): "Logikát morfondírozó gépek: Neural Networks and the Distorted Automation of Intelligence as Statistical Inference," *Glass Bead journal*, Site "1, Logic Gate: The Politics of the Artifactual Mind".
- Perez, Carlos E (2017): "Why AlphaGo Zero is a Quantum Leap Forward in Deep Learning". *Medium.com*. Online: <https://medium.com/intuitionmachine/the-strange-loop-in-alphago-zeros-self-play-6e3274fcdd9f> [Utolsó hozzáférés: 2018/03/06].
- Reichert, Ramón, szerk. (2014): *Big Data. Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie*, Bielefeld: transcript.
- Ribeiro, Marco Tulio, Sameer Singh és Carlos Guestrin (2016): "Introduction to Local Interpretable Model-Agnostic Explanations(LIME)A techniquetoexplain the predictions of any machine learning classifier". *O'Reilly*. Augusztus 12. On-line: <https://www.oreilly.com/learning/introduction-to-local-interpretable-model-agnostic-explanations-lime> [Utolsó elérés: 2018/03/06].
- Rodriguez, Jesus (2017): "Technology Fridays: OpenAI Gym Makes Reinforcement Learning Real." *Medium.com*. Online: <https://medium.com/@jrod>

- thoughts/technology-fridays-openai-gym-makes-reinforcement-learning-real-bcf762c16774 [Utolsó hozzáférés: 2018/03/06].
- Searle, John. R. (1980): Searle: "Minds, brains, and programs." (Elmék, agyak és programok). *Behavioral and Brain Sciences* 3 (3): 417-457.
- Sprenger, Florian (2015): *A mikrodöntések politikája: Edward Snowden, Net Neutrality, and the Architectures of the Internet*. Lüneburg: Meson P.
- Sudmann, Andreas (2016): "Wenn die Maschinen mit der Sprache spielen." *Frankfurter Allgemeine Zeitung* Nr. 256,2.11., N2.
- Thielmann, Tristian (2013): "Jedes Medium braucht ein Modicum". *ZMK Zeitschrift für Medien- und Kulturforschung* 4/2: "ANT und die Medien", 111-127.
- Trask Andrew, David Gilmore, Matthew Russell (2015): "Modeling order in neural word embeddings at scale". *Proceedings of the 32nd International Conference on Machine Learning (ICML-15)*; Lille, Franciaország. Július 6-11.
- Vincent, James (2016): "A Twitter kevesebb mint egy nap alatt megtanította a Microsoft AI chatbotját rasszista seggfejnek lenni." *The Verge*. Március 24. <https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist> [Utolsó elérés: 2018/03/06].
- (2018): "Elon Musk kilép a mesterséges intelligencia biztonsági csoport vezetéséből, hogy elkerülje az összeférhetlenséget a Teslával." *The Verge*. Febr. 21. Online: <https://www.theverge.com/2018/2/21/17036214/elon-musk-openai-ai-safety-leaves-board> [Utolsó elérés: 2018/03/06].
- Vogl, Joseph (2008): Vogl: "Becoming-media: Galilei távcsöve." *Grey Room* 29 (Winter): 14-25.
- Voosen, Paul (2017): "Hogyan törik fel az AI detektívek a mélytanulás fekete dobozát". *Science Mag*. Július 6. Online: <http://www.sciencemag.org/news/2017/07/how-ai-detectives-are-cracking-open-black-box-deep-learning> [Utolsó hozzáférés: 2018/03/06].
- Weizenbaum, Joseph (1976): Weizenbaum: *Computer Power and Human Reason. Az ítélőképességtől a Számítás*. New York: W. H. Freeman.
- Wiener, Norbert (1961 [1948]): Cambridge, Massachusetts: MIT P.
- Zaller, John (1999): *A médiapolitika elmélete. Hogyan érvényesülnek a politikusok, a sajtó és a média érdekei? listák, és a Polgárok alakítják a híreket*. Chicago, IL: Chicago UP.

Hogyan védjük meg a mesterséges intelligenciát

Ina Schieferdecker/Jürgen Großmann/Martin A. Schneider

1. Bevezetés

A mesterséges intelligencia az informatikán belül egy olyan tudományág, amely olyan szoftveralapú rendszerek fejlesztésével foglalkozik, amelyek olyan funkciókat látnak el, amelyek a jellemzően (emberi) intelligenciának nevezett funkciók végrehajtását igénylik. Mivel azonban az emberi intelligenciának nincs széles körben elfogadott definíciója, a mesterséges intelligenciának, amelyet néha gépi intelligenciának is neveznek, sincs széles körben elfogadott definíciója (Legg, 2007). A mesterséges intelligencia a logika, a valószínűségelmélet és a folytonos matematika módszereit és eszközeit használja annak érdekében, hogy szoftveralapú rendszereken keresztül érzékelést, érvelést, tanulást és cselekvést biztosítson (Russell, 2016). És már most számos gyakorlati alkalmazást biztosít a közlekedésben, az energiaellátásban, az egészségügyi szolgáltatásokban, a pénzügyekben és a bankszektorban, valamint a jogban és a szabályozásban: "A mesterséges intelligencia-technológiák már most is áthatják az életünket. Ahogy a társadalom központi erejévé válnak, a terület az egyszerűen intelligens rendszerek építésétől az emberre figyelő és megbízható intelligens rendszerek építése felé mozdul el". (Stone, 2016)

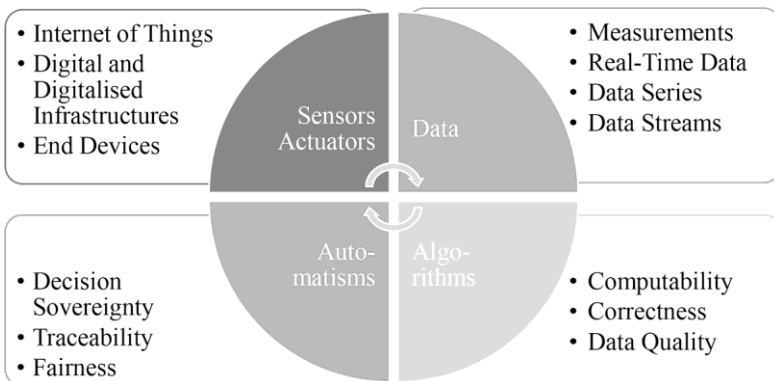
Sr Speech Recognition	Si Speech Identification							
Ar Audio Recognition	Ai Audio Identification	Pi Predictive Inference	Pl Planning					
Fr Face Recognition	Fi Face Identification	Ei Explanatory Inference	Ps Problem Solving		Lr Relationship Learning			
Ir Image Recognition	Ii Image Identification	Sy Synthetic Reasoning	Dm Decision Making	Lg Language Generation	Lc Category Learning	Ml Mobility Large		Cm Communication
Gr General Recognition	Gi General Identification	Da Data Analytics	Te Text Extraction	Lu Language Understanding	Lt Knowledge Refinement	Ms Mobility Small	Ma Manipulation	Cn Control

1. ábra: A mesterséges intelligencia funkcionális összetevői Hammond (2016) szerint: Beszéd (Sr), hang (Ar), arc (Fr) és kép (Ir) felismerése és általános felismerés (Gr), beszéd (Si), hang (Ai), arc (Fi) és kép (Ii) azonosítása és általános azonosítás (Gi); adatelemzés (Da) és szövegkiemelés (Te); prediktív következtetés (Pi), tervezés (Pl), magyarázó következtetés (Ei), problémamegoldás (Ps), szintetikus következtetés (Sr) és

döntéshozatal.

(Dm); Nyelvi generálás (Lg) és megértés (Lu); Kapcsolattanulás (Rl), Kategóriatanulás (Cl) és Tudásfinomítás (Kr); Mozgás nagyban (Ml) és kicsiben (Ms); Manipuláció (Ma), Kommunikáció (Cm) és Irányítás (Cn), amelyek önállóan vagy kombinálva is használhatók, pl. jövőbeli események előrejelzésére a technikai rendszerek hangjainak felismerésével és/vagy a rendszer állapotát ábrázoló képek azonosításával és/vagy adatok korrelálásával és konkrét tények felismerésével.

Számos olyan technológia létezik, amelyet a mesterséges intelligencia gépi tanúlással (röviden ML) történő létrehozására használnak, amely a problémamegoldás pontosságának vagy hatékonyságának javítását jelenti azáltal, hogy megtanulunk valamit jobban csinálni. A gépi tanulás például a tanulás típusa szerint csoportosítható felügyelt, felügyelet nélküli vagy félig felügyelt tanulási módszerekre, illetve a tudás szimbolikus számítással vagy szubszimbolikus feldolgozással történő kinyerése szerint. Csoportosíthatók a fő megközelítés mentén is, például regression, instance-alapú, regularizációs, döntési fa, Bayes-elmélet, klaszterezés, neurális hálózat, mélytanulás és még számos más algoritmus szerint. Ezek alapján hasonlóan számos mesterséges intelligencia-alkalmazás fejleszthető. Hammond (2016) bemutatta az AI funkcionális összetevőinek első taxonómiáját (1. ábra). Függetlenül attól, hogy milyen funkcionális komponenseket használnak, a mesterséges intelligencia alapú rendszerek szoftverek alkalmazásával vagy a környezettel való összekapcsolódást biztosító szenzorok és aktuátorok használatával is megvalósulnak (2. ábra). A szoftverek algoritmusok által értelmezett adatokat használnak fel, hogy automatizmusokat biztosítsanak a műszaki rendszerek egyes részei vagy teljes folyamatai számára, mint például az autómotorok vezérlésében, vagy a társadalmi-technikai rendszerek számára, mint például az autonóm vezetés esetében.



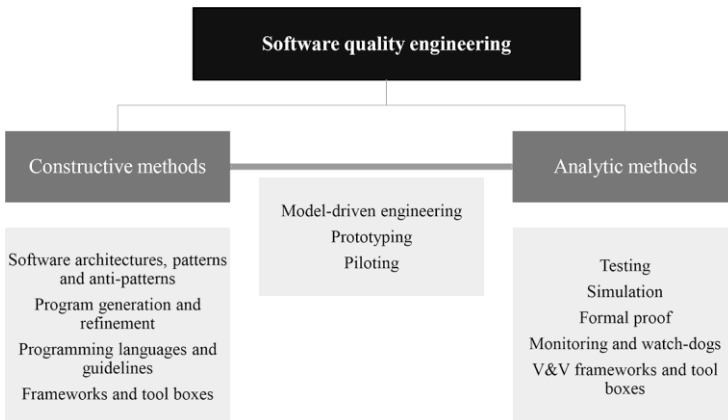
2. ábra: A szoftveralapú rendszerek elemei (WBGU, 2019). Az érzékelők a tárgyak internetének részét képezik, és különböző típusú adatokat, például méréseket, méréssorozatokat vagy adatfolyamokat generálnak. Az algoritmusok ezeket az adatokat használják fel számításaikban vagy képzési adatként. Az algoritmusokat a

*komplexitás, a kiszámíthatóság és a
a teljesítménykorlátok és esetleg a megvalósított számítási logika (nem) helyessége és a
(nem) elfogult (képzési) adatok. Ennek eredményeképpen a szoftware-alapú rendszerek a
fer*

automatizmusok, amelyek esetében elengedhetetlen a döntési szuverenitás, a nyomon követhetőség és a méltányosság egyeztetése (és biztosítása). A környezettel kapcsolatos minden döntés végül a software-en keresztül (a kibertérbe) és a működtetőkön keresztül (a környezetbe) táplálható.

2. Szoftverellenőrzés és -validálás

Mivel minden mesterséges intelligencia egyben szoftveralapú rendszer is, meg kell vizsgálni, hogy a mesterséges intelligencia milyen mértékben ellenőrizhető és validálható a szoftverekre általánosságban alkalmazott, bevett ellenőrzési és validálási (röviden V&V) módszerekkel. A szoftverekre vonatkozó V&V-módszerek már 1968-ban, a szoftverválság idején megjelentek (Wirth, 2008), amikor a soft-ware engineering kifejezést megalkották. Ez arra mutatott rá, hogy az adott erőforrásokkal és az adott időn belül milyen nehéz hasznos és megbízható szoftvereket tervezni és fejleszteni: "A szoftverválság fő oka, hogy a gépek nagyságrendekkel erősebbek lettek!...(A)ameddig nem voltak gépek, a programozás egyáltalán nem jelentett problémát; amikor néhány gyenge számítógépünk volt, a programozás enyhe problémává vált, és most, hogy gigantikus számítógépek vannak, a programozás ugyanolyan gigantikus problémává vált." A programozás nem jelentett problémát. (Dijkstra, 1972). Az újonnan kitalált kifejezés pedig arra mutatott rá, hogy a szoftverfejlesztéshez gyakorlatias és skálázható mérnöki módszereket kell kidolgozni. Azóta konstruktív és analitikus módszereket dolgoztak ki a szoftverminőség-fejlesztésre. Ezek közé tartoznak a szoftverfejlesztési folyamatokra, a szoftverfejlesztési eszközökre és a szoftverre mint olyanra vonatkozó módszerek. E módszerek durva áttekintését a 3. ábra mutatja be.



3. ábra: Áttekintés a szoftverminőség-fejlesztési módszerekről. A szoftverminőség a szoftvertervezéssel kezdődik, amelyet a szoftverarchitektúrák képviselnek, amelyek felhasználhatják a szoftvermintákat. A programok (részben) ezekből a szoftvertervekből

generálhatók és/vagy finomíthatók. A programok jellemzően magas szintű programozási nyelveket használnak

melyek a legjobb programozási gyakorlatra vonatkozó fer irányelvek, és melyek a programozási keretrendszerek és eszközök által támogatottak. Az elért szoftverminőséget jellemzően tesztelik, szimulációval ellenőrzik vagy formálisan bizonyítják. A futó szoftware monitorozható, és a figyelőprogramok futás közben ellenőrizhetik a korlátozások megsértését. Mindezek az elemzési módszerek V&V keretrendszerekkel és eszközökkel automatizálhatók. Három konkrét módszer(csoport) használható konstruktívan és analitikusan is: a modell használata a sof tware tervezésben, a sof tware (vagy a V&V sof tware) korai prototípezálása és a sof tware (vagy a V&V megoldások) tesztelése.

A szoftver (program vagy kód) megmondja a számítógépnek, hogy mit tegyen, *"de ez lehet, hogy sokkal másképp lesz, mint amit Ön elképzelt"*. (Joseph Weizenbaum, informatikus, 1923-2008). A szoftverminőség-technikai módszerek szisztematikus alkalmazásával azonban a szoftvereket úgy lehet fejleszteni, hogy azok biztonságosak, megbízhatóak és megbízhatóak legyenek, valamint hogy a szoftver több adatot tudjon elemezni és kiszámítani, mint bármely ember, és mindezt megbízhatóbban tegye.

A szoftverminőséget számos nemzetközi szoftverfejlesztési szabvány alapozza meg, mint például az ISO/IEC 25010 (ISO, 2011) a szoftverminőségi követelményekre és értékelésre (SQuaRE) és a szoftverminőségi modellekre vonatkozó szabványok. Ez a szoftverek használati minőségéről, külső és belső minőségéről érvel, és különbséget tesz a funkcionális alkalmasság, megbízhatóság, használhatóság, biztonság, kompatibilitás, portál- bilitás, karbantarthatóság és teljesítmény/hatékonyság között.

Bár ezek mind fontos szoftverminőségi szempontok, amelyek évtizedek alatt fejlődtek ki, érdekes módon a mesterséges intelligencia számára új szempontok merülnek fel a társadalmi-technikai rendszerekben való alkalmazásuk során. Nyilvánvalóan,

- érthetőség, azaz a felhasználók és az üzemeltetők megismerhetik a rendszerek jellemzőit és szolgáltatásait,
- értelmezhetőség, azaz a felhasználók és az érintettek hozzáférhetnek az eredmények és azok lehetséges hatásainak tisztázásához,
- nyomon követhetőség, azaz a felhasználók és az érintettek hozzáférhetnek az eredmények részletesebb elemzéséhez egy adott helyzet/probléma kapcsán,
- magyarázhatóság, azaz a felhasználók és az érintettek leírást, érvelést és indoklást kapnak az eredményekről, valamint
- a méltányosság, azaz az érintettek azonos bánásmódban részesülnek a közösen elfogadott kezelési szabályok szerint, sokkal nagyobb lendületet kap.

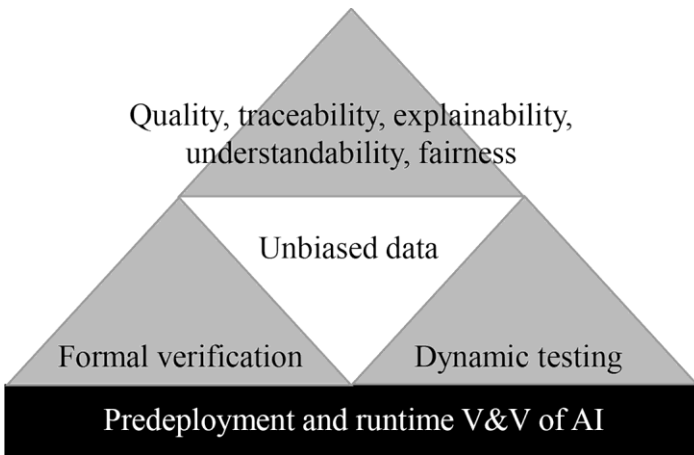
3. AI ellenőrzés és validálás

A mesterséges intelligencia ugyanis bizonyos mértékig további módszereket és eszközöket igényel a V&V-hez (Van Wesel, 2017), mivel a jól bevált tesztelési technológiák hiányoznak a mesterséges intelligencia V&V-jéhez. Ez nemcsak a további szociotechnikai minőségi szempontok miatt igaz (lásd fentebb), hanem a logikai alapú szoftverek (az eddigi szoftverek többsége általában és az AI egy része) és a statisztikai alapú szoftverek (az AI többsége, különösen a gépi tanulásban) eltérő jellege miatt is. A tesztelésnek korlátai vannak az ML dinamikája, a problémataromány pusztá mérete és a mögöttes orákulumprobléma tekintetében (Xie, 2011).

Ezenkívül a mesterséges intelligencia nagy részét az adatok irányítják. Ebben az értelemben a neurális háló egy általános funkcióközelítő, amelynek szerkezete csak nagyon kis mértékben tükrözi a tényleges funkcionáltságot. Ezért a forráskód-orientált V&V technikák, mint például a statikus elemzés vagy a white-box tesztek csak korlátozottan használhatóak ebben a kontextusban. Másrészt az adatok megbízhatósága és minősége központi kérdéssé válik a rendszerek általános minősége szempontjából.

Mivel azonban a szoftverek szisztematikus dinamikus tesztelése a legismertebb és leghatékonyabb V&V módszer, valószínűleg az ML tesztelésének is ez lesz a fő alapja. Az elmúlt évtizedekben a kutatás ipari szintű technikákat fejlesztett ki a tesztelés minőségének, hatékonyságának és megbízhatóságának növelésére. Ide tartoznak különösen a dinamikus tesztelésre vonatkozó automatizálási stratégiák, mint például a tesztvégrehajtás automatizálása olyan teszttechnológiákkal, mint a TTCN-3 (Testing and Test Control Notation [Grabowski, 2003]), a tesztek generálásának automatizálására szolgáló modellalapú tesztelés (MBT [Utting, 2012]), valamint a keresési és optimalizációs algoritmusok alkalmazása az automatizált teszt kiválasztásra és a tesztcsomagok csökkentésére (Harman, 2015). Ezen túlmenően a dinamikus tesztelés olyan verifikációs megközelítésekkel, mint a forráskód-elemzés, a modellellenőrzés és a szimbolikus végrehajtás kombinálása olyan fejlesztéseket tesz lehetővé a tesztelésben, amelyek a verifikációs folyamatok szigorát a dinamikus tesztelés skálázhatóságával ötvözik (Godefroid, 2018). Ezeket a technikákat a funkcionális és a funkcionáltságon kívüli tulajdonságok, például a teljesítmény vagy a biztonság tesztelésére is alkalmazzák (Schiefer-decker, 2012). Végül a tesztelés szoros integrációja a rendszerfejlesztési folyamatokkal és a kockázatkezeléssel (Felderer, 2014) javította a tesztelés hatékonyságát és átláthatóságát, így a tesztelés az egyik legfontosabb szoftverminőségi intézkedéssé érett az iparban. Mégis, a teszt automatizálás, valamint a modellek használata a tesztelésben még mindig alulkielégített: bár erős teszt automatizálásra van szükség, a szoftverteszteléssel foglalkozó szakemberek kevesebb mint 14%-a állítja, hogy MBT-t használ (Binder, 2015). A kockázatalapú tesztelésben rejlő lehetőségeket a tesztelési folyamatok bizonytalanságok alapján történő

irányítására különösen a kritikus rendszerek területén mutatták ki a biztonság és a védelem szempontjából, ami hasonlóképpen alkalmazható lesz a mesterséges intelligenciára is (Erdogan, 2014).



4. ábra: Az AI V&V piramis. A mesterséges intelligencia-alapú rendszereket mind a telepítést megelőző fázisokban, mind pedig a futás idején ellenőrizni és validálni kell. A formális verifikáció és a dinamikus tesztelés V&V módszereinek kombinációja ajánlott, különösen a biztonság- és védelemkritikus AI-alapú rendszerek esetében. A V&V segít a minőségi és a megmagyarázhatósági követelmények biztosításában, valamint lehetővé teszi a mesterséges intelligenciában használt (képzési) adatok torzításának igazolását.

Az ML hitelesítésére és validálására szolgáló célzott módszerekkel kapcsolatos kutatások még csak a kezdeteknél tartanak. Ennek ellenére a tesztelés már része az ML általános képzési felépítésének, a legtöbb tesztelés a kezdeti képzési célokhoz képest pontosabb modellek elérése érdekében történik. A felügyelt tanulásban például tesztelési és validálási adathalmazokat használnak az ML-modell értékelésére. A validációs adathalmazokat tipikusan a képzés során használják a modell paramétereinek finomhangolására, míg a tesztadathalmazokat a végleges modellen használják az általánosítási hibák mérésére. Mivel azonban az egyes tesztadathalmazok csak a modell egyszeri értékelését biztosítják, és csak korlátozottan képesek az eredmények bizonytalanságának jellemzésére, a modell kiválasztásához fejlettebb statisztikai tesztelési megközelítéseket, például keresztvalidálást használnak.

Ghosh et al (2016) úgy kombinálja az ML-t és a modellellenőrzést, hogy ha a kívánt logikai tulajdonságokat egy betanított modell nem elégíti ki, akkor a modellt ("modelljavítás") vagy azokat az adatokat, amelyekből a modellt tanulták, szisztematikusan módosítják ("adatjavítás"). Fulton és Platzer (2018) a formális verifikáció és az ellenőrzött futásidejű felügyelet kombinálását javasolja, hogy garantálni lehessen a biztonságos tanulást. Az ap- proach beavatkozik a tanulási folyamatba, amikor a biztonsági tulajdonságok sérülnek, és úgy irányítja a tanulási folyamatot, hogy az eredmény megfeleljen a verifikációs modellnek. Az olyan megközelítések, mint a DeepXplore (Pei, 2017), a DLFuzz (Guo, 2018) és a

Ten- sorFuzz (Odena, 2018) metrikákat biztosítanak a neurális lefedettség számszerűsítéséhez és egyszerűsítik a tesztek automatizálását. A DeepTest (Tian, 2018) lehetővé teszi a szisztematikus tesztelést.

neurális hálózatok reálisan változó környezeti feltételek mellett, különösen az autópárhban való felhasználásra.

Az ML egyik társadalmi-technikai korlátja az átláthatóság hiánya, azaz a "fekete doboz" megközelítés. Ennek kezelése érdekében különböző megközelítéseket dolgoztak ki, például a következőket

- modellértelmezés a képosztályozásokhoz, pl. az aktivitásmaximalizálás megértése a saliency térképekkel (Simonyan, 2013),
- modellmagyarázat érzékenységi elemzéssel és helyi magyarázó vektorokkal, hogy bármely osztályozási módszer döntéseinek okait bemutassa (Baehrens, 2010),
- modellbontás az általános többrétegű neurális hálózatok értelmezéséhez, a hálózat osztályozási döntésének a hálózat beillesztett elemeinek hozzájárulásaira való bontásával (Montavon, 2017),
- döntési fák kinyerése képzett neuronhálózatokból származó bemeneti adatokból (Krishnan, 1999),
- a relevancia terjedése nemlineáris osztályozók pixelenkénti bontásával (Bach, 2015), és
- dekonvolúciós módszerekkel, hogy betekintést nyerjünk a köztes szerkezeti rétegek működésébe és az osztályozók működésébe (Zeiler, 2014).

Egy másik jól bevált módszer a tesztforatókönyvek, azaz tesztesetek és tesztadataik felhasználása az ML-döntések magyarázatára. Az ML másik szociotechnikai korlátja a méltányosság esetleges hiánya, azaz az esetleges elfogultság. Ebben segít a tesztadatok generálásából ismert, a szükséges kategóriákat és tulajdonságokat jól lefedő (képzési) adatok szisztematikusan generálása (Nguyen, 2016).

A mesterséges intelligencia hatékony tesztelésének képessége alapvető fontosságú lesz a széles körű elfogadásához, és központi szerepet játszik az olyan biztonságkritikus területeken, mint a közlekedés és az autópárh, az egészségügy vagy az ipari automatizálás. A tesztelési technológiák, eszközök, tesztforatókönyvek, tesztesetek és tesztadatok biztosítása a mesterséges intelligencia számára nemcsak a V&V szilárd alapja lesz, hanem segít a mesterséges intelligencia magyarázatában és átláthatóbbá és elfogulatlanabbá tételében is. Ezek az eszközök a mesterséges intelligencia futás közbeni biztonságának és védelmének biztosítására is felhasználhatók.

És végül, de nem utolsósorban, a mesterséges intelligencia védelmét szolgáló eszközök is hozzájárulnak a mesterséges intelligencia demokratizálódásához: ezek képezik az eredmények megerősítésének vagy tanúsításának alapját, amikor a mesterséges intelligencián alapuló rendszereket kell elszámoltatni. Az AI-alapú rendszerek összehasonlításának és teljesítményértékelésének digitális közös eszközeivé is válhatnak, és ezáltal hozzájárulhatnak az AI közös tudáslapjához.

Visszaigazolás

Ez a cikk a Diana Serbanescu által vezetett, a mesterséges intelligencia-alapú rendszerek kritikájával foglalkozó kutatócsoporttal, a Dirk Messner és Sabine Schlacke által vezetett, a Globális Változások Német Tanácsadó Testületével, valamint a "The Democratization of AI. Net Politics in the Era of Learning Algorithms" című, Andreas Sudmann vezette 2018 bochumi konferencián, valamint a Jürgen Großmann vezette "Industrial-grade Verification and Validation of Evolving Systems (IVVES)" című ITEA3 projekt tagjaival.

Hivatkozások

- Bach, Sebastian, Alexander Binder, Grégoire Montavon, Frederick Klauschen, Klaus-Robert Müller és Wojciech Samek (2015): "On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation", *PloS one* no10., e01301407,.
- Baehrens, David, Timon Schroeter, Stefan Harmeling, Motoaki Kawanabe, Katja Hansen és Klaus-Robert Müller (2010): "How to Explain Individual Classification Decisions", *Journal of Machine Learning Research* no11., Jun, 1803-31.
- Binder, Robert V., Bruno Legeard és Anne Kramer (2015): "Model-Based Testing: Where Does It Stand?", *Communications of the ACM* no58., 2, 52-56.
- Dijkstra, Edsger W. (1972): Dijkstra: "The Humble Programmer", *Commun. ACM* 15, no. 10, 859-66.
- Erdogan, Gencer, Yan Li, Ragnhild Kobro Runde, Fredrik Seehusen és Ketil Stølen (2014): "Megközelítések a kockázatelemzés és a tesztelés kombinált használatához: A Systematic Literature Review", *International Journal on Software Tools for Technology Transfer* 16, no. 5, 627-42.
- Felderer, Michael és Ina Schieferdecker (2014): "International Journal on Software Tools for Technology Transfer 16, no. 5, 559-68.
- Fulton, Nathan és André Platzer (2018): "Safe Reinforcement Learning Via Formal Methods: Toward Safe Control through Proof and Learning", A harmincketedik AAAI Conference on Artificial Intelligence konferencián előzetesen közzétett tanulmány.
- Ghosh, Shalini, Patrick Lincoln, Ashish Tiwari, Xiaojin Zhu és Wisc Edu (2016): "Trusted Machine Learning for Probabilistic Models", Az ICML Workshop on Reliable Machine Learning in the Wild című konferencián bemutatott tanulmány.
- Godefroid, Patrice és Koushik Sen (2018): "In Handbook of Model Checking, 613-49: Springer.

- Grabowski, Jens, Dieter Hogrefe, Réthy György, Ina Schieferdecker, Anthony Wiles és Colin Willcock (2003): "An Introduction to the Testing and Test Control Notation (Ttcn-3)", *Computer Networks* no42., 3, 375-403.
- Guo, Jianmin, Yu Jiang, Yue Zhao, Quan Chen és Jiaguang Sun (2018): "Dfuzz: Differential Fuzzing Testing of Deep Learning Systems", Paper presented at the Proceedings of the 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering.
- Hammond, Kris (2016): "The Periodic Table of AI", <https://www.datasciencecentral.com/profiles/blogs/the-periodic-table-of-ai>. Hozzáférés: 2019. július 11.
- Harman, Mark, Yue Jia és Yuanyuan Zhang (2015): "Achievements, Open Problems and Challenges for Search Based Software Testing", Paper presented at the IEEE 8th International Conference on Software Testing, Verification and Validation (ICST).
- ISO/IEC (250102011): Rendszer- és szoftvertervezés - Rendszer- és szoftverminőségi követelmények és értékelés (tér)- Rendszer- és szoftverminőségi modellek. Nemzetközi Szabványügyi Szervezet.
- Krishnan, R, G Sivakumar és P Bhattacharya (1999): "Extracting Decision Trees from Trained Neural Networks", *Pattern recognition* 32, 12. sz.
- Legg, Shane és Marcus Hutter (2007): "Univerzális intelligencia: Minds and machines" no17., 4, 391-444.
- Montavon, Grégoire, Sebastian Lapuschkin, Alexander Binder, Wojciech Samek és Klaus-Robert Müller (2017): "Explaining Nonlinear Classification Decisions with Deep Taylor Decomposition", *Pattern Recognition* 211-2265..
- Nguyen, Anh, Alexey Dosovitskiy, Jason Yosinski, Thomas Brox és Jeff Clune (2016): "Synthesizing the Preferred Inputs for Neurons in Neural Networks Via Deep Generator Networks", Paper presented at the Advances in Neural Information Processing Systems, 3387-3395.
- Odena, Augustus és Ian Goodfellow (2018): "TensorFuzz: Debugging Neural Networks with Coverage-Guided Fuzzing", arXiv preprint arXiv:1807.10875.
- Pei, Kexin, Yinzhi Cao, Junfeng Yang és Suman Jana (2017): "DeepXplore: Automated Whitebox Testing of Deep Learning Systems", Paper presented at the proceedings of the 26th Symposium on Operating Systems Principles, 1-18, ACM.
- Russell, Stuart J és Peter Norvig (2016): *Mesterséges intelligencia: A Modern Approach*. Malajzia; Pearson Education Limited.
- Schieferdecker, Ina, Jürgen Grossmann és Martin Schneider (2012): "Model-Based Security Testing", arXiv preprint arXiv:1202.6118.
- Simonyan, Karen, Andrea Vedaldi és Andrew Zisserman (2013): "Deep inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps", arXiv preprint arXiv:1312.6034.

- Stone, Peter, Rodney Brooks, Erik Brynjolfsson, Ryan Calo, Oren Etzioni, Greg Hager, Julia Hirschberg, Shivaram Kalyanakrishnan, Ece Kamar, Sarit Kraus, Kevin Leyton-Brown, David Parkes, William Press, AnnaLee Saxenian, Julie Shah, Milind Tambe és Astro Teller (2016): "Mesterséges intelligencia és élet a mesterséges intelligencia 2030.százéves tanulmányában. A 2015-ös tanulmányozó testület jelentése", 52. Stanford, Kalifornia: Stanford Egyetem.
- Tian, Yuchi, Kexin Pei, Suman Jana és Baishakhi Ray (2018): "DeepTest: Automated Testing of Deep-Neural-Network-Driven Autonomous Cars", Paper presented at the Proceedings of the 40th International Conference on Software Engineering, 303-314. ACM.
- Utting, Mark, Alexander Pretschner és Bruno Legeard. "A Taxonomy of Model-Based Testing Approaches", Software Testing, Verification and Reliability 22, no. 5 (2012): 297-312.
- Van Wesel, Perry és Alwyn E Goodloe (2017): "Challenges in the Verification of Reinforcement Learning Algorithms", NASA/TM-2017-219628.
- WBGU (2019): A közös digitális jövőnk felé. Zászlóshajójelentés. Berlin: German Advisory Council on Global Change.
- Wirth, Niklaus. "A Software Engineering rövid története", IEEE Annals of the History of Computing 30, no. 3 (2008): 32-39.
- Xie, Xiaoyuan, Joshua WK Ho, Christian Murphy, Gail Kaiser, Baowen Xu és Tsong Yueh Chen (2011): "Testing and Validating Machine Learning Classifiers by Metamorphic Testing", Journal of Systems and Software no84., 4, 544-58.
- Zeiler, Matthew D és Rob Fergus (2014): "Visualizing and Understanding Convolutional Networks", Paper presented at the European conference on Computer Vision, 818-833, Springer.

Mesterséges intelligencia, demokrácia és jog

Christian Djeffal

A digitális technológiák átalakítják demokráciánkat. Miközben eligazodást és útmutatást keresünk ebben a folyamatban, a technológia és a demokrácia közötti kapcsolat nem egyértelmű, és úgy tűnik, hogy fluxusban van. A technológia és a demokrácia tükrözik egymást?¹ Az internetet először úgy üdvözölték, mint valóban demokratikus technológiát és a demokrácia végső eszközét. Ma már gyakran a demokráciát fenyegető komoly veszélynek tekintik. A mesterséges intelligencia (AI) története talán éppen ellenkezőleg alakulhat. Míg a mesterséges intelligenciáról sokan úgy vélekednek, mint a demokrácia fenyegetéséről vagy akár a demokrácia végéről,² néhányan a mesterséges intelligenciában rejlő demokratikus lehetőségeket emelik ki.³ Ahogy az gyakran előfordul, a kutatási eredmények a kutatás alapjául szolgáló premisszáktól függenek. Ez a fejezet abból az állításból indul ki, hogy a technológiák és a média nagymértékben alakítják az emberi ügyeket, de a technológiát viszont az emberi döntések és választások is alakítják. Hatalmas lehetőség van a demokratikus folyamatok veszélyeztetésére, kijátszására vagy akár megszüntetésére. Ellenkezőleg, a demokrácia előmozdítására is adódhatnak lehetőségek. Ezért az, hogy a mesterséges intelligencia milyen mértékben befolyásolja a demokráciát, attól függ, hogy milyen utat választanak a kutatás, a fejlesztés és a társadalomban való alkalmazás során.

E fejezet fő célja, hogy rávilágítson a mesterséges intelligencia konstruálásában rejlő választási lehetőségekre és annak a demokrácia jövőjére gyakorolt hatásaira. Azt is megvizsgálja, hogy a jog és a jogtudomány hogyan viszonyul ezekhez a kérdésekhez. Ebből a szempontból a mesterséges intelligenciának a demokráciára gyakorolt jelenlegi hatásai fontos iránymutató funkciót töltenek be. A különböző szintű találmányok és szabályozási intézkedések további lehetőségeivel szemben azonban ezek csak előfutárai annak, ami lehetséges lesz és lehetségesnek kell lennie. Ebben az értelemben ez a fejezet kísérlet arra is, hogy olyan fejleményekkel és találmányokkal foglalkozzunk, amelyeket még nem tudunk felfogni. A fő érv az, hogy mégis lehetséges lehet befolyásolni őket. Ezért a fejezet jogi és jogtudományi szempontból a mesterséges intelligencia demokratizálásának lehetőségéről és szükségességéről fog elmélkedni. Ezt követően a mesterséges intelligencia demokratizálásának különböző módjait vizsgálja meg.

1 Erről a kérdésről lásd Hofmann (2018).

2 Hofstetter (2016); O'Neil (2016).

3 Helbing (2019); Ennals (1987).

I. A mesterséges intelligencia demokratizálása: lehetőség és szükségszerűség

A. A mesterséges intelligencia nyitottságának és erejének megértése

A mesterséges intelligencia és a demokrácia közötti kapcsolat megértéséhez tisztázni kell a mesterséges intelligencia fogalmát. A fogalom döntő aspektusa nem a mesterséges intelligencia egyértelmű meghatározásában, hanem nyitottságában rejlik. A mesterséges intelligencia nagyon tág fogalom, és talán ez az oka annak, hogy ez a fogalom megelőzött más elképzeléseket, például a kibernetikát, és ma a tudományban, a politikában és a gazdaságban használt általános fogalom. A mesterséges intelligencia egy olyan kutatási kérdést jelölő kifejezés, amely ma az informatika egy egész részdiszciplínáját in- tézi. Ezt a kutatási kérdést a következőképpen foglalták össze: Képesek-e a rendszerek önállóan megoldani komplex problémákat?⁴ A nyitottság már a fogalom kezdeti, 1955-ös meghatározásában, a Rockefeller Alapítványnak benyújtott támogatási javaslatban is megmutatkozik:

Javasoljuk, hogy a New Hampshire állambeli 1956Dartmouth College-ban (Hanover, New Hampshire) a nyár folyamán egy 2 hónapos, a mesterséges intelligenciáról szóló 10tanulmányt folytassanak. A tanulmányt azon feltételezés alapján kell lefolytatni, hogy a tanulás vagy az intelligencia bármely más jellemzőjének minden aspektusa elvileg olyan pontosan leírható, hogy egy gép képes lesz azt szimulálni. Megkísérlik kideríteni, hogyan lehet a gépeket nyelvhasználatra, absztrakciók és fogalmak kialakítására, olyan problémák megoldására, amelyek ma az ember számára vannak fenntartva, és önmaguk fejlesztésére készíteni.⁵

Okos dolog úgy megfogalmazni egy támogatási javaslatot, hogy az ösztönözze az olvasók fantáziáját. A mesterséges intelligencia használata itt pontosan ezt teszi. A mesterséges intelligencia nyitottságával kapcsolatban az első szempont, amely ebből a meghatározásból levezethető, az, hogy a mesterséges intelligencia egy kutatási kérdés. Nem egy elmélet, amely magyarázatokat kínál. Nem egy általános hipotézis vagy egy elképzelés, amely bizonyos szempontokat meghatározott módon keretbe foglal. Az általános kutatási kérdés, hogy a rendszerek képesek-e önállóan megoldani komplex problémákat, bizonyos feltételezéseken alapul, de ezek minimálisra redukálódnak. Az a tény, hogy a mesterséges intelligencia egy kérdés, magyarázatot adhat a kifejezés hosszú életére is. A mesterséges intelligencia annyi hullámvölgyet élt meg, hogy a kommentátorok "mesterséges intelligencia térről" és "mesterséges intelligencia nyárról" beszélnek.⁶ Mindaddig, amíg a mesterséges intelligencia alapjául szolgáló általános kutatási kérdést nem oldják meg úgy, hogy az ne legyen továbbfejleszhető, addig a mesterséges intelligencia továbbra is érdekes marad a kutatók számára. A mesterséges intelligencia nyitottságának másik aspektusa az alapfeltevésekkel kapcsolatos. John McCarthy, az egyik pályázó és a mesterséges intelligencia kutatásának fontos szereplője megjegyzései arra utalnak, hogy a mesterséges intelligencia

kifejezést azért találták ki, hogy elkerüljék a kibernetikai feltételezéseket.

4 Mainzer (2019: 3).

5 McCarthy/Minsky/Shannon (1955).

6 Sudmann (57).

ics kutatás és Norbert Wiener hatásának megkerülése érdekében.⁷ Bár Wiener kétségtelenül nagyban hozzájárult a számítástechnika területéhez, és a mesterséges intelligencia számos fontos, ma is aktuális kérdését érintette, mindezt más szemszögből tette. A kibernetikáról alkotott elképzelése, amelyet sok más fontos kolléga is általánosan vallott, egy általános elmélet, erős feltételezésekkel. Ezzel szemben a mesterséges intelligencia kifejezés hagyományosan egészen más nézeteket takar. Az egyik általános nézeteltérést erős és gyenge AI-hipotézisnek nevezték el.⁸ Az erős AI tézis eltért attól az elképzeléstől, hogy az AI képes megismételni vagy akár felülmúlni az ember intelligenciáját. Ezzel szemben a gyenge AI tézis csak azt követeli meg, hogy a gépek úgy viselkedjenek, mintha intelligensek lennének. Általában bizonyos konkrét megoldandó problémákra összpontosít.

A mesterséges intelligencia nyitottságának másik aspektusa, hogy nem egyetlen technológiához, hanem technológiák egész sorához kapcsolódik.⁹ Jelenleg a gépi tanulás¹⁰ technológiái vagy a technika jelenlegi állásának, vagy akár "valódi mesterséges intelligenciának" tekinthetők. A mesterséges neuronhálózatok például bizonyos feladatokat látnak el, például a képfelismerést. Ezeket nagy mennyiségű képzési adat alapján képzik ki, amelyeket címkéznek, hogy a tanulás alapjául szolgáló matematikai modelleket folyamatosan adaptálni és fejleszteni lehessen. Ezzel szemben a generatív adverszális hálózatok emberi képzési adatok nélkül, adverszális módon fejlesztik magukat. Még mindig számos olyan általános ötlet és architektúra létezik, amelyek a múltban népszerűbbek lehettek, mint például a döntési fák, vagy amelyek a jövőben népszerűbbek lehetnek, mint például az evolúciós mesterséges intelligencia. Mivel a mesterséges intelligencia nyitott az új megközelítések és áttörések felé, a mesterséges intelligencia kutatása továbbra is mozgó célpont. Azok a rendszerek, amelyek egy adott időpontban a legmodernebb mesterséges intelligenciát képviselik, később már nem minősülnek igazán intelligensnek. A különböző technológiák különböző erőforrásokat igényelnek. Bár a mesterséges intelligenciát néha olyan big data-alkalmazásokkal hozzák összefüggésbe, amelyek hatalmas mennyiségű adat képzésére vagy elemzésére támaszkodnak, a big data nem feltétlenül szükséges követelmény. Vannak kis adatmennyiségű alkalmazások vagy olyan alkalmazások is, amelyek egyáltalán nem igényelnek jelentős képzési adatokat. Az erőforrások ennek megfelelően változnak. A mesterséges neurális hálózatoknak nagy mennyiségű képzési adatra, az adatok tárolásához elegendő memóriaterületre és a számításához elegendő teljesítményre van szükségük. Különösen fontos megjegyezni, hogy a képzési adatokat embereknek kell megjegyezniük. Legyen szó akár röntgenfelvételek értelmezéséről, bőrrák felismeréséről vagy a zebrák felismeréséről az automatizált vezetés keretében, a mély neurális hálózatok betanításához szükséges adatok emberi inputtól függenek. A nagy emberi erőforrás-állományok még fontosabbak voltak a régi, népszerű szakértői rendszereknél.

- 7 McCarthy (1989).
- 8 A vitát lásd Russell/Norvig/Kirchner (2012: 1020).
- 9 Gasser/Almeida (2017: 59).
- 10 A különböző nézőpontokból készült áttekintéseket lásd Shalev-Shwartz/Ben-David (2014); Sudmann/Engemann Goodfellow/Bengio/Courville (2016).

az 1990-es években. A szakértőknek sok esetben döntési fákat kellett tervezniük, amelyek aztán másoknak segítséget nyújtottak.

Továbbá a mesterséges intelligencia általános céljai is nyitottak. Bár gyakran azt állítják, hogy a mesterséges intelligencia az automatizálás szinonimája, valójában nem egységes a vélemény abban a tekintetben, hogy a mesterséges intelligencia célja inkább a kiegészítés, mint az automatizálás. Míg az automatizálás az emberek gépekkel való helyettesítésére vonatkozik, addig a bővítés az ember-gép interakcióra összpontosít az emberi képességek felerősítése érdekében. Ez az augmentációs paradigma a számítástechnika különböző területein bizonyult befolyásosnak. Már a legbefolyásosabb képviselőjének, Douglas Engelbartnak a legkorábbi kutatási programja is azt mutatja, hogy egyértelmű kapcsolat van a mesterséges intelligencia kutatási programjával:

Az "emberi értelem bővítése" alatt azt értjük, hogy növeljük az ember képességét arra, hogy megközelítsen egy összetett problémás helyzetet, hogy a sajátos igényeinek megfelelő megértést nyerjen, és hogy megoldásokat találjon a problémákra. A megnövekedett képesség ebben a tekintetben a következők keverékét jelenti: gyorsabb megértés, jobb megértés, a korábban túl bonyolult helyzetben a megértés hasznos fokának megszerzésének lehetősége, gyorsabb megoldások, jobb megoldások, és a korábban megoldhatatlannak tűnő problémák megoldásának lehetősége. És az "összetett helyzetek" alatt értjük a diplomátok, vezetők, társadalomtudósok, élettudósok, természettudósok, ügyvédek, tervezők szakmai problémáit - függetlenül attól, hogy a problémás helyzet húsz percig vagy húsz évig áll fenn. Nem elszigetelt okos trükkökről beszélünk, amelyek bizonyos helyzetekben segítenek. Egy olyan életmódra utalunk egy olyan integrált területen, ahol a megérzések, a vágás és próbálkozás, az intangibilis és az emberi "helyzetérzék" hasznosan együtt élnek erős fogalmakkal, egyszerűsített terminológiával és jelölésekkel, kifinomult módszerekkel és nagy teljesítményű elektronikus segédeszközökkel.¹¹

Az ilyen automatizálásra képes rendszereknek "kifinomultnak" kell lenniük, és képeseknek kell lenniük a komplexitás kezelésére. Ezeket a rendszereket emberi intelligenciával kell kombinálni. Nem az a céljuk, hogy helyettesítsék azt. A mesterséges intelligencia általános célja tehát a kiegészítés és az automatizálás tekintetében is nyitott. A mesterséges intelligencia és az ember kapcsolatában az általános cél nyitottsága tükröződik a mesterséges intelligencia rendszerek felhasználási céljainak sokféleségében. A mesterséges intelligencia kifejezéssel leírt technológiák összessége az úgynevezett általános célú technológiák (GPT). Bár a GPT-k fogalmát elsősorban a közgazdaságtanban alkalmazták,¹² a technológia társadalmi hatásainak elemzésére is jól alkalmazható kategória. A szén és az acél felhasználásának sokféle célját a "kardok szántóvá válása" kifejezéssel fogalmazták meg. Talán ugyanez lehet

- 11 Engelbart (1963: 1).
- 12 Rousseau (2009).

a mesterséges intelligenciáról, amely a halálos autonóm fegyverrendszereket és a segítő robotokat is táplálhatja. A mesterséges intelligencia megértéséhez rendkívül fontos az adott technológiák általános célú jellegének megértése. A mesterséges intelligenciát alkotó technológiák nem kötődnek kizárólagosan sem bizonyos kockázatokhoz és kihívásokhoz, sem pedig bizonyos lehetőségekhez és előnyökhöz. Erre a tételre számos ellenkező értelmű példa van, de az adatvédelem és a magánélet védelme ismét nagyon szemléletes ebben a tekintetben. A mesterséges intelligencia rendszerek kétségtelenül veszélyt jelenthetnek a magánéletre és az adatvédelemre, mivel sok személyes információ kinyerését teszik lehetővé. Az egyik érdekes aspektus a mesterséges intelligenciával támogatott árnyékprofilkészítés. Ez azt jelenti, hogy az embereket anélkül profilozzák, hogy jelentős saját tevékenységet végeznének. Az információkat a körülöttük lévő emberek szolgáltatják. A közvetett bizonyítékok, például a közösségi hálózatban lévő más személyek keresési lekérdezései lehetővé teszik az intelligens rendszerek számára, hogy rekonstruálják egy személy profilját az adott hálózaton belül, és összegyűjtsék a releváns személyes információkat anélkül, hogy a személy bármit is felfedett volna személyesen. A mesterséges intelligencia azonban a magánélet védelmét javító technológiaként is szolgálhat. Általános igény van például az olyan chatbotokra, amelyek egy rövid és egyszerű beszélgetés során megtanulják egy személy adatvédelmi preferenciáit, majd ezt követően az adott személy által használt összes hálózatban és online szolgáltatásban adaptálják az adatvédelmi preferenciákat. A mesterséges intelligencia rendszerek céljai tehát egyszerre erősíthetik és veszélyeztethetik a magánélet védelmét. Amint azt a későbbiekben bemutatjuk, ugyanez igaz más elvekre és értékekre, például a demokráciára is.

A nyitottság aspektusa	Alternatívák
Kutatási kérdés	Gyenge mesterséges intelligencia tézis, erős mesterséges intelligencia tézis
Technológiák	Gépi tanulási technológiák (mesterséges neurális hálózatok, generatív adverzális hálózatok), a jó öreg mesterséges intelligencia (AI).
Források	Adatok, józan ész, számítások...
A felhasználás céljai	Automatizálás vs. kiegészítés
Célok	Általános célú technológiák: sokféle célt szolgálhatnak, például átláthatóságot és adatvédelmet.

1. táblázat: A mesterséges intelligencia nyitottságának dimenziói

Ennek a nyitottságnak a fontosságát akkor tudjuk jobban értékelni, ha

felismerjük, hogy a technológia milyen módon alakíthatja a demokráciát. Először is, a demokrácia jelentésének különböző értelmezései és konstrukciói léteznek. Míg a nép önrendelkezésének közös vonása létezik, addig a demokrácia és a demokrácia eltérők.

nézetek arról, hogy hogyan kell ezt az önkormányzatot gyakorolni. A demokrácia a társadalom tényleges gyakorlatában konstituálódik. A technológia mindig is hatalmas szerepet játszott a demokrácia tényleges gyakorlatában. A demokrácia és a technológia összefonódik. "A demokráciát nem megalkotják, majd közvetítik. A közvetítés aktusai révén alakul ki. A közvetítés technológiái mindig is a demokrácia társadalmi megvalósulásának velejárói voltak és vannak."¹³ El lehet tehát menni odáig, hogy bizonyos technológiák használatának gyakorlatát a demokrácia sajátos elképzeléseihez kössük.¹⁴ A technológiák használata a demokráciát kon- figurálja. A mesterséges intelligencia esetében, mint általános célú technológiák összessége, ez a konfiguráció általában nyitott. A demokrácia inkább egy folyamat, mint egy rögzített és elérhető állapot. Folyamatosan meg kell valósítani, olyan eszközökkel, mint a technológiai innováció, az intézmények, a piacok és a verseny, a jog és a közigazgatás.¹⁵ E nyitottságra való tekintettel érdekes megvizsgálni a mesterséges intelligencia jelenlegi és lehetséges felhasználási módjait a demokrácia kontextusában.

B. Empirikus meglátások

Bár az olyan általános célú technológiák, mint az internet vagy a mesterséges intelligencia nagyon különbözőképpen működhetnek, általában egy bizonyos módon írják le őket. Az internetről és a demokráciáról szóló diskurzus az internet demokráciára gyakorolt lehetséges jótékony hatásainak üdvözlésével kezdődött.¹⁶ A mesterséges intelligenciával kapcsolatban úgy tűnik, hogy ez éppen fordítva van: elsősorban a demokráciát fenyegető veszélynek tekintik. Úgy látják, hogy a mesterséges intelligencia akadályozhatja a demokratikus folyamatokat, például a választásokat és a szavazásokat. Attól is tartanak, hogy a mesterséges intelligencia sok esetben átveszi a döntéshozatalt. Ahhoz, hogy árnyaltabb képet alkothassunk, értékelnünk kell a technológia esetleges voltát és azt, hogy nagyon különböző módon lehet használni. Az internettel foglalkozó szakirodalom ma már elismeri az internet demokráciára gyakorolt pozitív és negatív hatásait.¹⁷ Az internet esetlegessége azt jelenti, hogy "mint minden korábbi médiumot, az ábécétől a televízióig, [az internetet] is az alakítja, hogy a társadalom hogyan használja a rendelkezésre álló eszközöket".¹⁸

A mesterséges intelligencia általános célú jellege a demokratikus folyamatokhoz való viszonyában is tükröződik, különösen a választások összefüggésében. Ebben a tekintetben a mesterséges intelligenciát általában fenyegetésnek tekintik. Számos kísérlet történt a választások befolyásolására különböző feladatokat ellátó automatizált rendszereken keresztül. A választásokkal összefüggésben álhíreket terjesztenek, hogy blokkolják és akadályozzák a politikai diskurzust, és hogy a választókat szemcsés szinten célozzák meg, hogy bevonják vagy elvonják őket a választásoktól.

¹³ Coleman (2017: 27).

- 14 Bozdag/van den Hoven (2015).
- 15 Irrgang (2002: 173).
- 16 Pernice (2016).
- 17 Áttekintést lásd Ceron/Curini/lacus (2017: 6).
- 18 Coleman (2017: vii).

szavazás.¹⁹ A mesterséges intelligencia és más digitális technológiák által ösztönzött tevékenységek egyike a mikrocélzás. A mikrocélzás az emberek viselkedésének befolyásolására tett kísérleteket jelenti, amelyek személyes profilok és a profil sajátos jellemzőire alapozott cselekvések alapján történnek. Ezek a profilok konkrét információkat szolgáltatnak bizonyos személyekről; az embereket így a közösségi médiában megjelenő hirdetésekkel egyénileg lehet megcélozni, ahelyett, hogy politikai plakátokkal vagy televíziós reklámokkal egy csoport részeként szólítanák meg őket. Ezek az akciók a demokratikus diskurzus befolyásolására vagy akadályozására irányuló kísérletektől az egyének tényleges döntéshozatalának befolyásolásáig vagy akadályozásáig terjedhetnek. Míg az első mikrocélzási törekvéseket a közvélemény-kutatási kampányok során alkalmazták, amelyek során emberek házról házra jártak a választók befolyásolása érdekében, a mesterséges intelligencia szerepet játszhat az egyes emberek részletes profiljain alapuló akciókban is. Az ilyen technológiák alkalmazásáról már több jelentés is született. Míg az Egyesült Államokban és Brazíliában a választások, valamint a Brexit-szavazás került be a hírekbe, addig olyan államokban, mint Svájc és Ausztria, szintén viták tárgyát képezte a használatuk.²⁰ A mesterséges intelligencia-rendszerek különböző szinteken növelhetik a mikrocélzás lehetőségeit. A mesterséges intelligencia segíthet az információk kinyerésében az internet feltérképezésével és a strukturálatlan adatok egyéb forrásainak elemzésével. A mesterséges intelligencia-rendszerek segíthetnek az emberek profilozásában is. A mesterséges intelligencia-rendszerek továbbá a különböző csatornákon, például a közösségi médián keresztül automatikusan megszólíthatják a személyeket a profiljuk alapján. A mikrocélzást alkalmazó kampányok több szempontból is problémásak.²¹ Először is, a vonatkozó adatokat gyakran nyilvános forrásokból gyűjtötték, egyes esetekben illegálisan. Ez sérti az érintettek adatvédelemhez való jogát, ha az adatokat jogellenesen gyűjtötték és használták fel. Ez sérti a személyes autonómiához való jogukat is, mivel az összegyűjtött adatok alapján befolyásolják őket. A mikrocélzásból való kilépés egyelőre nem lehetséges. Mi több, a mikrocélzást manipulációs célokra is fel lehet használni. A témával kapcsolatos kutatás a lehetséges előnyös hatásokat is megemlíti - például azt, hogy a választók a számukra releváns információkat kapják meg.²² Ez a választások szempontjából is relevánsabbá teheti az egyes témákat, és növelheti bizonyos csoportok jelentőségét, különösen akkor, ha azok szétszórtak és nem szervezettek.²³ A mesterséges intelligencia tehát segíthet a gyengének és kevésbé erősnek tartott csoportoknak abban, hogy több és jobb információhoz jussanak.²⁴

19 Bodó/Helberger/Vreese (2017: 3).

20 Eidgenössischer Datenschutzbeauftragter/Konferenz der schweizerischen Datenschutzbeauftragten (2018); Der Standard (2019).

21 A fenyegetések feltérképezése megtalálható Zuiderveen Borgesius/Möller/Kruikemeier/Ó

Fathaigh/Irion/Dobber/Bodo/Vreese (2018: 87) Ugyanezen az oldalon gyűjtöttek referenciát a magánélet védelmére és a manipulációs trendekre.

22 Zuiderveen Borgesius/Möller/Kruikemeier/Ó Fathaigh/Irion/Dobber/Bodo/Vreese (2018: 84ff).

23 Ibid.

24 Ennals (1987: 14).

Ez ismét rámutat a mesterséges intelligencia általános célú jellegére és arra, hogy nehéz azt egy dobozba zární. A mikrocélzás lehet káros, de lehet hasznos is a demokrácia számára. A ma létező alkalmazások azonban csak előzetes képet adnak arról, hogy mi minden lehetséges. A technológiai fejlesztések, de ami még fontosabb, a technológia kreatív és innovatív felhasználása még inkább elősegítheti a mesterséges intelligencia megoldások demokráciára gyakorolt hatását. Az AI-megoldások jelenthetnek valami igazán újat, vagy felforgathatják a meglévő lehetőségeket. Az egyik példa erre a választóknak a jelöltek célzása és profilozása révén történő befolyásolása. Egy intelligens keresőmotor segíthetne azonosítani azokat az információkat, amelyek arra vonatkoznak, hogy a pártok vagy jelöltek hogyan gondolkodnak bizonyos kérdésekről. A választók még nagyobb felhatalmazása révén olyan mesterséges intelligencia-rendszerekkel lehetne előállni, amelyek előre jelzik a jövőbeli kormányzati magatartást. Megpróbálhatnánk kiszámítani annak valószínűségét, hogy a pártok vagy jelöltek bizonyos ígéretek szerint cselekszenek. Valóban, nem tűnik teljesen lehetetlennek megjósolni annak a kérdésnek a valószínűségét, hogy bizonyos ígéreteket a jövőben be fognak-e tartani. Ez a profilalkotás teljesen más módon történő felhasználása lenne. Bár a jelöltek és pártok ilyen jellegű profilalkotása egy sor problémát és kérdést vet fel, azt mutatja, hogy a mesterséges intelligencia használata nagyban változhat, és támogathatja a választók felhatalmazását is. Megnyithatná döntéshozatali lehetőségeiket, ahelyett, hogy beszűkítené azokat. Miközben jelenleg nagy aggodalomra ad okot a mesterséges intelligencia választásokkal és szavazásokkal kapcsolatos alkalmazása, a mesterséges intelligencia jövőbeli hatása valójában nyitott.

C. Jog és technológia: Korlátozás, motiváció, tervezés

A jog és a technológia kapcsolata sokrétű. Ez a kapcsolat nagyjából három funkcióban foglalható össze: korlátozás, motiváció és tervezés. A jognak a technológia és a demokrácia kapcsolatára gyakorolt hatását ezek mentén magyarázzuk. A jog több szempontból is hozzájárulhat a mesterséges intelligencia demokratizálódásához. Különösen fontos, hogy mindezeket a funkciókat bevonjuk a képbe, mivel ezek különböző perspektívákat emelnek ki, amelyek a legalkalmasabbak arra, hogy teljes képet alkossunk a mesterséges intelligencia demokráciával kapcsolatos kihívásairól és lehetőségeiről.

1. Jogi korlátok és demokrácia

Az emberi jogok korlátozzák a mesterséges intelligencia alkalmazását, különösen a hatóságok által. Az emberi jogok a demokratikus indoklás szükségességét is kiváltják. Ezáltal tovább korlátozzák a mesterséges intelligencia lehetséges felhasználását. A jognak a technológiát korlátozó funkciója talán a legismertebb funkciója. Az adatvédelemből eredő jogi kötelezettségek például több szempontból is korlátozzák a technológia használatát. Az adatvédelmi jogszabályok betilthatják a képzési adatok gépi

tanulásban való felhasználását, mivel előfordulhat, hogy nincs jogi alapja az ilyen felhasználásnak, vagy a meglévő engedélyek nem fedik le az adott célt. Az EU általános adatvédelmi rendelete (GDPR) értelmében például az adatprodukciónak az adatvédelemre vonatkozó

A Bizottság a rendelet 5. cikke alapján a 6.²⁵ Az automatizált döntéseket hozó rendszereknek meg kell felelniük az Art. GDPR 22. CIKKÉNEK. Ez a rendelkezés csak akkor teszi lehetővé az ilyen döntéseket, ha a szakaszokban és 2 foglalt követelmények 3teljesülnek.²⁶ A szakaszok és 2 a szakaszok a szerződésen, jogszabályon vagy kifejezett hozzájáruláson alapuló döntésekre vonatkoznak.³

2. Motiváció

A jog különböző formákban is ösztönözheti a technológia használatát. Ez a motiváció vonatkozhat "a technológia fejlesztésére, fejlesztésére és alkalmazására a közigazgatás által, vagy akár kötelezővé is teheti azt".²⁷ A demokrácia mint jogi elv különböző módon motiválhatja a technológia és konkrétan a mesterséges intelligencia alkalmazását. A demokráciáról szóló nemzetközi eszközöket vizsgálva különböző kontextusokban szembesülhetünk a technológiával kapcsolatos kérdésekkel.²⁸ Az emberi jogi jogban számos olyan jog van, amely a demokratikus kormányzásra utal. Egyes emberi jogi eszközök kifejezetten rámutatnak a technológia döntő fontosságára a demokrácia erősítése érdekében.²⁹ Az egyik terület, ahol ez különösen fontos, az olyan személyek bevonása, akik még nem képesek hatékonyan részt venni a demokratikus eljárásokban és a demokratikus diskurzusokban. A fogyatékossgal élő személyek jogairól szóló ENSZ-egyezmény. Art. 4 para. 1. g) pontja kötelezi "a fogyatékossgal élő személyek számára alkalmas új technológiák - beleértve az információs és kommunikációs technológiákat, a mobilitási segédeszközöket, eszközöket és asszisztens technológiákat - kutatásának és fejlesztésének megkezdésére vagy előmozdítására, valamint az új technológiák rendelkezésre állásának és használatának előmozdítására, elsőbbséget biztosítva a megfizethető költségű technológiáknak". Ez egy példa a progresszív emberi jogi záradékra, amely arra ösztönzi az államokat és más szereplőket, hogy az emberi jogok előmozdítása érdekében alkalmazzanak technológiákat. Számos mesterséges intelligencia-technológia segíti a fogyatékkal élőket, különösen a vakokat és a halottakat. Ezek a technológiák arra is képessé teszik a megfelelő felhasználóikat, hogy részt vegyenek a demokratikus diskurzusban. Ezért az emberi jogokról és az emberi jogokról szóló 4. cikk (1) bekezdésének g) pontja hatással van az emberek demokratikus befogadására.

3. Tervezés

A jogszabály másik funkciója a tervezési folyamat strukturálása és irányítása. A törvény meghatározza a tervezési célokat, megmutatja, hogyan lehet egyensúlyt teremteni a különböző célok között, és még a műszaki szintű problémák megoldásának lehetőségeit is kiemeli. Erre jó példa a tervezési jogról szóló záradék az Art. sec251,:

25 A cikk úgy rendelkezik, hogy az adatfeldolgozás csak akkor jogszerű, ha a követelmények teljesülnek. 26 Abel (2018); Martini (2018).

27 Djeffal (2019: 16. bekezdés).

28 Ez a kutatás Ehm/Walter (2015) dokumentumgyűjteményén alapul.

- 29 Lásd például ga-Res. 68/164. Az ENSZ szerepének megerősítése az időszakos és valódi választások és a demokratizálódás előmozdításában, ENSZ A/ RES/68/164, elfogadva a Közgyűlés által 2013. december 18-án (70. plenáris ülés).

Figyelembe véve a technika állását, a megvalósítás költségeit, az adatkezelés jellegét, hatályát, összefüggéseit és céljait, valamint a természetes személyek jogait és szabadságait érintő, különböző valószínűségű és súlyosságú kockázatokat, az adatkezelő mind az adatkezelés eszközeinek meghatározásakor, mind magának az adatkezelésnek az időpontjában, megfelelő technikai és szervezési intézkedéseket - például álnevesítés - hajt végre, amelyek célja az adatvédelmi elvek - például az adatok minimalizálása - hatékony végrehajtása, valamint a szükséges biztosítékok beépítése a feldolgozásba az e rendeletben foglalt követelmények teljesítése és az érintettek jogainak védelme érdekében.

Art. A 25. cikk (1) bekezdése közvetlen kötelezettséget ír elő arra, hogy a magánélet védelmére vonatkozó megfontolásokat figyelembe kell venni az alkalmazás tervezése vagy elfogadása során. Lehetséges azonban inkább közvetett kötelezettség is. Nemrégiben azt állították, hogy az olyan alkotmányos elveket, mint az emberi jogok, a jogállamiság és a demokrácia, szintén be kell vonni a mesterséges intelligencia tervezésének folyamatába.³⁰ Ez elősegítené a jog azon funkcióját, hogy már nagyon korai szakaszban befolyásolja a technológiákat. Ezeket a kötelezettségeket a rendszereket fejlesztőknek közvetlenül is be kell tartaniuk. E kötelezettségek teljesítése érdekében különböző területeken számos módszertant találtak ki. Míg az alkotmányos értékekkel kapcsolatban különböző szabványosítási eljárások léteznek, a mesterséges intelligenciával és a demokráciával kapcsolatban még nincs konkrét szabvány. A mai napig semmi sem írja elő azt az általános kötelezettséget, hogy a demokrácia elvét be kell építeni a mesterséges intelligencia tervezésébe.³¹

D. Jogi okok és tanulságok a mesterséges intelligencia demokratizálódásához

Ez a szakasz a mesterséges intelligencia demokratizálásának főbb jogi okait, valamint a jog és a demokrácia kapcsolatának néhány tanulságát vázolja fel. A demokrácia mint alapelv számos állam alkotmányában szerepel, akár hallgatólagosan, akár kifejezetten; a demokrácia olyan nemzetközi szervezetek számára is alapvető érték, mint az Európa Tanács.³² Egy ilyen alkotmányos elv megköveteli a megvalósítását a közszférában. E nagyon általános demokratikus követelményen kívül vannak konkrétabb tanulságok is, amelyeket a jog működéséből le lehet vonni. Három felismerést az alábbiakban részletesebben tárgyalunk.

30 Nemitz (2018).

31 A jelenlegi értékérzékeny tervezési szabványok megtalálhatók a vonatkozó ISO-projekteknel és a következő oldalakon
IEEE P7000 sorozat.

32 Lásd például az Európa Tanács májusi, 5ETS1949, sz. 001.

1. Indoklás

Amint fentebb említettük, az emberi jogok egy újabb réteggel bővítik a technológia korlátozását. Abszolút korlátokat szabnak a hatóságok viselkedésének, és rákényszerítik őket az emberi jogok megvalósítására. Az emberi jogok a demokratikus döntéshozatalhoz is kapcsolódnak. Ha egy intézkedés érinti az emberi jogokat, csak akkor lehet jogszerű, ha demokratikus indoklással van alátámasztva. A Polgári és Politikai Jogok Egyezségokmánya a következő kifejezést használja: "a jogok nem vehetők alá semmilyen korlátozásnak, kivéve azokat, amelyeket a törvény előír". Eközben az Emberi Jogok Európai Egyezménye a "törvénynek megfelelően" kifejezést használja. Ez azt jelenti, hogy az emberi jogok korlátozásáról törvényben kell rendelkezni.³³ Ahhoz, hogy igazolásnak minősüljön, a hatást a törvénynek az egyén számára érthető módon kell leírnia. A törvény itt a demokratikus *ex ante* döntéshozatal helyettesíti. Az emberi jogokra gyakorolt bármilyen hatást demokratikus döntésnek kell megelőznie, amely lehetővé teszi a pontos hatást, és biztosítékokat nyújt a túlzott és önkényes felhasználás ellen. Egy másik példa az Európai Unió Alapjogi Chartája, amely előírja, hogy "[a]z e Charta által elismert jogok és szabadságok gyakorlásának korlátozásáról törvényben kell rendelkezni".³⁴ Ez a rendelkezés kifejezetten előírja a demokratikus indoklás szükségességét. Ilyen indoklás hiányában az intézkedés szükségszerűen jogellenes. A törvény az emberi jogok érvényesítésének eszköze. Egyben a demokratikus döntések médiuma is. Az emberi jogok és a demokrácia közötti, a jog által közvetített szoros kapcsolat a mesterséges intelligencia és a demokrácia közötti kapcsolatot is érinti. Ha a mesterséges intelligenciát alkalmazó rendszerek hatással vannak az emberi jogokra, akkor alkalmazásukat indokolni kell.

A demokratikus igazolás szükségessége csak attól függ, hogy az emberi döntéshozatali képességet befolyásolja-e a mesterséges intelligencia rendszer. A demokratikus indoklást nem csak a konkrét emberi jogok váltják ki. A mesterséges intelligencia rendszerek igazolásának szükségessége minden bizonnyal az úgynevezett automatizált döntéshozatali rendszerekre (ADM) is vonatkozik, amelyek gyakran állnak a tudományos figyelem középpontjában. Ez csak egy a sok lehetőség közül, ahol az emberi jogokról lehet szó. Amikor a mesterséges intelligenciát az informatikai biztonság vagy a kritikus infrastruktúra karbantartása során megfigyelőként használják, az emberi jogok megvalósítása szempontjából kulcsfontosságú. Míg a magánélethez való jog és az önrendelkezési jog a legnyilvánvalóbb példái lehetnek az ilyen hatásoknak, más, finomabb hatásokat is figyelembe kell venni. Például az ADM-eket gyakran szabályozzák, de ez a szabályozás soha nem vonatkozik az ajánlórendszerekre. Mégis, ezek az ajánlórendszerek jelentős hatással lehetnek az emberi jogokra. Kiemelve a mesterséges és az emberi

33 Lásd például Greer (1997: 9) magyarázatát.

34 Az Európai Unió Alapjogi Chartája (HL C 391-407. o326,26.10.2012.,). A demokratikus

jogalkalmazás és az emberi jogok közötti kapcsolat nem minden emberi jogi eszközben ilyen általános. Az egyetemes emberi jogi egyezmények például csak bizonyos esetekben követelik meg a demokratikus jogalkalmazás igazolását.

a társadalmi intelligencia és az emberi jogok között meglehetősen világos kritériumot állapíthatunk meg a demokratikus meghatározottság szükségességére vonatkozóan. Ez nem a gép "hatalma" a döntéshez, hanem az emberi jogokra gyakorolt hatás. Az emberi jogok és a demokrácia közötti összefüggés korlátozhatja a mesterséges intelligencia rendszerek nyilvános használatát. Ha van ilyen kapcsolat, akkor demokratikus indoklásra van szükség - függetlenül attól is, hogy az emberi jogokat hogyan érinti.

2. Felsőbbrendűség

Az emberi jogokat érintő kérdések demokratikus igazolásának szükségességéhez kapcsolódik a legitim demokratikus döntések elsőbbségének gondolata. Ez az elképzelés abban az elképzelésben ölt testet, hogy a legmagasabb demokratikus legitimitással rendelkező szervezet által hozott normák elsőbbséget élveznek más normákkal szemben. Ezért számos olyan jogrendszer, amely a parlamentet tekinti a legmagasabb demokratikus hatóságnak, a "parlament szuverenitására" támaszkodik, és a parlamenti törvényeket minden más jogalkotási aktussal szemben elsőbbséget élveznek. Más jogrendszerek olyan normahierarchiát írnak le, amelyben az alkotmány áll az első helyen, a parlamenti aktusok pedig a második helyen. Míg az alkotmányjog a *pouvoir constituant*ól nyeri legitimitását, addig a törvényi jog a parlament legitimitására támaszkodik, és mégis más normák kevésbé legitimált szereplőktől származnak. A magasabb normák elsőbbséget élveznek az alacsonyabb normákkal szemben, konfliktus esetén az alacsonyabb normák vagy érvénytelenné válnak, vagy alkalmazhatatlanok. A jogi normák hierarchiája általában a demokratikus legitimitás különböző szintjein alapul.³⁵ Azokban az esetekben, amikor a technológia normatív erővel bír, ez az általános elképzelés azt követeli meg, hogy a jog mint a demokratikus döntés közvetítője elsőbbséget élvezzen a technológia funkcionális követelményeivel szemben, és ténylegesen a demokratikus döntéseket kell irányítania.³⁶

3. Demokratikus egyensúly helyreállítása

Jogi szempontból a demokrácia fogalma nyitott. Miközben sokféleképpen lehet megérteni és felépíteni, hogy mit kellene jelentenie a demokráciának, az alkotmányjog általában nyitott a demokrácia többféle értelmezése és elmélete felé. Ez a nyitottság lehetővé teszi a jog számára, hogy alkalmazkodjon a különböző kontextusokhoz és helyzetekhez, különösen akkor, amikor változásokról és reformokról van szó. Az ilyen reformok különböző szinteken történhetnek, de mindig megváltoztatják az előzetes demokratikus folyamatokat, sőt néha magát a demokrácia fogalmát is. Az egyik mintát, amely a bíróságok e kérdésekkel való foglalkozásának módjából kirajzolódik, az újbóli egyensúlyozás módjaként lehetne leírni. A bíróságok továbbra is flexibilisek és nyitottak maradnak a meglévő

- 35 Jogpozitívista szempontból ugyanerre a következtetésre lehetne jutni az érvényességgel kapcsolatban is. Ekkor azzal kellene érvelni, hogy az érvényesség alapvető oka a demokrácia.
- 36 Lásd például Schulz/Dankert (2016) Fontos azonban megjegyezni, hogy egy ilyen hierarchiának a demokratikus legitimáción kell alapulnia, nem pedig az elsődleges és másodlagos szabályok formális megkülönböztetésén.

folyamatok, de olyan aktív lépésekre van szükség, amelyek demokratikus szempontból újra egyensúlyba hozzák a helyzetet. Ez az egyensúly helyreállítása azt jelentheti, hogy vannak olyan intézkedések, amelyek hatékonyan demokratizálják az új intézményi berendezkedést. Ezt két, más kontextusból származó példa illusztrálhatja. Az európai integráció folyamatában számos szerződésmódosítás történt, amelyek új hatásköröket hoztak létre, vagy hatásköröket ruháztak át nemzeti szintről európai szintre. A német szövetségi alkotmánybíróságnak többször kellett foglalkoznia hatáskörök létrehozásával és átruházásával. A híres lisszaboni ítéletében a bíróság engedélyezte a hatáskörök átruházását, de a német jogrendben intézményi intézkedéseket is megkövetelt, amelyek lehetővé teszik a törvényhozás számára, hogy ténylegesen szerepet játszhasson az európai politikában. Így, bár hozzájárult a nemzetek feletti hatáskör-átruházáshoz, ezt csak azzal a feltétellel tette, hogy a nemzeti törvényhozás befolyásolni tudja a magasabb szintű politikát.³⁷ Egy másik esetben a baden-württembergi alkotmánybíróságnak a professzori kollégiumtól az egyetem elnökére történő hatáskörátruházással kellett foglalkoznia. A bíróság engedélyezte ezt a hatalomátruházást, de csak azzal a feltétellel, hogy az elnök a professzori kollégiumnak tartozik elszámolással, ami a gyakorlatban azt jelentette, hogy demokratikus választási eljárást kellett létrehozni.³⁸ Ezek az esetek azt mutatják, hogy a demokratikus folyamatokra hatást gyakorló változásokat és reformokat - jogi szempontból - nem lehet az "igen" vagy "nem" bináris módon értékelni. A változások néha demokratikus egyensúly helyreállítását teszik szükségessé. Ha viták vannak arról, hogy hogyan lehet ezeket a változásokat demokratikusan kiegyensúlyozni, akkor ezeket a vitákat végső soron jogi eljárásokban lehet rendezni. Az újbóli kiegyensúlyozás ezen kérdései fontos szerepet játszanak, amikor a cselekvéseket és döntéseket nagyobb léptékben delegálják a mesterséges intelligencia rendszerekre. Ahelyett, hogy azzal érvelnénk, hogy ez nem lenne demokratikus, a kérdés inkább az, hogy a mesterséges intelligencia rendszerekre történő átruházás kiegyensúlyozható-e? Ez a különböző demokráciaesetekben jelenlévő flexibilis szemlélet is képes arra, hogy elmozdítsa a mesterséges intelligencia és a demokrácia közötti kapcsolatot. Ahelyett, hogy azt kérdeznénk, hogy demokratizálni kell-e a mesterséges intelligenciát, a kérdés az, hogy hogyan lehet demokratizálni, és hogy a megfelelő intézkedések elegendőek-e.

II. Hogyan demokratizáljuk a mesterséges intelligenciát

Ha szükség van a mesterséges intelligencia demokratizálására, hogyan lehet azt a gyakorlatban megvalósítani? E kérdés ösztönző megközelítése először is megvizsgálná azokat az eseteket, amelyekben demokratikus döntések születnek, másodsor pedig azt, hogy miként lehet ezeket a döntéseket meghozni. Ahogy a

-
- 37 BVerfG, A második szenátus 2009. júniusi 30. ítélete-2 BvE 2/08-paras. (I-421), http://www.bverfg.de/e/es20090630_2bve000208en.html 273ff. pont.
- 38 Landesverfassungsgericht Baden-Württemberg, ítélet VB14.11.2016,I 16/15, elérhető a következő címen: https://verfgh.baden-wuerttemberg.de/fileadmin/redaktion/m-verfgh/dateien/161114_IVB16-15_Urteil.pdf 43ff. o.

minden más demokratikus döntésnél különböző eszközök állnak rendelkezésre, a parlamenti tanácskozások és döntések or- dinárius folyamataitól kezdve a demokratikus részvétel közvetlenebb formáihoz. Mindegyik módszer sajátos előnyökkel járhat egy adott környezetben. Egy ilyen demokratikus eszköztár többek között a következő elemeket tartalmazhatja:

- rendes parlamenti eljárások a mesterséges intelligencia megvitatására és szabályozására
- parlamenti szakbizottságok alkalmazása bizonyos kérdések eldöntésére
- a szakértők felhatalmazása arra, hogy bizonyos döntéseket előre meghatározott elvek szerint hozzanak meg
- a polgárok közvetlen bevonása bizonyos kérdésekben
 - részvételi módszerek
 - szelekció: véletlenszerűen kiválasztott állampolgárok csoportjainak bevonása egy hivatal betöltése vagy bizonyos döntések meghozatala érdekében.
 - véletlenszerű mintavételes szavazás: az egyes kérdésekről való szavazáshoz a lakosság reprezentatív mintáját választják ki.

A technológiákkal és a konkrét technikai artefaktumokkal kapcsolatos döntések körének megértése érdekében hasznos, ha analitikusan különbséget teszünk a különböző laikusok között, annak ellenére, hogy a különböző laikusok közötti összefüggések nyilvánvalóak. A technikai artefaktumokkal kapcsolatos konkrét döntésekre összpontosítva, vannak olyan döntések, amelyek inkább technikai jellegűek, és vannak olyanok, amelyek inkább társadalmi jellegűek. Ezért különbséget kell tenni egy társadalmi és egy technikai réteg között. Továbbá, egyes döntések nem egy adott műtárgyra, hanem inkább egy technológiára vonatkozóan születnek. Ezek a döntések a kormányzás rétegében helyezkednek el. Minden rétegben vannak konkrét kérdések, amelyeket fel kell vázolni.

A. Technológiai réteg

1. Tervezési választások

A technológia demokratikus meghatározásának fontos lépése a technológia feltalálása vagy alkalmazása során hozott döntések megértése. A fejlesztés során számos tervezési döntés születik. Néhány tervezési döntést szándékosan hoznak, néhánynak fontos következményei vannak. Demokratikus szempontból meg kell érteni és ki kell emelni a konkrét döntéseket. Ezek a döntések az architektúrára, az alkalmazásokra és az alkalmazott technológiák minden más jellemzőjére vonatkoznak. Ahol van alternatíva, ott van választás. A választási lehetőségek megértéséhez demokratikus gondolkodásmódra is szükség van, amely nyitott több lehetőségre anélkül, hogy automatikusan előnyben részesítene bizonyos eredményeket. Különösen az informatikusok, akiket konkrét célok, például a hatékonyság elérésére képeztek ki, rendszeresen nem látják az általuk preferált

értéket maximalizáló választási lehetőségek mögé.

A gépi tanulás esetében a választás megbecsülése érdekében az optimalizálás kérdései nagyon érdekesek.³⁹ A gépi tanuló rendszerek bizonyos célok elérésére optimalizálódnak, visszajelzést kapnak, és ennek megfelelően módosítják a modelljüket. Sok esetben a célok, amelyekre egy modellt optimalizálnak, nincsenek kőbe vésve, hanem inkább esetlegesek. Egy algoritmus, amely egy területen belül bizonyos iskolákba osztja be a gyerekeket, különböző célfüggvények szerint optimalizálható: Az egyik lehet a legrövidebb út az iskolába. Egy másik lehet a legbiztonságosabb út az iskolába. Ugyanakkor más célokat is meghatározhatunk, például az iskolában tanulók jó keveredését etnikai vagy gazdasági szempontból. Az ilyen választások gyakran kompromisszumokat eredményeznek. Aktív választást igényelnek. Az egyik kompromisszum, amely az utóbbi időben egyre ismertebbé vált, az adatok felhasználása és a megkülönböztetés megértése közötti választás. A ma- chine learning modelleket gyakran olyan adatokon képzik ki, amelyek implicit előítéleteket tartalmaznak - ugyanakkor a képzési adatok nem tartalmazhatnak explicit utalásokat az életkorra, a nemre vagy más kritériumokra. Így a döntő információ nincs jelen, és lehetetlenné válik annak megértése, hogy van-e torzítás az adatokban és következésképpen az algoritmusban is, és hogy lehetséges-e a korrekció. Ugyanakkor több adat, pl. az életkor vagy a nemi hovatartozás feltüntetése hatással van a magánélethez való jogra és az adatvédelemre. Különösen a diszkrimináció lehetséges eseteiben gyakran szükség lenne a személyes adatok különleges kategóriáinak, például a faji vagy etnikai származást feltáró adatoknak a felhasználására, amelyek számos adatvédelmi rendszerben fokozottan védettek.⁴⁰ Ezért e tekintetben mérlegelni kell a magánélet és az adatvédelem és a méltányosság közötti egyensúlyt. Az átláthatóság és a pontosság mérlegelésekor egy másik kompromisszum is létrejöhét. Lehetséges, hogy egyes algoritmusok magasabb pontszámot érnek el, mint a hasonló alternatívák, de olyan bonyolult modelleken alapulnak, hogy azok az emberek számára nem érthetőek. A számítástechnikai közösségben egyre inkább tudatosul, hogy a döntések nem csak a meglévő technológiák használata során születnek, hanem a kutatás és fejlesztés során is. Ugyanúgy, ahogyan a magánélet védelmét javító technológiák is megjelentek, új közösségek alakultak, amelyek a mesterséges intelligencia bizonyos irányokba történő fejlesztésére irányuló kutatásokat folytatnak. Ilyen például az ACM Fairness, Accountability, and Transparency (ACM FAT) konferencia, amely kifejezetten a társadalmi-technikai rendszerek igazságosságával, elszámoltathatóságával és átláthatóságával kapcsolatos új kutatásokkal foglalkozik. Az AI-panelekkel foglalkozó hasonló konferenciák vagy sávok azt mutatják, hogy a kutatás és a fejlesztés hogyan irányulhat kifejezetten bizonyos célok felé. Ismét elmondható, hogy a technológiák létrehozásakor vagy fejlesztésénél is van választási lehetőség. Ebben az esetben ezeket a döntéseket a kutatók gyakorolhatják, de a finanszírozó ügynökségek is befolyásolhatják. A választás eleme gyakran különböző szakaszokban van jelen.

39 Haferkamp (2017).

40 Lásd például a következő cikket: Art. 9 GDPR.

1. A tervezhetőség elve

A tudósok és intézmények a mesterséges intelligencia kontextusában a beépített demokrácia szerepeltetésére szólítottak fel.⁴¹ Az értékérzékeny tervezés gondolatával összhangban a demokratikus értékeket be kell vonni a tervezési folyamatba. Nemcsak a tervezési döntéseket kell demokratikus módon meghozni, hanem magának az alkalmazás működésének is demokratikusnak kell lennie. Ez az általános elképzelés azonban számos nehézségbe ütközik. Az egyik probléma az, hogy a demokráciának különböző fogalmai vannak, és ezek az alkalmazás tervezésénél egészen eltérő módon érvényesülhetnek.⁴² A demokratikus legitimitás különböző formáit úgy lehet strukturálni, hogy input-, output- és folyamat-legitimitásra osztsuk őket. A technikai követelmények igen kifinomultak lehetnek. Attól függően, hogy milyen kontextusban használják a mesterséges intelligencia alkalmazást, a demokrácia nagyon különböző szereplőket is bevonhat: az intelligens városok kontextusában a demokratikus döntések gyakran a városi lakosság döntéshozatalát vagy részvételét igénylik. Nemzeti környezetben inkább a parlament bevonásáról lesz szó a döntésekbe. Ezen okok miatt a demokrácia tervezési úton történő érvényesítése sok bizonytalanságot jelent a fejlesztők számára. Technikai szempontból egy olyan elvre lenne szükség, amelyet a fejlesztők meg tudnak ragadni, és amely támogatja a demokratikus értékeket a tervezési folyamatokban anélkül, hogy a demokrácia bizonyos felfogását előítéletekkel terhelné.

Javaslatom e kihívás kezelésére a tervezhetőség tervezési elvének megfogalmazása lenne. A tervezhetőség elve arra irányul, hogy az általános demokratikus értékeket általános és működőképes módon ültessük át a tervezésbe. Az elvnek legalább két szintre kellene kiterjednie, amelyekkel a fejlesztőknek foglalkozniuk kell: Az első szint a rendszer változtathatósága. A második szint a rendszer érthetősége. A demokrácia különböző elképzelései azon az elképzelésen alapulnak, hogy nyitottak és flexibilisek a változások különböző formáira: kormányváltásokra, tájékozott diskurzus utáni véleményváltozásokra és így tovább. Ez különösen igaz akkor, ha bizonytalanság van azzal kapcsolatban, hogy egy döntés hogyan érvényesül a gyakorlatban. Ilyen helyzetben a változtathatóság a demokratikus részvétel feltétele. Az ilyen változtathatóságot azonban a tervezéssel fokozni kell. Ez történhet egy meghatározott architektúra kiválasztásával vagy meghatározott módszerek alkalmazásával. Tekintettel arra, hogy a gépi tanulás magában foglalja az alkalmazkodás lehetőségét, definíció szerint változtatható. A tervezhetőség másik szintje a rendszer érthetősége. Az intelligenciát az informatikában nem az általános értelemben használják, azaz nem az adott rendszer cselekvései mögött álló logika megértésének lehetőségét. Az érthetőséget demokratikusan kell kialakítani. Általános cél lehet itt az, hogy egy rendszer érthető legyen minden olyan ember számára, akit a rendszer tevékenységei érintenek. Bár valójában nem mindenki fog dönteni arról, hogy az adott mesterséges intelligencia-rendszert használja-e és hogyan, az ideális az lenne, ha mindenkinek lehetősége lenne rá. Ez a szabvány az in-

- 41 Lásd például Nemitz (2018); Die Bundesregierung (2018: 4433.); High Level Expert Group on Artificial Intelligence (19).
- 42 Bozdag/van den Hoven (2015).

A csak egy adott csoportot célzó rendszerek esetében a támogathatóság meglehetősen szűk körű lehet. Ezzel szemben az általánosan alkalmazható mesterséges intelligencia rendszereknek meg kell felelniük az érthetőség általános követelményeinek. Ezért a demokratikus érthetőség szintje illeszkedik az átláthatóságról szóló jelenlegi diskurzusokhoz. A tervezhetőséggel összefüggésben azonban az érthetőség nem korlátozódik a rendszer által végzett konkrét cselekvésekre vagy döntésekre. A rendszer által érintett embereknek meg kell érteniük a rendszert és az annak alapjául szolgáló döntéseket. Tudniuk kell, hogy a rendszer megváltoztatható-e és hogyan. Mint minden tervezési elv, a tervezhetőséget is aligha lehet teljes mértékben megvalósítani. De a helyes irányba terelheti a fejlesztőket. Míg az érthetőség a demokratikus mérlegelés lehetőségére mutat rá, addig a változtathatóság szintje a változtatás lehetőségét jelzi, és lehetőségeket nyit a műszaki artefaktum hatékony irányítására.

B. Szociális réteg

A mesterséges intelligenciát nem csak technikai szinten tervezik, a mesterséges intelligencia-rendszereket körülvevő számos társadalmi konstrukció döntő szerepet játszik.⁴³ Ezek a társadalmi konstrukciók nem elkerülhetetlenek, hanem a technológiát és a társadalmat egyszerre alakító döntések és feltételezések szövedékei. A jog egy olyan mechanizmus, amely a társadalmi-technikai döntéseket demokratikus meghatározás alá vonhatja.

1. A hatások megértése

Fontos értékelni a technológia társadalmi hatását, de azt is meg kell érteni, hogy az ilyen hatások elismerése maga is társadalmi konstrukció. A közelmúltban különböző módszereket javasoltak a mesterséges intelligencia hatásainak értékelésére.⁴⁴ A hatásvizsgálat előfeltétele a technológiai szintű döntések feltárásának. Néha a megfelelő döntések csak akkor válnak nyilvánvalóvá és érthetővé, ha ismerjük a társadalmi hatásokat. A mesterséges intelligencia méltányosságáról szóló vita akkor indult el, amikor több kutató kritizálta az algoritmikus rendszerek diszkriminatív hatásait. Ugyanez igaz az átláthatóságra is. A technológiák következményeinek megismerése még azelőtt, hogy a károk és sérelmek bekövetkeznének, korántsem könnyű. Ahogy a technika története mutatja, a technológiák következményeinek ismerete gyakran túl későn érkezik. A sugárzás felfedezése sokatmondó és szomorú példa erre, mivel a technológiát felfedező tudósok közül sokan nem tudtak a technológia veszélyes hatásairól, és később rákban haltak meg. Eltartott egy ideig, amíg megértették a hatásokat. Sok más esetben a technológia és a hatás közötti ok-okozati összefüggés nem volt ennyire nyilvánvaló vagy vitatott. Ezekben az esetekben a jognak mélyreható hatása van a technológia társadalmi konstrukciójára.

43 Stamper (1988).

44 Reisman/Schultz/Crawford/Whithattaker (2018); ECP (2018).

Először is, az emberi jogi jog konszenzust teremthet abban, hogy egy bizonyos megfontolás védelemre érdemes. Ahhoz, hogy tudjuk, mi minősül hatásnak, meg kell alkotni egy védendő értéket. A jog konszenzust teremthet arról, hogy mi is ez. A magánélethez való jog jó példa egy olyan jogra, amelyet más jogintézményekből evolúciós módon levezetve találtak ki.⁴⁵ Amint létrejön a megállapodás arról, hogy mi az, ami emberi jogként védendő, egy különleges védelem lép életbe. Amint azt már bemutattuk, ez a védelem magával vonja az emberi jogokat érintő döntések demokratikus indoklásának szükségességét. A jog másik fontos jellemzője, hogy képes a hatásokat holisztikus módon felismerni és kiegyensúlyozni. A hatások nem eleve negatívak. Lehetnek előnyösek is. Miközben fontos, hogy kritikusán viszonyuljunk az új fejleményekhez, és megértsük az új veszélyeket és hátrányokat, ugyanilyen fontos, hogy értékeljük az előnyöket és a potenciális lehetőségeket. A technológiák hatásainak értékeléséhez elengedhetetlen, hogy az összes jövőbeli lehetőséget szem előtt tartsuk. Ez emberi jogi szempontból is igaz. Amint azt a fentiekben bemutattuk, a technológiák az emberi jogok előmozdítására is alkalmasak. Ezért a következményeket egymással szemben kell mérlegelni. Az ilyen helyzetek jogi eljárásokban történő értékelésére számos joghatóság kidolgozott egy arányossági tesztet.⁴⁶ Ez egy praktikus módja annak, hogy egy intézkedést holisztikusan értékeljünk, és az érvelést úgy strukturáljuk, hogy számos megfontolásra és azok egymás elleni mérlegelésére legyen lehetőség. Emellett gyakorlati következtetésekhez jut, amelyeket közölnék a döntések által érintettekkel. Az arányosság elve valójában lehetővé teszi a társadalmi-technikai értékelést különböző szinteken.

2. A mesterséges intelligencia tervezése társadalmi konstrukciókon keresztül

Van azonban egy még szélesebb értelemben vett társadalmi konstruáltsága a mesterséges intelligencia hatásainak. Ez a mesterséges intelligencia-rendszerek hatásának nagy részére vonatkozik. Különösen az adatelemzés esetében lehetnek különböző célok és célkitűzések: bizonyos összefüggések feltárása, bizonyos cselekvések valószínűségeinek feltárása, vagy éppen annak valószínűségének megmutatása, hogy bizonyos alternatív cselekvések hogyan alakulhatnak.⁴⁷ Bár igaz, hogy ezeknek a rendszereknek lehetnek mélyreható normatív hatásai, ezek a hatások gyakran a rendszer társadalmi felépítéséből erednek, ahelyett, hogy tévesen a technológiában rejlenek minősítenénk őket. Míg például a nagy adatelemző eszközök bizonyos valószínűségeket számolnak ki, e valószínűségek jelentését és az általuk játszott szerepet aktívan konstruálják.⁴⁸ Egy szemléletes példa a pontszámok visszaélészerű felhasználása a hitelesítésre...

45 Lásd az egyik korai érvet a Warren/Brandeis (1890) című könyvben.

46 Klatt/Meister (2012).

47 Ennek alapján megkülönböztetünk diszkrét, preszkrét és leíró analitikát. Hoffmann-Riem (2017).

48 Lásd például Schlaudt (2018).

itworthiness mint megbízhatósági pontszám az alkalmazottak számára.⁴⁹ Nyilvánvaló, hogy egy olyan rendszer, amelyet arra terveztek, hogy kiszámítsa egy személy adósságtörlesztési valószínűségét, nem arra készült, hogy értékelje az adott személy megbízhatóságát, amikor a munkáról van szó. Mégis, az a döntés, hogy a rendszert más kontextusban használják, semmiképpen sem olyan döntés, amelynek köze van a rendszer kialakításához. Ez inkább egy társadalmi választás egy másik társadalmi kontextusba való átvitelre.

Ugyanez igaz bizonyos valószínűségek használatára is. A törvény sok esetben megmutatja, hogy a valószínűségeknek a különböző kon- textusokban teljesen eltérő jelentésük van. A rendőrségi és biztonsági jogban is vannak különböző valószínűségi követelmények, amelyek társadalmi szempontból fogalmazódnak meg. Az emberi jogokat kevésbé érintő intézkedéseknek alacsonyabb valószínűségi küszöbnek kell megfelelniük, míg a nagyobb potenciális hatásokkal járó intézkedéseknek magasabb valószínűségi követelményeknek kell megfelelniük. Aktív választás és demokratikus döntés, hogy egy adott hatósági hatáskört egy bizonyos valószínűséghez kötünk.

A mesterséges intelligencia rendszerek kimeneteinek jelentését számos módon lehet megkonstruálni. A törvény nem csupán explicitte teszi ezt a jelentést, hanem megnyitja a technológia társadalmi konstruálását a demokratikus tanácskozás és a demokratikus döntéshozatal előtt. A mesterséges intelligencia-rendszerek kimenetei illegálissá és irrelevánssá tehetők. Emberi felügyelet és emberi döntéshozatal alá vonhatók. Sőt, a törvény erejével is felruházzhatók. A német jogban a köztisztviselők értékelése, a döntések nem alapulhatnak bizonyos személyiségjegyek teljesen automatizált értékelésén.⁵⁰ A fent említett Art. 22 GDPR rendelkezik az emberi felügyelet jogáról, és a teljesen automatizált döntéseket emberi döntésekhez köti. Ugyanakkor vannak olyan rendelkezések, amelyek egyértelművé teszik, hogy a teljesen automatizált döntéseknek joghatása van. Vegyük például a szövetségi közigazgatási eljárási törvénykönyv 35a. §-át. A rendelkezés kimondja: "A közigazgatási aktus teljes egészében elfogadható automatikus rendszerekkel, feltéve, hogy azt jogszabály engedélyezi, és nincs mérlegelési vagy mérlegelési lehetőség." Ez a rendelkezés egyértelművé teszi, hogy létezhetnek teljesen automatizált közigazgatási aktusok, azaz meghatározott személyek vagy csoportok számára jogerővel rendelkező határozatok. Ez alapvetően azt jelenti, hogy ezek a rendszerek jogerős és végrehajtható határozatokat hozhatnak. Két példa az ilyen döntésekre az intelligens közlekedési rendszerek, amelyek automatikusan sebességkorlátozásokat állítanak be, vagy előzési tilalmat rendelnek el, ha az időjárás vagy a forgalom miatt veszély fenyegeti a járművezetőket. Egy másik példa a teljesen automatizált sebességtúllépési bírságok, amelyeket a megfelelő értesítéseket automatikusan elküldő detektívrendszerek állítanak ki.

290 Christian Djeffal

49 O'Neil (2016: 147-149).

50 Lásd a német szövetségi köztisztviselőkről 4szóló törvény paragrafusát. I 14

3. A mesterséges intelligencia mint szokásjog

A mesterséges intelligencia rendszereknek lehetnek valós hatásai, amelyek nagymértékben függenek attól a társadalmi konstrukciótól, amely ezeket a következményeket a rendszernek tulajdonítja. Ez elvezet ahhoz a kérdéshez, hogy milyen követelményeket kell támasztani az ilyen elfogadáshoz. Ezzel a kérdéssel jelenleg a számítógépes társadalmi választás területe foglalkozik.⁵¹ A mesterséges intelligencia tervezésének folyamatában rejlt rejtett erkölcsi döntések az egyik fő motivációja annak, hogy a társadalmi választás és az informatika összefonódásával foglalkozzunk. A számítógépes társadalmi választás hívei tehát megpróbálnak kritériumokat találni a mesterséges intelligencia-rendszerek legitím módon történő tervezéséhez. A gépi tanulás egyik szembetűnő jellemzője, hogy valójában olyan adatokon alapul, amelyeket gyakran azok állítanak elő, akikre a rendszer vonatkozik. A kutatási projektek például megkérdéseket és szimulációkat használtak annak érdekében, hogy felhasználói adatokat szerezzenek arról, hogy az automatizált autóknak hogyan kellene reagálniuk bizonyos helyzetekben.⁵² Ennek az etikus tervezési fókuszának a demokratikus szemlélete azonban bizonyos problémákat tár fel: Az első probléma az, hogy a különböző feltevések egészen eltérő eredményekhez vezethetnek, amelyek mindegyike igényt tarthat arra, hogy etikus legyen. A különböző etikai elméletek akár ellentétes eredményeket is hozhatnak. Vegyük például az utilitarizmust és az elvi etikát. Míg bizonyos, egy személy számára káros, de a többség számára előnyös cselekedetek az utilitarista nézőpontból etikusnak tekinthetők, addig a principikus nézőpontból etikátlannak tekinthetők. Végül soron sok alternatíva közül kell választani. Ha azt állítjuk, hogy csak egyetlen helyes és erkölcsös megoldás létezik, amelyet minden más megoldással szemben előnyben kell részesíteni, azzal minden más lehetséges megoldást diszkriminálunk. Ez figyelmen kívül hagyja a különböző megközelítéseket és az egyetlen kérdésre adott különböző megoldásokat. Egy ilyen helyzetben nem marad hely a választásnak. Egy másik kérdés, hogy a mesterséges ágensek valóban képesek-e morális döntéseket hozni, vagy csak szimulálják azokat. Erkölcsi szempontból a tényleges ítélőképesség kérdése a legfontosabb. Ez a probléma ahhoz a kérdéshez kapcsolódik, hogy a gépek valóban képesek-e gondolkodni, amely Turingtól Searle-ig vitatott reflexiót váltott ki.⁵³

E szakasz alapvető érve az, hogy a számítógépes választási teoretikusoknak erkölcsi helyett jogi fogalmakban kellene gondolkodniuk. Kantra építve a külső hatásokkal járó cselekedeteket a jognak tulajdoníthatnánk, míg a belső hatású kérdések az etika területére tartoznak. A mesterséges intelligencia-rendszereknek gyakran mélyreható normatív hatásai vannak. Míg a legtöbb etikai megfontolás a kimeneti legitimitásra összpontosít, a komputáció és a jog oly módon egyesíthető, hogy a demokratikus bemeneti legitimitást jogi eszközökkel érjük el. A gépi tanulási alkalmazásokat általában olyan adatokkal képzik ki, amelyek bizonyos szereplők viselkedését tükrözik. Bár nincs általános formalizált szabály arra vonatkozóan, hogy mi a jelentősége az ilyen gyakorlatnak, azt az érvet szeretném felhozni, hogy

51 Brandt/Conitzer/Endriss/Lang/Procaccia (2016) A mesterséges intelligenciával kapcsolatos szakirodalom áttekintését Prasad (2019) adja.

52 Awad/Dsouza/Kim/Schulz/Henrich/Shariff/Bonnefon/Rahwan (2018).

53 Turing (1950); Searle (1980).

hogy a gépi tanulás - bizonyos feltételek mellett - szokásjognak tekinthető. Ez rávilágítana azokra a számítástechnikai és társadalmi választási lehetőségekre, amelyek lehetővé teszik a demokratikus kifejezést egy mesterséges intelligencia rendszeren keresztül. Bizonyos jogalkotási gyakorlatok analógiájára építve lehetséges lenne a mesterséges intelligenciával mint a demokratikus döntések médiumával szemben támasztott követelmények megfogalmazása.

A szokásjog korábban nagyon fontos szerepet játszott egyes közösségek irányításában, amelyek bizonyos gyakorlatokat kötelezőnek tekintettek. A modern társadalmak növekvő összetettsége és az új nyomdatechnológiák lehetőségei miatt a szokásjog sokat veszített jelentőségéből. Főként az idővel kialakult kisebb közösségek íratlan gyakorlatára támaszkodott. Míg a common law-országok bíróságai továbbra is az egykor kialakult elvekre támaszkodtak, és azokat olyan érvekre alakították át, amelyekre az igazságszolgáltatás építhetett, addig az egyik olyan jogrendszer, amelyben a szokásjog megőrizte jelentőségét, a nemzetközi jog. A nemzetközi jogban még mindig van egy kezelhető számú résztvevő, akiknek a gyakorlata szokásnak minősíthető. A digitalizáció számos trendje új tudásdimenziót segít, amely a szokásjog újjáélesztéséhez vezethet különböző területeken. Először is, az adatosítás új utakat nyit az egyes szereplők viselkedésének tárolására és megértésére. A big data azt jelenti, hogy hatalmas mennyiségű adatot lehet tárolni és elemezni. Másodszor, az olyan trendek, mint a dolgok internete lehetővé teszik az adatok folyamatos, automatizált és mindenütt jelenlévő gyűjtését. A dolgok internete a különböző emberi környezetekben található, hálózatba kapcsolt eszközök trendjét jelenti. A mesterséges intelligencia-technológiák segíthetnek az adatok elemzésében és megértésében oly módon, hogy a gyakorlat érthetővé és érthetővé váljon. Ezek a technológiák együttesen láthatóvá teszik az emberek tényleges gyakorlatát.

Azonban továbbra is kérdéses, hogy ez a szokás a generalizálták az emberi cserében. A számítógépes társadalmi választás tudósai elgondolkodtak ezen a kérdésen, és olyan kritériumokat állítottak fel, amelyeket figyelembe kellett venni a gyakorlatot reprezentáló mesterséges intelligencia megalkotásának folyamatában. Baum például három általános kritériummal dolgozott:

1. **Állva:** Ki vagy mi tartozik a csoportba, hogy értékeit figyelembe vegyék a mesterséges intelligenciában?
2. **Mérés:** Milyen eljárást alkalmaznak a kiválasztott csoport minden egyes tagjának értékeinek meghatározására?
3. **Összevonás:** Hogyan kombinálják az egyes csoporttagok értékeit az összesített csoportértékek kialakításához?⁵⁴

A szokásjog követelményei bizonyos értelemben kiegészítik, bizonyos értelemben eltérnek a fenti kérdésektől. A formális kritériumok kialakulásának

a szokásjog egy gyakorlat (*consuetudo*) és az a meggyőződés, hogy ez a gyakorlat jognak tekintendő (*opinio iuris sive necessitatis*). A gyakorlatnak következetesnek és általánosnak kell lennie, még ha ez nem is jelenti azt, hogy a gyakorlat egységes és univerzális.⁵⁵ A szokásjoggal összefüggésben az általános gyakorlattal kapcsolatos legfontosabb kérdés az, hogy van-e elegendő képviselő. Ennek oka, hogy egyes szereplők hallgatólagosak maradnak, és nem vesznek részt a gyakorlatban. A második kritérium az úgynevezett *opinio iuris*. Ez azt a meggyőződést jelenti, hogy az adott gyakorlat az adott módon való cselekvés jogi kötelezettségén alapul. Ez a kritérium tulajdonképpen legitimálja a gyakorlat normatív erejét. Az *opinio iuris* kritériumának teljesítése érdekében az érintetteknek az adatokat azzal a céllal kell a tudomásukra hozniuk, hogy befolyásolják az ezen adatok alapján eljáró rendszert. Ez a kritérium a felhasználók szuverén döntésétől teszi függővé a mesterséges intelligencia rendszer legitimitását. A rendszer egyszerűen megtanulja, hogy mi az emberek gyakorlata. Megtanulja, hogy az érintettek mit akarnak, hogy ez a gyakorlat az legyen. Ebben a környezetben az információs önrendelkezés nem csupán a személyes adatok hatalmát jelenti, hanem a hatalom tudatos gyakorlását az adatokon keresztül. Az adatalany nem olyan erőforrás, amelyből a személyes adatokat kivonják. Ebben a környezetben az adatok előállítása olyan demokratikus aktussá válik, mint a szavazás.

C. Irányítási réteg

A mesterséges intelligencia demokráciára gyakorolt hatásának elemzéséhez nem elég, ha csak bizonyos rendszereket vizsgálunk. Makroszintű elemzésre van szükség, amely a technológiákra vagy akár az AI egészére összpontosít. Ezt nevezzük itt kormányzási rétegnek.

1. Keretezés

A mesterséges intelligencia demokratikus irányítását befolyásolja az, hogy milyen módon keretezik a mesterséges intelligenciát. A mesterséges intelligenciát rendszeresen meghatározott kontextusba helyezik, vagy egy bizonyos módon látják. Gyakran beszélnek a tudósok a mesterséges intelligencia etikájáról, egy ⁵⁶másik irányzat pedig a mesterséges intelligenciáról és az emberi jogokról. Míg a tudósok egy-egy kereten belül vitatkoznak és elemeznek, viszonylag kevés vita folyik a keretek közötti választásról. Pedig a kereteknek jelentős hatásuk van. Vegyük például az etikai és a politikai keret közötti választást.⁵⁷ A keretek a technológiáról való gondolkodás teljesen különböző módozataihoz vezetnek. Hasonlítsuk össze az összejtemelést és az 5G hálózati infrastruktúra létrehozását. Az összejtervezést túlnyomórészt etikai kérdésként értelmezik, míg az utóbbit általában politikai kérdésként fogják fel. Természetesen az összejtkutatással összefüggésben számos olyan kérdés van, amelyet politikai kérdésként fognánk fel, és az 5G infrastruktúra kiépítése során is számos etikai kérdés merülhet fel. Konstruktivista

55 Crawford (2012: 23ff.).

56 Mittelstadt/Allo/Taddeo/Wachter/Floridi (2016).

57 Erről az elmékedésről lásd Djeffal (2019).

a tudósok kiemelték, hogy a keretek és elméletek befolyásolják a tudományos vizsgálat tárgyát. Ezért aktív választás, hogy a mesterséges intelligenciát bizonyos kontextusba helyezzük, és a mesterséges intelligencia etikáját vagy politikáját vizsgáljuk, vagy a mesterséges intelligencia és az emberi jogok kapcsolatát. Ez a választás szükségszerűen tartalmaz bizonyos preferenciákat, amelyek eredendően vagy az adaptált keretből következnek. Minden keret bizonyos vakfoltokat is előidéz. Egyes szempontok láthatatlanná válnak.

Az egyik kísérlet a mesterséges intelligencia társadalomra gyakorolt hatásának általános leírására az "algokrácia" fogalma. Ez a kifejezés a kormányzás más formáit, például a demokráciát vagy a monarchiát állítja szembe egy olyan rendszerrel, amelyben a hatalmat (egyre inkább) automatizált rendszerek gyakorolják.⁵⁸ Az algokrácia kifejezést többnyire kritikusan használják.⁵⁹ Rávilágít arra, hogy az algoritmusok egyre fontosabbá válnak a kormányzás kérdéseiben. Ahelyett, hogy hozzáadnám a témával kapcsolatos egyre növekvő szakirodalomhoz, az algokrácia konstruktivista jellegét szeretném kiemelni. Ez elvezet ahhoz a kérdéshez, hogy mit emel ki ez a kifejezés, és mi marad ki a képből. Az aktorhálózat-elmélet (ANT) alapvető meglátásaira építve azt állítom, hogy az algokrácia kerete hajlamos elmosódni és elrejteti az emberi cselekvőképességet. Az algokrácia kiemeli a gépi hatalmat, de figyelmen kívül hagyja, hogy az emberek hogyan befolyásolják az automatizált akciókat. Az ANT egyik alapvető érve, hogy figyelmen kívül kell hagyni az alanyok és a tárgyak közötti különbséget, és a technológiát a társadalom részeként kell értékelni, egy olyan hálózatban, amelyben emberi szereplők használják.⁶⁰ Ez az elemzés lehetővé tette az ANT hívei számára, hogy feltárják a technikai artefaktumok cselekvőképességét. Az én alapvető érvem az, hogy ezt az elméletet ma már fordítva is lehetne használni, hogy a gépi cselekvés helyett az emberi cselekvést tárjuk fel. Az algokrácia elmélete a mesterséges intelligencia diszkurzusának kritikus részét képviseli, amely a mesterséges intelligenciát kifejezetten automatizált döntéshozó rendszerekként keretezi, és a megnövekedett hatalmukat vizsgálja. Az ilyen rendszerek növekvő képességére és hatalmára való összpontosítással néha elfelejtkezünk arról, hogy hogyan használják ezeket a rendszereket, és hogyan fonódnak össze az emberi cselekvőképességgel. Mint fentebb vázoltuk, a társadalmi környezet sokféleképpen meghatározza a mesterséges intelligencia alkalmazások tervezését. Sok esetben a jog is része a mesterséges intelligenciát normatív erővel felruházó konstruktív erőfeszítéseknek. Az algokráciát kiegészítő keret nem kizárólag azt a tényt vizsgálná, hogy egyre több döntést delegálnak, hanem azt is, *hogyan* delegálják, és hogy *ki* ellenőrzi és befolyásolja az automatizált rendszereket. Ahogy az ANT számos támogatója érvelt, a hangsúly nem a szereplők egyetlen osztályára, hanem inkább a szereplők egymáshoz való viszonyára helyeződne.

58 Yeung (2018).

59 Danaher (2016).

60 Latour (2000: 180).

2. Szervezeti szempontok

A mesterséges intelligencia fejlesztését és alkalmazását szervezeti intézkedésekkel is lehet befolyásolni. A közelmúltbeli mesterséges intelligencia stratégiák közül sok tartalmaz ilyen intézkedéseket. Egyrészt a szervezeti változások célja a mesterséges intelligencia területén a technológiai fejlődés fokozása. Új intézményeket alapítanak, akár a kutatásban és fejlesztésben való közvetlen részvételre, akár az ilyen tevékenységek finanszírozására, akár a már meglévő szervezetek hálózatának bővítésére. Az Egyesült Arab Emírségek a mesterséges intelligenciáért felelős miniszterrel került a címlapokra⁶¹, a német kormány pedig nemrég alapított egy "innovációs ugrásokért" felelős ügynökséget, amelynek feladata az úttörő innovációk kutatásának és fejlesztésének finanszírozása és a megvalósítás fokozása. Másrészt az újonnan alapított szervezetek is felügyeletet gyakorolnak a mesterséges intelligencia rendszerek felett. Valójában már most is számos, ezzel a feladattal felruházott szervezet létezik. Az olyan szervezetek, mint az Egyesült Államok Szövetségi Gyógyszerügyi Hivatala, vagy annak európai és máshol működő megfelelői, részt vesznek az orvosi termékeknek minősülő mesterséges intelligencia rendszerek tanúsításában. Több felügyeleti intézményt is követelnek.⁶² A kanadai példák nyomán néhány államban megalakultak a mesterséges intelligencia megfigyelőközpontok, amelyek célja, hogy feltárják a mesterséges intelligencia társadalmi következményeit. A munka jövője az egyik olyan kérdés, amellyel gyakran foglalkoznak ebben az összefüggésben.⁶³

A szervezeti változás nem mindig csak új szervezetekben nyilvánul meg. Néha a szervezetek belülről változnak az új feladatokhoz való alkalmazkodással. Az egyik fontos fejlemény ebben a tekintetben az a kérdés, hogy szükség van-e új munkaköri profilra a szervezeteken belül. Az adattudósok az egyik olyan profil, amely jelenleg felemelkedőben van. Egyesek mégis úgy vélik, hogy az algoritmusművészek teljesen új profiljára lehet szükség.⁶⁴ Ennek lényege, hogy legyenek speciális technikai készségekkel rendelkező emberek, hogy egy szervezet megőrizze az ügynökséget, amikor mesterséges intelligencia rendszerekkel kell foglalkoznia. Ennek az elképzelésnek az az érdekes aspektusa, hogy a szakértelem olyan szervezetek számára is elérhetővé válna, amelyek eddig nem kapcsolódtak technológiai szakértelemhez. Az algoritmus-szakértő munkaköri profilja demokratizálhatja az algoritmusokkal kapcsolatos kérdésekkel kapcsolatos ügynöki tevékenységet. A mesterséges intelligencia-rendszerekkel kapcsolatos tudás általánosan elérhetővé válna. Ettől a konkrét profiltól független kérdés lenne az egyes mesterséges intelligenciával kapcsolatos kérdéseken dolgozó csoportok interdiszciplináris összetétele. Ha a mesterséges intelligenciát konkrét kontextusokban használják, akkor lehet, hogy a szerepek és nézőpontok minimális követelménye a jelenlét. Ezért az AI-rendszereket fejlesztő, használó vagy értékelő szervezeteknek el kell gondolkodniuk azon, hogy mi lenne e csapatok megfelelő összetétele. Bár az informatikusok szükséges alkotóelemei az ilyen csapatoknak, soha nem elegendők. Mindent egybevetve, a szervezeti

kihívások és változások nagyon jó példája annak, hogy az algo-

61 Tendersinfo (2017).

62 Tuttt (2017).

63 Lásd például Die Bundesregierung (2018: 26).

64 Mayer-Schönberger/Cukier (2013: 189-192); Hill (2015: 284).

rithmusok hatással vannak társadalmi környezetükre, és hogyan járulhatnak hozzá hatékonyan a mesterséges intelligencia rendszerek társadalmi-technikai kontextusának változásai a megfelelő kormányzáshoz.

III. Következtetések

A német szövetségi állam egyik egységének, Brémának az alkotmánya 1947 tartalmaz egy nagyon érdekes rendelkezést az ember és a gép kapcsolatáról. Az alkotmány a következő cikkelyt tartalmazza. 12. szakasza szerint: "Az ember magasabb rendű, mint a gép és a technika". Ez a rendelkezés az iparosodás folyamatának tapasztalataival foglalkozik, amelynek során a gépek, a technológiák és a termelés új lehetőségei egyre nagyobb jelentőségre tettek szert. Érdekes, hogy az alkotmány alapítói szükségét érezték annak, hogy emlékeztessék az embereket és a hatalmon lévőket arra, hogy az embernek magasabb rendűnek kell lennie. Az iparosodás során ez nem foglalkozott a gépek egyre növekvő képességeivel, hogy olyan intelligensen cselekedjenek, hogy akár személynek is tekinthetők. Inkább arról volt szó, hogy mint termelési kapacitásoknak, olyan nagy jelentőséget tulajdonítottak nekik. Az alapgondolat tehát az volt, hogy a technikai artefaktumok hatalmas társadalmi és gazdasági jelentősége ellenére is emberközpontú szemlélet mellett érveljenek. Ez az alapgondolat lefordítható a digitalizáció folyamatára is, amelyben a gépek olyan problémák megoldásában vesznek részt, amelyek az intelligencia olyan fokát igénylik, amelyet korábban kizárólag az ember számára fenntartottnak tekintettek. Az ember e normatív centrikusságának egyik aspektusa az emberi jogok kizárólagos hordozói státusza. Ugyanilyen fontos szempont az emberek tényleges önrendelkezésének szempontja a technológiák növekvő lehetőségeivel szemben. A magasabb rangúság nem csak azt jelenti, hogy az embereket nem érheti kár az új technológiai lehetőségek miatt. Azt is jelenti, hogy az embereknek a vezetőülésben kell ülniük. Ez úgy is felfogható, mint a különböző szinteken megvalósuló hatékony önrendelkezésre való felhívás.

Ha a mesterséges intelligencia továbbra is megfelel a magas elvárásoknak, és folyamatos hatást gyakorol a

társadalmi fejlődés, még fontosabb lesz a mindenre kiterjedő értékérzékeny fejlődés szempontjából. A brémai alkotmány szempontjából az egyik szükséges elem a mesterséges intelligencia demokratizálásának átgondolása lenne. Ehhez döntő fontosságú lesz, hogy a mesterséges intelligenciát olyan általános célú technológiák összességeként értelmezzük, amelyek nagyon különböző körülmények között és nagyon különböző módon használhatók többféle feladat elvégzésére. Miközben fontos megérteni, hogy a mesterséges intelligencia jelenleg hol fenyegeti a demokráciát, ugyanilyen fontos értékelnünk a benne rejlő lehetőségeket is. A technológia használatának és lehetőségeinek nyitottságának megértése lehetővé teszi számunkra, hogy eldöntsük, továbbfejlesztjük-e a technológiát, és milyen utat válasszunk. A mesterséges

intelligencia demokratizálásával kapcsolatban néhány általános igazság érvényes a demokráciáról: A demokrácia egy folyamat, nem pedig egy elérhető eredmény. Nagyon könnyen elveszíthető, és mindenkinek folyamatosan dolgoznia kell érte az út során. Ha egyszer abbahagyjuk az érte való törekvést, eltűnik. Ebből a szempontból a mesterséges intelligencia csak egy újabb kihívás, amelynek a

a társadalom közelebb kerülhet ahhoz az eszményhez, amely a brémai alkotmány cikkelye¹², valamint számos más demokratikus rendelkezés mögött áll: az embereket kell a közhatalom normatív középpontjába helyezni.

Hivatkozások

- Abel, Ralf B. (2018): "Automatisierte Entscheidungen im Einzelfall gem. Art. 22 DS-GVO. Anwendungsbereich und Grenzen im nicht-öffentlichen Bereich", in: Zeitschrift für Datenschutz 8, pp. 304-307.
- Awad, Edmond/Dsouza, Sohan/Kim, Richard/Schulz, Jonathan/Henrich, Joseph/Shariff, Azim/Bonneton, Jean-François/Rahwan, Iyad (2018): "The Moral Machine experiment", in: Nature pp563., 59-64.
- Baum, Seth D. (2017): "A biztonságos és társadalmilag hasznos mesterséges intelligencia előmozdításáról", in: AI & SOCIETY 32, pp. 543-551.
- Bodó, Balázs/Helberger, Natali/Vreese, Claes H. de (2017): "Political micro-targeting: a manchurian candidate or just a dark horse?", in: Internet Policy Review (IPR).
- Bozdog, Engin/van den Hoven, Jeroen (2015): "Breaking the filter bubble: democracy and design", in: Ethics and Information Technology pp17., 249-265.
- Brandt, Felix/Conitzer, Vincent/Endriss, Ulle et al. (Hg.) (2016): Handbook of computational social choice, Cambridge: Cambridge University Press.
- Ceron, Andrea/Curini, Luigi/Iacus, Stefano M. (2017): Politika és nagy adatok. Nowcasting és választások előrejelzése a közösségi médiával, London, New York: Routledge.
- Coleman, Stephen (2017): Cambridge, UK, Malden, MA: Polity.
- Crawford, James (2012): *Brownlie's Principles of Public International Law*, Oxford: Oxford Univ. Press.
- Danaher, John (2016): "Az algokrácia fenyegetése: Reality, Resistance and Accommodation", in: Philosophy & Technology pp29., 245-268.
- Der Standard (2019): Post löscht alle Informationen zu Parteipräferenzen, <https://derstandard.at/2000095874780/Post-loescht-alle-Informationen-zu-Partei-affinitaet>.
- Die Bundesregierung (2018): Eckpunkte der Bundesregierung für eine Strategie Künstliche Intelligenz, https://www.bmbf.de/files/180718%20Eckpunkte_KI-Strategie%20final%20Layout.pdf.
- Djeflal, Christian (2019): In: "Normatív iránymutatások a mesterséges intelligenciához", in: Thomas Wischmeyer/Timo Rademacher (Hg.), *Regulating Artificial Intelligence*, Wien, Berlin, New York: Springer, megjelenés alatt.

- ECP (2018): Artificial Intelligence Impact Assessment, <https://airecht.nl/s/Artificial-Intelligence-Impact-Assessment-English.pdf>.
- Ehm, Frithjof/Walter, Christian (Hg.) (2015): Nemzetközi demokrácia-dokumentumok. Szerződések és más eszközök gyűjteménye, Leiden, Boston: Brill Nijhoff. Eidgenössischer Datenschutzbeauftragter/Konferenz der schweizerischen Datenschutzbeauftragten (2018): Leitfaden der Datenschutzbehörden von Bund und Kantonen zur Anwendung des Datenschutzrechts auf die digitale Bearbeitung von Personendaten im Zusammenhang mit Wahlen und Abstimmungen in der Schweiz, <https://www.edoeb.admin.ch/dam/edoeb/de/dokumente/2018/Leitfaden%20Wahlen.pdf.download.pdf/Leitfaden%20Wahlen%20und%20Kampagnen%20final.pdf>.
- Engelbart, Doug (1963): "A Conceptual Framework for Augmentation of Mans Intellect", in: *Vistas in Information Handling* 1, pp. 1-29.
- Ennals, Richard (1987): In: "Socially useful artificial intelligence" (Társadalmilag hasznos mesterséges intelligencia): *AI & SOCIETY* 1, pp. 5-15.
- Gasser, Urs/Almeida, Virgilio A.F. (2017): "A Layered Model for AI Governance", in: *IEEE Internet Computing* 21, pp. 58-62.
- Goodfellow, Ian/Bengio, Yoshua/Courville, Aaron (2016): *Deep Learning*, Michigán: MIT Press.
- Greer, Steven C. (1997): Az emberi jogok európai egyezményének 8-11. cikkei alóli kivételek, [https://www.echr.coe.int/LibraryDocs/DG2/HRFILES/DG2-EN-HRFILES-15\(1997\).pdf](https://www.echr.coe.int/LibraryDocs/DG2/HRFILES/DG2-EN-HRFILES-15(1997).pdf).
- Haferkamp, Björn (2017): "Was ist optimum? Nutzen und Fallstricke der Optimierung", in: Björn Bergh (Hg.), *Big Data und E-Health*, Berlin: Erich Schmidt Verlag, pp. 59-68.
- Helbing, Dirk (2019): "Machine Intelligence: Blessing or Curse? It Depends on Us!", in: Dirk Helbing (Hg.), *Towards Digital Enlightenment. Esszék a digitális forradalom sötét és világos oldaláról*, Cham: Springer International Publishing, pp. 25-39.
- Mesterséges intelligenciával foglalkozó magas szintű szakértői csoport: <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>.
- Hill, Hermann (Hg.) (2015): *Auf dem Weg zum Digitalen Staat - auch ein besserer Staat? (= Verwaltungsressourcen und Verwaltungsstrukturen, Band 30)*, Baden-Baden: Nomos.
- Hoffmann-Riem, Wolfgang (2017): In: "Verhaltenssteuerung durch Algorithmen - Eine Herausforderung für das Recht", in: "Verhaltenssteuerung durch Algorithmen - Eine Herausforderung für das Recht": *AöR (Archiv des öffentlichen Rechts)* 142, pp. 1-42.
- Hofmann, Jeanette (2018): Digitalisierung und demokratischer Wandel als Spiegelbilder?", in: "Digitalisierung und demokratischer Wandel als Spiegelbilder?": Franziska Martinsen (Hg.), *Wissen - Macht - Meinung. Demo-*

- kratie und Digitalisierung die Hannah-Arendt-Tage20. Weilerswist2017,; Velbrück Wissenschaft, pp. 14-21.
- Hofstetter, Yvonne (2016): *Das Ende der Demokratie. Wie die künstliche Intelligenz die Politik übernimmt und uns entmündigt*, München: Bertelsmann.
- Irrgang, Bernhard (2002): *Technischer Fortschritt. Legitimitätsprobleme innovativer Technik (= Philosophie der Technik, Band 3)*, Paderborn: Schöningh.
- Klatt, Matthias/Meister, Moritz (2012): *Oxford: The Constitutional Structure of Proportionality*: Oxford Univ. Press.
- Latour, Bruno (2000): *Latour: Pandora reménye. Esszék a tudománytudományok valóságáról*, Cambridge, Mass.: Harvard University Press.
- Mainzer, Klaus (2019): *Künstliche Intelligenz - Wann übernehmen die Maschinen? (= Technik im Fokus)*, Berlin, Heidelberg: Springer.
- Martini, Mario (2018): "Art. 22", in: Boris Paal/Daniel Pauly (Hg.), [Duplikat] *Datenschutz Grundverordnung: DS-GVO*, München: C.H. Beck.
- Mayer-Schönberger, Viktor/Cukier, Kenneth (2013): *Big Data. Die Revolution, die unser Leben verändern wird*, München: Die Revolution, die unser Leben verändern wird: Redline.
- McCarthy, John (1989): Bloomfield által szerkesztett *The Question of Artificial Intelligence*", in: "Review of *The Question of Artificial Intelligence* edited by Brian Bloomfield": *Annals of the History of Computing*.
- McCarthy, John/Minsky, Marvin/Shannon, Claude (1955): *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html> vom 31.03.2017.
- Mittelstadt, Brent D./Allo, Patrick/Taddeo, Mariarosaria/Wachter, Sandra/Floridi, Luciano (2016): "Az algoritmusok etikája. A vita feltérképezése", in: *Big Data & Society* 3, 1-21.
- Nemitz, Paul (2018): in: "Alkotmányos demokrácia és technológia a mesterséges intelligencia korában": *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences* 376.
- O'Neil, Cathy (2016): *A matematikai tisztítás fegyverei. Hogyan növelik a nagy adatok az egyenlőtlenséget és fenyegetik a demokráciát*, New York: Crown.
- Pernice, Ingolf (2016): "E-Government and E-Democracy. A legitimációs deficitek leküzdése a digitális Európában", in: *A digitális Európa: HIIG Discussion Paper Series*.
- Prasad, Mahendra (2019): "Social Choice and the Value Alignment Problem", in: "Social Choice and the Value Alignment Problem": *Social Choice and the Value Alignment Problem: Roman V. Yampolskiy (Hg.)*, *Artificial intelligence safety and security*, Boca Raton: CRC Press, pp. 291-314.
- Reisman, Dillon/Schultz, Jason/Crawford, Kate/Whittaker, Meredith (2018): *Algoritmikus hatásvizsgálatok. Gyakorlati keretrendszer a közhatalom és az elszámoltathatóság számára. AI NOW*, <https://ainowinstitute.org/aiareport2018.pdf>.

- Rousseau, Peter L. (2009): "Általános célú technológiák", in: Steven Durlauf/L. Blume (Hg.), *Economic Growth*, London: Palgrave Macmillan UK, pp. 74-79.
- Russell, Stuart/Norvig, Peter/Kirchner, Frank (2012): *Künstliche Intelligenz. Ein moderner Ansatz*, München: Pearson Higher Education.

- Schlaudt, Oliver (2018): Die politischen Zahlen. Über Quantifizierung im Neoliberalismus (= Klostermann Rote Reihe, Band 102), Frankfurt am Main: Vittorio Klostermann.
- Schulz, Wolfgang/Dankert, Kevin (2016): A "dolgozók általi kormányzás" mint a jogi szabályozás kihívása, in: Internet Policy Review 5, x.
- Searle, John R. (1980): "Agyak, agyak és programok", in: Searle Searle, John Searle: "Minds, brains, and programs": Behavioral and brain sciences 3, pp. 417-424.
- Shalev-Shwartz, Shai/Ben-David, Shai (2014): Shav Shav Shav Shav Shav Shav Shav Shav Shav Shaw: A gépi tanulás megértése. Az elmélettől az algoritmusokig, Cambridge: Cambridge University Press.
- Stamper, Ronald (1988): "A mesterséges intelligencia patológiái: a mesterséges intelligencia felelős használata a szakmai munkában, in: A mesterséges intelligencia patológiái: a mesterséges intelligencia felelős használata a szakmai munkában: AI & SOCIETY, pp2., 3-16.
- Sudmann, Andreas: Sudmann: "Szenarien des Postdigitalen: Deep Learning als Medien Revolution", in: Andreas Sudmann/Christoph Engemann (Hg.), Machine Learning - Medien, Infrastrukturen und Technologien der Künstlichen Intelligenz, 55-74.
- Sudmann, Andreas/Engemann, Christoph (Hg.): Machine Learning - Medien, Infrastrukturen und Technologien der Künstlichen Intelligenz.
- Tendersinfo (2017): Egyesült Arab Emírségek. A mesterséges intelligenciáért felelős miniszter Miniszter előadást tart az AI-ról a DPC rendezvényén, <http://www.tendersinfo.com/> vom 06.01.2017.
- Turing (1950): Turing: "Computing Machinery and Intelligence", in: A Quarterly Review of Psychology and Philosophy, 433-460. 059,,: Mind A Quarterly Review of Psychology and Philosophy.
- Tutt, Andrew (2017): Tutt: "An FDA for Algorithms", in: Közigazgatási Jogi Szemle 69, pp. 83-123.
- Warren, Samuel D./Brandeis, Louis D. (1890): "A magánélethez való jog", in: HLR (Harvard Law Review) 4, pp. 193.
- Yeung, Karen (2018): "Algoritmikus szabályozás: A critical interrogation. A Critical Interrogation", in: Szabályozás és kormányzás 12, pp. 505-523.
- Zuiderveen Borgesius, Frederik J./Möller, Judith/Kruikemeier, Sanne/Ó Fathaigh, Ronan/Irion, Kristina/Dobber, Tom/Bodo, Balazs/Vreese, Claes de (2018): "Online politikai mikrocélzás: In: "Mikrocélú politikai célcsoportok: ígéreték és veszélyek a demokrácia számára", in: Utrecht Law Review 14, pp. 82.

A tudás problémájának újragondolása a vállalati gigantizmus korában¹

Frank Pasquale

Friedrich von Hayek, a laissez-faire egyik legkiemelkedőbb teoretikusa a "tudásproblémát" a központi tervezés leküzdhetetlen akadályának nevezte. A készletek és a munkaerő árával, valamint a fogyasztók fizetési képességével és hajlandóságával kapcsolatos ismeretek annyira szétszórtak és szerteágazóak, hogy még a legavatottabb hatóságok sem férhetnek hozzá mindehhez. Senki sem tud mindent arról, hogy egy gazdaságban hogyan kellene árazni az árukat és szolgáltatásokat. Egyetlen központi döntéshozó sem képes megragadni egyének millióinak sajátos preferenciáit, értékeit és vásárlóerejét. Hayek szerint ez a fajta tudás *elosztott*.

A mesterséges intelligencia és a tömeges megfigyelés korában azonban a központi tervezés csábítása újra felbukkant - ezúttal hatalmas cégek formájában. Mivel az Amazon több milliárd tranzakciót naplózott és elemzett, intim részleteket tud minden vásárlójáról és beszállítójáról. Gondosan kalibrálhatja a képernyők kijelzőit, hogy a vásárlókat bizonyos termékek vagy vásárlási szokások felé terelje, vagy hogy az eladókat saját, olcsóbb, saját kínálatával másolja. Mark Zuckerberg a fogyasztói vágyak mindentudására törekszik, szinte mindenkit profiloz a Facebookon, az Instagramon és a Whatsappon, majd ezt az adathalmazt kihasználva követi a felhasználókat a weben és a való világban (a mobilhasználaton és az eszköz-ujjlenyomatokon keresztül). Valójában nem kell használnod az említett alkalmazások egyikét sem ahhoz, hogy a Facebook/Instagram/Whatsapp fájlokban végezd - a profilok hozzád rendelhetők. A Google "szándékainak adatbázisa" legendás, és a trösztellenes hatóságok világszerte egyre aggasztóbbnak tartják, hogy a Google képes kiszorítani a riválisokat a keresési eredményekből, ha egyszer érdekeltséget szerez az üzleti tevékenységükben. A Google nemcsak azt tudja, hogy a fogyasztók mire keresnek, hanem azt is, hogy más vállalkozások mit keresnek, mit vásárolnak, mit e-maileznek, mit terveznek - az adatfeldolgozási kapacitás és a nyers kommunikációs folyamatok valóban páratlanul jól illeszkednek egymáshoz.

Ez a logika nem korlátozódik az online kontextusra. A koncentráció a div-

a legnagyobb bankok (amelyekről széles körben feltételezik, hogy túl nagyok ahhoz, hogy csődbe menjenek), és a nagybankok

I Ez az esszé eredetileg "Tech Platforms and the Knowledge Problem" címmel jelent meg az *American Affairs* nyári számában. Az 2018. *American Affairs* szíves engedélyével nyomtatjuk újra.

az egészségbiztosítók (amelyek most harmonikaszerűen szorítják és bővítik az orvosi ellátási láncot). A digitális óriásokhoz hasonlóan ezek a pénzügyi és biztosítási cégek nemcsak közvetítőként működnek, részesedést szedve a tranzakciókból, hanem arra is törekuszenek, hogy az ügyfelek és szolgáltatók megfigyeléséből szerzett tudásukat kihasználva kiszorítsák őket, és közvetlenül nyújtsanak szolgáltatásokat és befektetéseket. Ha sikerrel jár, a CVS-Aetna fúzió intenzív vállalati konszolidációt jelez, amely a biztosítók, a szolgáltatók és a közvetítők barokkos sorozata (a gyógyszerészeti ellátások kezelőitől a csoportos beszerzési szervezetekig) további vertikális integrációját fogja látni az óriási egészségügyi szolgáltatókban. Egy CVS-orvos végül beutalhatja a beteget egy CVS-kórházba egy CVS-műtetre, amelyet a CVS által alkalmazott otthoni egészségügyi dolgozók követnek, akik CVS-gyógyszereket hoznak - mindezt a CVS/Aetna biztosítási terv fedezi, amely büntetheti a beteget, ha a CVS hálózatán kívüli szolgáltatókat vesz igénybe. Bár egy ilyen pánoptikus cég disztópikusnak hangozhat, logikus következménye az egészségügyi szolgáltatások ^{kutatóinak} lelkesedése az "integrált ellátórendszerek" iránt, amelyek állítólag hatékonyabban biztosítják az "ellátás koordinálását" és a "körkörös szolgáltatásokat", mint Amerika jelenlegi, széttagolt egészségügyi rendszere.

A nagyhatalmú közvetítők, mint a keresőmotorok és a biztosítók felemelkedése úgy tűnhet, hogy

mint a következő logikus lépés a kapitalizmus fejlődésében. A kritikusok egyre növekvő kórusa azonban megkérdőjelezi az ezeken a területeken vezető cégek méretét és hatókörét. Az Institute for Local Self-Reliance kiemeli, hogy az Amazon a törvények és a szerződések manipulálásával tisztességtelen előnyöket halmoz fel. A nemzetközi trösztellenes hatóságok a Google-t is megkérdőjelezték, hogy a vállalat keresőmotorját és Android operációs rendszerét agresszív módon használja saját szolgáltatásainak népszerűsítésére (és a riválisok lefokozására). Azt is megkérdőjelezzik, hogy a Google és a Facebook miért vásárol fel havonta legalább két vállalatot évek óta. A fogyasztóvédők a manipulatív reklámokra panaszkodnak. A pénzügyek tudósai azt kifogásolják, hogy a megabankok kihasználják az implicit támogatásokat, amelyeket a "túl nagy ahhoz, hogy csődbe menjen" státusz biztosít.

Össze tudnak-e ezek a különböző tiltakozási és kritikai szálak valami tartósabbá és következetesebbé válni? Ez az esszé a társadalmi és gazdasági elégedetlenség lehetséges formáit vizsgálja a következő évtizedekben. Azzal kezdem, hogy bemutatom, hol tartunk: egy hierarchikus, centralizált rendszerben, ahol a vállalatok hatalma óriási, és ahol a nagy nemzeti szabályozó apparátusok tűnnek az egyetlen olyan entitásoknak, amelyek képesek megfékezni ezt a hatalmat. Ezzel a gazdasági valósággal szemben jelenleg a politikai-gazdasági kritika két fontos irányvonalát látom.

A populista lokalizátorok a trösztellenes jogérvényesítés új korszakát akarják, hogy megfékezze az óriáscégeket. A nagy technológiai cégek, a megabankok és a nagy egészségügyi kombinációk e jeffersoni kritikusai decentralizálók. Hisznek abban, hogy a hatalmat egy igazságos társadalomban

szét kell osztani. Támogatják az erős helyi hatóságokat, amelyek közelebb vannak a saját polgáiraikhoz.

Mások a gigantizmust elkerülhetetlennek vagy kívánatosnak tartják, és azzal érvelnek, hogy egyszerűen jobb szabályokra van szükségünk a vállalati hatalommal való visszaélések megakadályozására. A mai Hamiltoni-

és azzal érvelnek, hogy a hatalmas adattárak kritikus fontosságúak a mesterséges intelligencia jövője - és ezáltal a gazdaság termelési dinamikusa - szempontjából. A vezető cégek feldarabolása helyett inkább a jobb szabályozásra összpontosítanak.

A Jefferson- és a Hamiltoni-féléknek nagyon eltérő hosszú távú nézeteik vannak arról, hogy hogyan néz ki egy optimális gazdaság. Hosszú távon elképzeléseik valószínűleg összeegyeztethetetlenek. Rövid távon azonban mindkét reformercsoport fontos tanulságokkal szolgál a hatalmas technológiai, pénzügyi és egészségügyi cégek hatalmával küzdő politikai döntéshozók számára. Ez az esszé ezeket a tanulságokat vizsgálja, és meghatározza, hogy mikor a Jefferson-féle megközelítés a legmegfelelőbb, és mikor Hamilton örökösei a jobb megközelítés.

A Jeffersoni/Hamiltoni szakadék

A technológiai politika gyakran sivár. A vállalatok által finanszírozott agytrösztök arra törekszenek, hogy a reformlehetőségeket a meglévő jogszabályok csipések és kisebb változtatások viszonylag szűk ablakában tartsák. Az akadémiai túlspecializálódás átka szintén rövid pórázon tartja a jogi és politikai professzorok nagy részét. Ennek ellenére vannak a területnek életképes zuga- tok. Két tábor alakult ki: egy decentralizáló tábor, amelyet én Jefferson-féle tábornak neveznék, és egy centralizálóbb, Hamiltoni irányzat, amely az ipari "nagysággal" is megbarátkozik.

A Jeffersoni iskola az Egyesült Államokban a trösztellenes szabályok laza betartásának és általában a verseny előmozdításának problémája körül tömörült. A Nyílt Piacok Intézete, amelyet a New American Foundationból kirúgtak, mert túlságosan ellenséges volt a Google-lel szemben, vezette a támadást. Az OMI vezetői, mint Matt Stoller és Barry Lynn, amellet érvelnek, hogy a Szövetségi Kereskedelmi Bizottságnak (FTC) fel kellene bontania a Facebookot, létrehozva az Instagramot és a WhatsAppot mint konkurens közösségi hálózatokat. Lina Khan, szintén az OMI munkatársa, kimerítő kritikát írt az Amazon gigantizmusáról, amely máris a *Yale Law Journal* egyik legtöbbet letöltött cikke. A szubszidiaritás hangsúlyozása a katolikus társadalmi gondolkodásban a decentralista elmélet egyik forrása is, amelyre a konzervatívok gyakran hivatkoznak a helyi hatóságok és a civil társadalmi intézmények autonómiájának védelmében.

A hamiltoniak között vannak hagyományos centristák (mint például Rob Atkinson, aki Michael Linddel közösen írta a *Big is Beautiful* című könyvet), valamint a politikai spektrum mindkét végéről érkező hangok. Peter Thiel *Zero to One* című könyve - Schumpeter monopolhelyzetének a növekedést ösztönző dicséretét megismételve - a monopolhatalom dicsőítmusa, amely a drámai innovációért járó jogos és szükséges jutalomként igazolja a monopólium előnyeit. A baloldalon

Evgeny Morozov nem szeretné, ha a Google-hoz és a Facebookhoz hasonló cégek adattárolói a szolgáltatások tucatnyi különböző változatában szétszóródnának. Szerinte inkább arról van szó, hogy ezek a vállalatok valószínűleg természetes monopóliumok: minden egyes feladatban egyre jobbak és jobbak lesznek, ha egyre több és több összevont adathoz férnek hozzá az általuk végzett *összes* feladatból. A végső baloldali logika itt a teljesen automatizált luxus felé mutat.

kommunizmus, ahol hatalmas cégek a gépi tanulás és a 3D nyomtatás segítségével megoldják az éhezést, megmentik a környezetet, és véget vetnek a hiány problémájának.² A baloldali centralizátorok azzal is érvelnek, hogy az olyan hatalmas problémákat, mint az éghajlatváltozás, csak a Hamilton-féle megközelítéssel lehet megoldani.

A Jeffersoni és a Hamiltoni elképzelések nagyon eltérő politikai ajánlásokhoz vezetnek a technológiai térben. A Jeffersoniak véget akarnak vetni a Google felvásárlási hullámának. Szerintük a cég egyszerűen túl nagy hatalomra tett szert. De még egyes progresszív szabályozók is átengednék a Google Waze (a forgalomfigyelő alkalmazás) felvásárlását, bármennyire is erősíti a Google hatalmát a térképezés területén, abban a reményben, hogy a vezetési adatok felgyorsíthatják az önvezető autók fejlesztését. A gyorsabb fejlődés ára a hatalom további koncentrációja lehet a Szilícium-völgyben. A jeffersoniak szerint azonban éppen a hatalom, a szabadalmak és a nyereségek megacégekénél történő koncentrációja az, ami visszatartja a kisvállalkozásokat attól, hogy kockázatot vállaljanak az áttörést jelentő technológiák kifejlesztése érdekében.

A Facebook dominanciája a közösségi hálózatok terén hasonló aggályokat vet fel. Az Egyesült Államok és Európa adatvédelmi szabályozó hatóságai vizsgálják, hogy a Facebook eleget tett-e annak érdekében, hogy megvédje a felhasználói adatokat a harmadik féltől származó alkalmazásokkal szemben, mint amilyeneket a Cambridge Analytica és szövetségesei használtak a gyanútlan Facebook-felhasználók tízmillióinak adatgyűjtésére. Megjegyzendő, hogy a Facebook maga is megszorította a harmadik felek hozzáférését az általa gyűjtött adatokhoz, 2013, részben azért, mert aggódott amiatt, hogy más cégek képesek voltak a híres "közösségi gráf" - az információk és kapcsolatok adatbázisa, amely a közösségi hálózatot a hirdetők számára oly értékessé teszi - kisebb, de még mindig erőteljes változatait létrehozni. A Jeffersonban a Facebooknak a külső fejlesztők felé történő adatáramlásra vonatkozó szigorítása gyanús. Úgy tűnik, hogy a közösségi hálózat megpróbál monopolizálni egy olyan adathalmazt, amely nélkülözhetetlen nyersanyag lehet a jövőbeli startupok számára. Hamiltoni szempontból azonban az adathalmaz egyetlen hatalmas cégben való biztosítása tűnik a legfelelősségteljesebb megoldásnak (feltéve, hogy a cég jól szabályozott). Miután az adatok véglegesen átkerülnek a Facebookról más cégekhez, gyakorlatilag nagyon nehéz lehet biztosítani, hogy ne éljenek vissza velük. A versenytársak (vagy Ariel Ezrachi és Maurice Stucke kifejezésével élve a "baráti ellenségek") nem férhetnek hozzá a Facebook szerverein biztonságban lévő adatokhoz - de a hackerek, zsarolók vagy a katonai szintű pszichológiai műveletekre szakosodott árnyék-adattörők sem. Ahhoz, hogy az "elszabadult adatok" ne hozzanak létre mindannyiunk számára egy teljes nyilvánosságra hozatalra épülő disztópiát, szükségesnek tűnik a "biztonsági feudalizmus".

Az elkövetkező években kiéleződik a politikai konfliktus a Jefferson- és Hamiltoni-pártiak, a "kicsi a szép" demokraták és a centralista bürokraták között.

Ahhoz, hogy megértsük az egyes irányzatok szerepét a digitális szférában, részletesebben meg kell vizsgálnunk a megközelítéseiket.

² Ilyen szerzők például Leigh Phillips és Michal Rozworski, *People's Republic of Walmart* (Verso, 2019); Aaron Bastani, *Fully Automated Luxury Communism* (Verso, 2019) és Peter Frase, *Four Futures* (2016).

A távollévők tulajdonjogának Jefferson-féle kritikája

A digitális kapitalizmus legnagyobb, legsikeresebb vállalatai általában platformként szolgálnak, és inkább rangsorolnak és minősítenek más entitásokat, mintsem közvetlenül kínálnának árukat és szolgáltatásokat. Ez a stratégia lehetővé teszi a platform számára, hogy a kockázatot kiszervezze az eladókra és a fogyasztókra, miközben megbízhatóan szedi be a részesedést minden egyes tranzakcióból. Ahogyan egy pénzügyi közvetítő is profitálhat a tranzakciós díjakból, függetlenül attól, hogy az egyes befektetések szárnyalnak-e vagy romlanak, úgy a platform a bevételeket a front- endről zsebeli be, függetlenül az általa közvetített kapcsolatok minőségétől.

Ez a közvetítői szerep számos lehetőséget teremt a platformok számára. Például felügyelik a tranzakciókat és eldöntik a vitás ügyeket, ami korábban a kormányok feladata volt. A platformoknak ezt az erőteljes új szerepét "funkcionális szuverenitásnak" nevezem, hogy jelezzem azt a hatalmi szintet, amelyet egy magánvállalat akkor ér el, amikor már nem egy a sok piaci szereplő közül, hanem a tényleges piaci szereplők fő felügyelője és szervezője. Az olyan platformok, mint az Amazon és a Google, funkcionális szuverenitással rendelkeznek egyre több piac felett, kvázi kormányzati szerepet játszanak, amikor a fogyasztók, a marketingesek, a tartalomszolgáltatók és a harmadik és negyedik felek egyre szélesebb köre közötti konfliktusokat bírálják el.

A személyre szabás a digitális stratégiák mantrája, akik hajlamosak azt feltételezni, hogy ez egy "win-win" ajánlat. A személyre szabott keresési eredmények például egyrészt megóvják a Google felhasználóit a figyelemeltereléstől, másrészt hajlamosak összekapcsolni őket a kívánt termékekkel. A vágyaink és "fájdalompontjaink", jövedelmi szintünk és vagyoni helyzetünk egyre jobb ismeretere épülő online piacok azonban könnyen a kizsákmányolás felé hajolhatnak. A platformoknak érdekében áll bizonyos digitális szférák intenzív figyelemmel kísérése és alakítása a profitmaximalizálás (és másodsorban saját hírnevük fenntartása) érdekében. Azonban szüntelen törekvésükben, hogy egyre több ágazatot csatoljanak saját ökoszisztémájukhoz, túl gyakran túl sokat haragnak el, mint amennyit meg tudnak rágni. Hajlamosak túlbecsülni az automatizálás erejét a modern piacok által generált összes igény feldolgozásában.

Ez egy másik, a monopolista vállalkozások történelméből ismert problémához vezetett: a távollévők tulajdonjoga. Amikor egy hatalmas cég megvásárol egy üzletet több ezer mérföldre a központjától, akkor az üzlet a tulajdonában van, és profitra törekszik belőle, de a teljesítményét csak nyersen értékelheti, és kevésbé érdeklí az a közösség, amelybe az üzlet beágyazódott. Az üzlet elhanyagolhatja a hagyományos funkciókat, amelyeket szolgált, egyszerűen azért, hogy maximalizálja a bevételeket, amelyeket a távollévő tulajdonos követel. Egy jelenlévő, a közösségben lakó tulajdonos nagyobb valószínűséggel fogja úgy működtetni az üzletet, hogy az megfeleljen a közösség érdekeinek és értékeinek,

mivel a jelenlévő tulajdonos maga is tapasztalni fogja, ha az üzlet javulást vagy romlást okoz a közösségben.

Hasonló dinamika alakul ki az interneten is. A Google birtokolja a legnagyobb online videógyűjteményt, de a YouTube leányvállalatának jövedelmezősége a kiszámított hanyagságtól függ.

a platform számos aspektusát. Az elmúlt két évben kritikuskok egész sora bírálta a céget, amiért zavaró, ízléstelen, sokkoló és sértő tartalmakat népszerűsít, még gyerekeknek is. A Google nemrégiben tett bejelentése, miszerint a YouTube által oly nagymértékben támogatott összeesküvés-elméleti videók leleplezése érdekében a Wikipe- dia linkjeit fogja népszerűsíteni, a kiszervezés újabb rétegeit jelenti - egy profitorientált vállalatról egy nonprofit szervezetre, amely viszont a tartalom feletti hatalmat önkéntesekre ruházza át, akiket egy árnyékos adminisztrátori réteg irányít.

A Jefferson-félék számára a válasz nyilvánvaló: nem szabad, hogy egyetlen, molyirtó vállalatnak legyen hatalma ennyi videó felett. A YouTube szerint a YouTube-nak szüksége van erre a méretre, hogy ingyenesen kínálhassa szolgáltatásait; a Jefferson-lakosok szerint a reklámvezérelt üzleti modell csak egy módja annak, hogy alulmúlják az előfizetéses szolgáltatásokat, amelyek jobban tudnak kezelni kínálatukat. A Jefferson-pártiak arra is rámutatnak, hogy nagyon nehéz megmondani, hogy az olyan szolgáltatások, mint a YouTube, milyen mértékben szolgálják ténylegesen a felhasználókat és a tartalomgyártókat, és milyen mértékben léteznek csupán a reklámbevételek maximalizálása érdekében.

A Hamiltoni perspektíva az új digitális közművekről

A jeffersoniak vezéreszméje az amerikai trösztellenes törvény eredeti szándéka - hogy a hatalmas vállalatok annyira képesek voltak uralni ügyfeleiket, alkalmazottaikat és közösségeiket, hogy szét kellett őket törni. Egy nagyvállalat kisebb részekre való felosztása "strukturális jogorvoslat", mert a társadalom alapvető tulajdonosi érdekeit és ellenőrzését érinti. A legnagyobb vállalatok feldarabolására irányuló populista követelés trösztellenes támadásokat inspirált a Stan- dard Olajtól a Brown Shoe-n át a Microsoftig számos cég ellen.

Az utóbbi időben azonban a trösztellenes hatóságok óvatosabbak a nagyvállalatok feldarabolásával kapcsolatban. Az Igazságügyi Minisztérium és a Szövetségi Kereskedelmi Bizottság is leszűkítette érdeklődését, és szinte kizárólag a nagyvállalatok fogyasztókra gyakorolt jelenlegi árhatásaira összpontosít. Így egy olyan hatalmas cég, amely a minőség csökkentésével a versenytársaknak alulmúlja a versenyt, kevésbé aggasztja őket. Az a lehetőség sem, hogy ugyanez a cég végül, ha monopolizál egy területet, drámai módon emeli az árakat a fogyasztók számára (vagy csökkenti a dolgozók bérét). Ehelyett a hatékonyságnak szentelik magukat - többet, kevesebbet, gyorsabban. Az ingyenes vagy alacsony árak rövid távon felülmúlják az egyéb megfontolásokat.

Hogy lássuk a rövidtávúság iránti megszállottság gyakorlati hatásait, képzeljük el, hogy a Google-ban rákeresünk az "időjárás" kifejezésre, és azonnal meglátjuk, hogy a saját időjárás-előrejelzésünk betölti a mobil képernyőjét. Ha három előrejelző oldalt linkelt volna a képernyőre, akkor talán több

megjelenést és reklámbevételt irányíthatott volna a különböző felületű, több vagy kevesebb információt tartalmazó vagy más variációkkal rendelkező oldalakra. Például a WeatherSpark oldal gyönyörűen pontos képet adott a viharok időbeli mozgásáról - tökéletes vizuális analógja az Accuweather min- dösségének.

percenkénti előrejelzések az esőre vagy a tiszta égre vonatkozóan. A WeatherSpark azonban már nem kínálja ezt a szolgáltatást, és ki tudja, hány más startup adta fel, hogy elfoglalja ezt a helyet. A trösztellenes hatóságok számára nincs ok a beavatkozásra - a fogyasztók a Google felületéről kapják meg az időjárás alapjait, ráadásul ingyen. Ez egy rövid távú szemléletmód, amely a hosszú távú megfontolásokat kihagyja a jelen idejű tudományosság nevében. Az ő világgépükben nincs helye vitának arról, hogy léteznek-e jobb vagy rosszabb alternatívák, vagy hogy létezniük kellene-e jobb vagy rosszabb alternatíváknak. Az Antirust elvileg a "versenyt, nem a versenytársakat" hivatott védeni - és a minőség iránti aggodalom egyedülálló hiánya a mélyeséges érdektelenségbe torkollik, hogy a jelenlegi vagy jövőbeli versenytársak jobb munkát végezhetnek-e, mint egy digitális behemót. De honnan tudhatnánk, hogy van-e verseny, ha nincsenek versenytársak, akik azt biztosítanák?

A trösztellenes jog e szűkítése nyomán a hamiltoni hangok a közhasznú jog újjáélesztését követelték a hatalmas online cégek hatalmának megfékezésére. A 20. század eleji közüzemi szabályozók nem akarták, hogy 10 különböző teleföntársaság ássa fel az utcákat, hogy versenyt teremtsen a telefonszolgáltatások terén. Nem gondoltak a helyi energiatermelésre sem (bármennyire is csábító lehet ez a lehetőség ma azok számára, akik egy elosztott, megújuló energiaforrásokra épülő hálózatra törekszenek). Ehelyett ezek a szabályozók elfogadták a távközlési, energia- és egyéb cégek tömegesedését, mint a modern gazdasági racionalizáció elkerülhetetlen aspektusát. Csak egy olyan államot (és szakszervezeteket) akartak, amely elég masszív ahhoz, hogy ellensúlyozó erőket kínáljon.

A hamiltoniak számára egy olyan ügynökség, mint a Szövetségi Hírközlési Bizottság, a strukturális jogorvoslatok viselkedési alternatíváját jelenti. A Szövetségi Keresőbizottság például figyelemmel kísérhetné, hogy a Google hogyan kezeli a konkurens cégeket a keresési eredményekben, és rákényszeríthetné, hogy a saját szolgáltatásai helyett alternatívákat kínáljon az ilyen eredményekben.³ Az európai versenyhatóságok hatékonyan létrehozhatnak egy ilyen ügynökséget, ha komolyan gondolják a Google vertikális kereső versenytársakkal való bánásmódjának ellenőrzését (azaz a bizonyos típusú áruk vagy szolgáltatások szűk körű keresését).

A hamiltoniak a technokrata bal-liberalizmussal azonosulnak. Olyan eszközöket akarnak alkalmazni, mint a költség-haszon elemzés és a fejlett adatelemzés, hogy kiszámítsák, mikor van értelme egy szolgáltatást konglomerátumba olvasztani, és mikor van értelme olyan szabályokat alkotni, amelyek a cégek függetlenségét feltételezik. Vannak azonban ideológiailag ambiciózusabbak is, amelyek az ipari méret és kiterjedés mellett érvelnek. Evgeny Morozov például óva int a Google vagy a Facebook feldarabolására irányuló törekvésektől, mivel a mesterséges intelligencia terén csak akkor lehet előrelépni, ha valóban hatalmas adatmennyiségeket vonnak össze. A Chapo Trap House szocialistái egy nemrégiben sugárzott podcastban azzal viccelődtek, hogy örülnek

az Amazon hatalomkonszolidációjának. Ha egyszer az ország minden üzletét átveszi, könnyű lesz "levágni a fejét", és egyszerűen im- pózálni a gazdaság feletti kormányzati ellenőrzést. "Ingyen Whole Foods hot bar minden-

3 O. Bracha & F. Pasquale, *Szövetségi Keresőbizottság: Access, Fairness, and Accountability in the Law of Search*, Cornell 93 L. Rev. (11492008).

egy!" - hangzott az elképzelt végkifejlet. Hasonlóképpen, ha az Egyesült Államokban az összes magán-egészségbiztosító egyesülne, a színpad végre készen állna az "egységes fizetős" rendszerre: a kormánynak csak az egyetlen megmaradt biztosítót kellene átvennie.

A *Jacobin* szerzői (köztük Alyssa Battistoni, Peter Frase, Christian Parenti) szintén egy neo-Hamiltoni megközelítést fogalmaznak meg a fejlett vállalati kapacitásról, amelyet a kormány és a szakszervezetek ellensúlyozó ereje mérésével. A németországi általános szakszervezeti tanácshoz és a megagyártókhoz hasonló nagy csúcsszervezetekbe való központosítás lehetővé tenné a korporatista tárgyalásokat az erőforrások masszív koncentrációja által lehetővé tett beruházási típusok zsákmányának felosztásáról. Németország legnagyobb szakszervezete nemrégiben tárgyalt arról, hogy tagjainak munkaidejét órákra 28 csökkentsé, miközben 4,3%-os béremelést is kapott - pontosan olyan alkut, amelyet az amerikai munkavállalók is kaphattak volna, ha a termelékenység növekedését az 1970-es évek vége óta széles körben megosztották volna, és ha az üzleti és a munkavállalói oldal hasonlóan szervezett lett volna.

A hamiltoni vízió a legambiciózusabb formájában a teljesen automatizált luxuskommu- nizmusból garantált egyetemes alapjövedelem álma felé tendál. A mesterséges intelligencia és a robotok a munkásokat utánozzák, akiket továbbra is azért az adatért fizetnek, amellyel ők (vagy elődeik) hozzájárultak a mesterséges intelligencia fejlődéséhez. A hamiltoni- anizmus a geomérnökség gazdasági megfelelője lehet - a radikálisan új és nagyszabású dolgok felkarolása, amely abból az érzésből fakad, hogy az egyenlőtlenségek és az éghajlatváltozás olyan hatalmas problémákat jelentenek, amelyeket csak a gyors technológiai fejlődés képes megoldani. A jeffersoniak az elővigyázatosság elvéhez hasonlóan ragaszkodnak, és megkérdőjelezzik, hogy bármely szervezetnek fel kellene-e halmoznia a szükséges hatalmat ahhoz, hogy mondjuk mindenki genomját összehasonlítsa, a munkavállalók millióinak mozgását robotokba programozható viselkedésmintákká alakítsa, vagy minden állampolgárról társadalmi hitelpontokat vezessen.

A Jeffersoni és Hamiltoni nézőpontok összeegyeztetése

Mindezek a tendenciák a 21. század gazdasági gondolkodásának új törésvonalait jelzik. E feszültségek enyhítése érdekében vissza kell térnünk a neoliberais projekt néhány alapvető feszültségéhez. Az 1930-as és 40-es években a Chicagói Egyetem közgazdásza, Henry

C. Simons arra figyelmeztetett, hogy a monopóliumok halálos veszélyt jelentenek a szabad és nyitott piacok klasszikus liberális eszméire. Simons *A Positive Program for Laissez Faire* című művében, amelyet Simons 1934, a következővel érvelt: "A demokrácia nagy ellensége a monopólium, annak minden formájában: gigantikus vállalatok, kereskedelmi szövetségek és más árszabályozó ügynökségek, szakszervezetek - vagy általában a hatalom

szerveződése és koncentrációja a funkcionális osztályokon belül". Az 1950-es évekre azonban George Stigler és Aaron Director kiszorította Simons-t Chicagóban, és a trösztellenes joggal kapcsolatban jóval gyakorlatiasabb megközelítést kínált. A koncentrált *állami és szakszervezeti* hatalmat sokkal nagyobb fenyegetésnek tekintették a társadalomra nézve, mint a koncentrált vállalati hatalmat. És mivel az előbbi

az utóbbiak elleni küzdelemhez szükséges, lekicsinyelték a hatalmas vállalatok által okozott károkat (a magatartás szűken körülhatárolt kategóriáján kívül, amely egyre kisebb lett, ahogy a chicagói tudósok, mint Robert Bork, csökkentették a trösztellenes jog területét és erejét).

Mi lett volna, ha Chicago Simons útját követte volna az igazgató helyett? A neoliberaisok talán kiegyensúlyozottabb megközelítést alkalmaztak volna a társadalomban a túlzott hatalommal való szembenézésre. A trösztellenes hatóságok jobban ellenálltak volna a behemót cégek azon törekvéseinek, hogy központosítsák az adatgyűjtést és a munkavállalók ellenőrzését. A politikai döntéshozók jobban kiegyensúlyozhatták volna az állami hatalom csökkentésére irányuló erőfeszítéseket az olyan párhuzamos erőfeszítésekkel, amelyek célja, hogy csökkentsék a vállalatok azon képességét, hogy akarataikat a közösségekre és a munkavállalókra kényszerítsék. Az 1950-es évek szakszervezeti hatalom csökkentésére irányuló politikai menetrendje a 2010-es években nevelésesnek tűnik, amikor a szakszervezetek sűrűsége meredeken csökkent, miközben a vállalati koncentráció nőtt.

A jeffersoniaknak megvan a maguk vakfoltja, amikor a munkáról van szó. A Jefferson-i irodalom túlságosan idealizálja a kisbirtokosokat, és a minden ember mint vállalkozó eszméjét képviseli. De a legtöbben közülünk életük nagy részében valaki másnak dolgoznak, és fognak is dolgozni. Ezért Atkinson és Lind joggal érvel a *Big is Beautiful* című könyvében emellett, hogy a kisvállalkozásokat számos olyan munka- és fogyasztóvédelmi törvénynek kellene betartani, amelyek ma csak a nagyobb vállalatokra vonatkoznak. Ellenkező esetben a franchise és a platformkapitalizmus varázslói egyszerűen új utakat találnak majd arra, hogy a meglévő vállalatokat kisebb egységekre bontják, hogy megkerüljék a szabályozást. Az alultökésített és ítélőképes kisvállalkozások a tökéletes üzleti törvényszegők, mivel kevés vesztenivalójuk van, ha elkapják őket.

Ugyanakkor tiszteletben kell tartani a jeffersoniak egyik alapvető felismerését: számos termék és szolgáltatás kezelésére valóban nincs "egyetlen legjobb módszer". A kérdés tehát az, hogyan lehet meghatározni a különböző iparágakban a vállalkozás optimális méretét és hatókörét. Ha egy cégnek jóhiszeműen szüksége van adatokra egy probléma megoldásához (például optimális útvonalak kiszámítása önvezető autókban álló flotta számára), az sokkal jobb indok a "nagyságra", mint az adatok egyszerű felhasználása a kereskedelmi tranzakciók saját előnyére történő átrendezésére. Stacy Mitchell, az Institute for Local Self-Reliance munkatársa megfigyelte, hogy "amikor a harmadik fél eladók új termékeket tesznek közzé, az Amazon nyomom követi a tranzakciókat, majd a legnépszerűbb termékek közül sokat elkezd eladni". Bármennyire is növeli ez a gyakorlat a gazdasági termelékenységet, elfogadhatatlan áron teszi ezt, mivel a hatalom egy cégben koncentrálódik, miközben a cégen kívüli vállalkozásokat elriasztja. A politikai döntéshozóknak meg kellene védeniük a kiszolgáltatott eladókat ettől.

A jeffersoniak strukturális aggályai a gazdaság túlzott koncentrációja elleni

első védelmi vonalat jelentik. A versenyhatóságoknak komolyan kell venniük ezeket, különösen akkor, ha a nagyságrendnek nincs érdemi *termelői* indoka. Ha az Amazonnak berendezésgyártókat kell vásárolnia, hogy a jobb Kindle előállítására érdekében vertikális integrációt folytasson, rendben - de ha más cégeket vásárol fel pusztán (vagy elsősorban) azért, hogy növelje a fogyasztókkal vagy beszállítókkal szembeni alkupozícióját, akkor ez nem lehetséges.

az egyesülések jogos indokai. Hasonlóképpen a hatóságoknak fel kell ismerniük, hogy a "jobb szolgáltatás" vagy a felhasználókra vonatkozó "olcsóbb következtetések" nevében történő fúzió túlnyomó alkupozíciót eredményezhet egy platform számára az általa kiszolgált hirdetőkkal szemben - és a felhasználók magánéletébe való beavatkozás lehetőségét. Ezek azok a fő okok, amelyek miatt az FTC-nek meg kellett volna akadályoznia a DoubleClick Google általi megvásárlását, valamint az Instagram és a WhatsApp Facebook általi felvásárlását.⁴

Politikailag nehéz lesz "kibogozni a jelenleg domináns cégek rántottáját". A hatóságok óvakodnak a fúziók és felvásárlások visszafordításától, még akkor is, ha azok utólag nyilvánvalóan problémásak. Míg a Jeffersoniak megakadályozhatják, hogy a digitális óriások még nagyobbak legyenek, a Hamiltoniaknak figyelemmel kell kísérniük a jelenlegi gyakorlatukat, és be kell avatkozniuk, ha azok átlépik a társadalmi normákat. Az algoritmikus elszámoltathatóság mozgalmának köszönhetően tudjuk, hogy az algoritmikus vállalati döntéshozatalt gyakran alkalmazzák a diszkriminációellenes, a tisztességes eljárásra vonatkozó és a médiatörvények megkerülésére. Az olyan ügynökségeknek, mint a Consumer Financial Protection Bureau, a Federal Communications Commission, a Federal Trade Commission és az állami főügyészek, szoros figyelemmel kell kísérniük a platformokat annak biztosítása érdekében, hogy valóban tisztességes esélyt adjanak a felhasználóiknak az ügyfelekhez, a reklámokhoz és a növekedéshez való hozzáféréshez. Ezek a cégek már nem egyszerű piaci szereplők. Piacot csinálnak, és úgy is kell kezelni őket. Még Mark Zuckerberg is elismerte nemrég, hogy a kérdés nem az, *hogyan* szabályozzuk-e a Facebookot, hanem az, *hogyan*. A többi technológiai vállalat vezérigazgatójának is hasonlóan nyitottnak kellene lennie a társadalmi értékek iránt, amelyeket oly sokáig elkerültek.

A kontextus számít

Folyamatos küzdelem folyik arról, hogy az online tér uralma milyen felelősséggel járjon. A befektetők a monopolizáció fantáziáját követelik: cégük nem csupán elfoglal egy területet, hanem "várárkokat" alakít ki a belépőkkel szemben, hogy garantálja mind a jelenlegi hozamot, mind a jövőbeli növekedést. A működési költségvetési megszorítások mindennapi valósága azonban ugyanezeket a cégeket az ab- sentee-tulajdonlás patológiái felé sodorja.

A jog segíthet feloldani ezeket a feszültségeket. A versenyjog a nagy technológiai platformok funkcionális szuverenitását veszi célba, csökkentve egy cég dominanciájának tétjét egy területen. A trösztellenes hatóságoknak legalábbis meg kellett volna akadályozniuk, hogy a Facebook felvásárolja az Instagramot és a Whatsappot, ahelyett, hogy hagyják, hogy a kommunikáció feletti dominanciájával eltíporjon néhányat azon kevés szervezetek közül, amelyek képesek alternatív online társulási módokat biztosítani. Tíz, húsz, vagy száz

közösségi háló-

4 A trösztellenes politika előnyeiről és hátrányairól lásd Frank Pasquale, *Dominant Search Engines: The Next Digital Decade* (2011).

a versenyjog megfelelő érvényesítése esetén a munkálatok előbb-utóbb létrejöhethetnek, és a teroperabilitási szabványok zökkenőmentes kapcsolatot biztosíthatnának a közösségi hálózatok szövetségei között, ahogyan az AT&T, a T-Mobile és a Verizon ügyfelei is zökkenőmentesen tudnak egymással beszélni. Ha ez a sokszínűség kialakulna, kevésbé aggódhatnánk amiatt, hogy a Szilícium-völgyben néhány személy lényegében a világ legfelsőbb bíróságaként dönti el, hogy melyik kifejezőmód megfelelő az úgynevezett "globális közösség" számára, és mit kell betiltani vagy eltakarni (gyakran titkos algoritmikus manipulációval).⁵

Ha az ipari óriásokat nem lehet feldarabolni, még mindig sokféleképpen lehet hatalmukat újratermelni. A közüzemi jellegű szabályozás enyhíti a hiányzó tulajdonosok legrosszabb hibáit, valamint a hatalmasok szeszélyeit. Az állam megkövetelheti a Google-tól, hogy bizonyos tartalmakat közvetítsen a YouTube-on, ahogyan a kábelhálózatoktól is megkövetelte, hogy a helyi híreket közvetítsék. Sőt, ha a politikai döntéshozók attól tartanak, hogy az olyan cégek, mint a Google, az Amazon vagy az Uber túl nagy részesedést vesznek ki az ügyletekből, akkor a biztosítási szabályozók példáját követhetik, akik gyakran korlátozzák a biztosítóknak, hogy a díjak 15-20%-át szedjék be (a többit az egészségügyi ellátásra kell fordítaniuk). Ez a fajta korlátozás elismeri e cégek szolgáltatásainak infrastrukturális minőségét. Nem szeretnénk olyan világban élni, ahol az áramszolgáltató végtelenül megemelheti a díjakat, hogy kihasználja a tőle való függőségünket. A digitális monopolistáknak is hasonló korlátozásokkal kellene szembenéznük.

Bár a Jefferson-féle bizalomromboló és a Hamiltoni közüzemi szabályozók nagyon különbözően látják a politikai gazdaságtant, mindketten szembeszállnak a digitális kapitalizmus híveinek korlátlan törekvéseivel (és kiábrándító hétköznapi valóságával). Segítenek megérteni azt is, hogy az óriáscégek mikor segíthetnek megoldani a Hayek által azonosított "tudásproblémát", és mikor súlyosbítják azt a homály és a ködösítés révén.⁶ Ha a konglomeráció és a vertikális fúziók valóban segítenek megoldani a valós problémákat - a gyorsabb közlekedést, a jobb élelmiszert, a jobb minőségű egészségügyi ellátást és a többit -, akkor a szerzőknek hagyniuk kellene, hogy továbblépjenek. Az ilyen ipari nagyság segít abban, hogy jobban megértsük és irányítsuk a természeti világot. Az államoknak azonban meg kellene akadályozniuk az alkupozíció és a tőkeáttétel puszta felhalmozását. Az ilyen lépések az emberek ellenőrzésére irányulnak - ami az ipari szervezet sokkal kevésbé üdvös célja. A termelékenységre és az inkluzív jólétre összpontosító gazdaságpolitika egyensúlyt teremt, és igazságot szolgáltat mind a Jefferson, mind a Hamiltoni kritikusok fontos meglátásainak, amelyek egyre inkább szklerotikus gazdaságunkkal kapcsolatban megfogalmazódtak.

- 5 Kate Klonick és Thomas Kadri, "How to Make Facebook's 'Supreme Court' Work", *N.Y. Times*, Nov. 17, 2018.
- 6 Walter Adams és James W. Brock, *The Bigness Complex* (Stanford University Press, 2004).

A mesterséges intelligencia és a művészet demokratizálódása

Jens Schröter

Gép akarok lenni
Andy Warhol

I. Bevezetés

A mesterséges intelligenciával kapcsolatos jelenlegi felhajtás annyiban túlzó, hogy a gépi tanulás sokat emlegetett eljárásai viszonylag speciális problémákat, például a képfelismerést vagy általánosabban a minta-megkülönböztetést célozzák. Az univerzális, általános mesterséges intelligencia nem cél, és talán nem is lehetséges. Mindazonáltal van egy kapcsolódó jelenség, amely manapság néha ideges vitákban is felbukkan: Ez a kérdés az, hogy a gépi tanulás rendszerei lehetnek-e "kreatívak". Ez a vita akkor forrósodott fel, amikor az Alpha Go képes volt legyőzni Lee Sedolt a Go-ban, különösen a 2016. március 10th-i második 37meccs híres lépése került szóba, amely még a tapasztalt Go játékosok számára is radikálisan újnak tűnt: Senki sem látta előre ezt a lépést - tehát kreatívnak tekinthető vagy sem?¹ Nyilvánvaló okokból ez az esszé nem tud a kreativitás elméletének vagy a "kreatív szubjektivitásnak" a mélyére hatolni², de az "új" mesterséges intelligenciáról szóló diskurzus egy olyan jelenségével foglalkozni, amelyben a kreativitás kérdése különösen sürgető: ez a művészet, a mesterséges intelligencia által létrehozott művészet területe.

1 Lásd Mersch (2019) alapvető kritikáját a gépi kreativitás gondolatáról.

2 Vö. többek között Sternberg (1999) és Reckwitz (2012).

The AI-Art Gold Rush Is Here

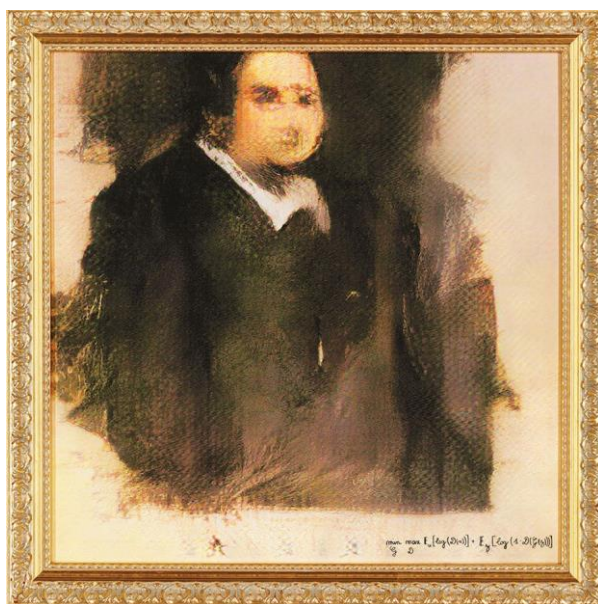
An artificial-intelligence "artist" got a solo show at a Chelsea gallery. Will it reinvent art, or destroy it?



AI-generated "faceless portraits" by Ahmed Elgarnal and AIGAN.

1. ábra: Az AI Art Goldrush itt van

Az *AI Art Gold Rush is here* - ez a címe Ian Bogost kritikái esszéjének (2019). Az ábrázolt képek valahogy úgy néznek ki, mint torzított renaissance-i portrék, talán egy kis Francis Baconnel (1. ábra).



2. ábra: Edmond de Belamy portréja

Különösen a portrék kissé homályos jellege hasonlít *Edmond de Belamy* híres *portréjára*, amelyet októberben a Christie's aukciósházban \$-ért 432.500 árvereztek el. Hogy 2018. ez a számítógép által generált kép viszonylag magas áron kelt el, meglepő volt, az algoritmussal mint aláírással való viccelés ismét vitát váltott ki az AI "alkotóképességéről", és felmerült a kérdés, hogy egy AI rendszer lehet-e művész vagy szerző. Emellett vita indult arról is, hogy a mesterséges intelligencia-rendszert használó *Obvious* művészeti kollektíva az "igazi" szerző, vagy akár az a programozó, aki az algoritmusok egy részét kifejlesztette.³ De a műalkotás más szempontból is meglepő volt: Egy meglehetősen konzervatív műfajhoz és egy meglehetősen konzervatív ábrázolási stílushoz folyamodott: úgy tűnt, hogy az elmosódottság nemcsak egy bizonyos "technicitást", hanem a művészet⁴ konzervatív felfogását is felidézi - lásd az impresszionisták "elmosódott" stílusáról szóló kissé furcsa vitákat (vö. Payne 2007). Tekintettel arra, hogy a festészet fejlődése a huszadik században olyan új formákat fejlesztett ki, mint az absztrakció, felmerül a kérdés, hogy miért van szükség a festészet konzervatív stílusára ahhoz, hogy az AI kreativitását demonstráljuk?

És még az aláírással való vicc is a művészet bizonyos hagyományos felfogását mutatja, amennyiben sok művész (lásd Andy Warhol híres fenti idézetét) valószínűsíti a művész hagyományos mítoszát: gondoljunk csak a szürrealista automatikus írásra, Cage aleatorikus eljárásaira vagy Elaine Sturtevant megkettőző gesztusaira, hogy csak néhányat említsünk. Talán *a* festményt tisztán ironikusan értelmezte - és éppen azzal akarta nevetségessé tenni az AI "újdonságát", hogy előtérbe helyezte annak konzervatív "ízlését". Az új technológiák "demokratizálása" gyakran konzervatív és bevett formák adaptálását igényli, hogy a tömegpiacok számára adaptálható legyen. Ezért egy absztrakt műalkotás elégtelenség tekinthető a mesterséges intelligencia kreativitásának bizonyítására, azon egyszerű oknál fogva, hogy sok embernek még mindig gondot okoz az absztrakt művészeti formák művészetként való elfogadása - vagy a valódi művészethez képest "túl könnyűnek" tartják azokat.

Ezért - és ez rövid esszém központi érve - kulcsfontosságú, hogy a "mesterséges intelligencia művészetről", a "gépi kreativitás" következményeiről, és így *a művészi munka* (lehetséges) *automatizálásáról* szóló vitát a tárgykörbe helyezzük: Néha elfelejtjük, hogy a számítógéppel, pontosabban a mesterséges intelligenciával kapcsolatos művészeti alkotás gondolata hozzáadódik a munka intelligens gépek általi automatizálásáról szóló, meglehetősen ideges diskurzushoz.⁵ Ha a gépek képesek művészetet előállítani - szól a feltételezés -, akkor tömegesen, sorozatban, iparilag is képesek lennének művészetet előállítani mindenki számára. Akkor nem lennének többé ritka zsenik által készített auratikus műalkotások. És senkinek sem kellene milliókat fizetnie a műalkotásokért.

3 Vö. Sudmann (2019) éleslátó cikkét.

4 Az AI-rendszereknek nem feltétlenül kell homályos képeket készíteniük - de a "modern művészetnek"

bizonyos értelemben muszáj.

bizonyos értelemben túllép a realizmus fogalmán.

- 5 Vö. a "Digitale Technologien und das Verschwinden der Arbeit" (2019) című cikkemet. Lásd még Bo- gost (2019): "Tekintettel a robotoktól való általános félelmekre, amelyek az emberi munkahelyek átvételétől tartanak, érthető, hogy egyes nézők éppen a képzőművészeknél látnak áldozatként egy mesterséges intelligenciát, amely átveszi a munkát".

A II. részben kifejezetten az "információs esztétikát" szeretném megvizsgálni, egy olyan, az 1960-as évekből származó diskurzust és gyakorlatot, amelyet már az a gondolat vezérelt, hogy a művészetet a számítógépek *segítségével* (és nem elsősorban a számítógépekkel mint eszközökkel) hozzák létre.⁶ A III. részben azt fogom megvitatni, hogy miért nem működött akkoriban a művészi munka automatizálásának (és így a "művészet" gépekkel való tömeges előállításának) gondolata.⁷ A IV. részben visszatérek az "AI art" közelmúltbeli "aranylázához", és újraolvasom azt az információs esztétika diszkurzusának fényében.

II. Rövid megjegyzések az információs esztétikáról

Az információs esztétika eredete az esztétika "mértékének" formális meghatározására tett kísérlet. David 1933, Birkhoff megfogalmazott egy egyenletet (3. ábra), amelyben az O változó egy adott mű rendezettségének mértékét, a C változó pedig a "komplexitást" jelöli. M annak a mértéke, hogy mennyire esztétikus egy műalkotás.

$$M = O/C$$

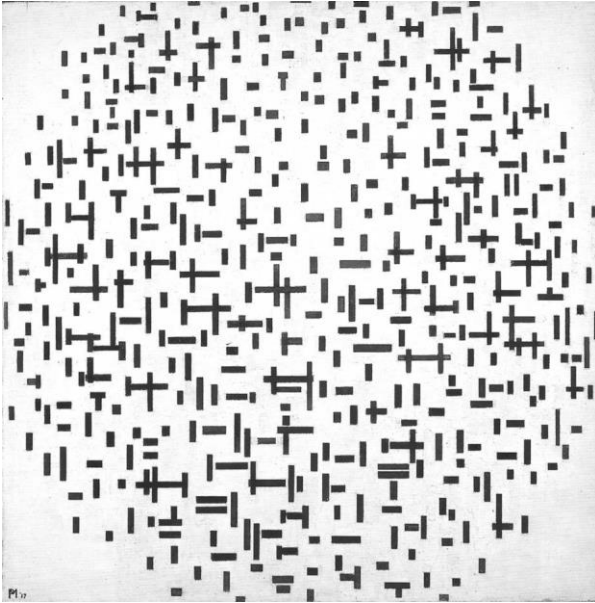
3. ábra: Birkhoff-egyenlet

Eszerint minél rendezettebb és minél kevésbé összetett egy műalkotás, annál esztétikusabb lenne. Eltekintve attól, hogy nehéz pontosan megérteni, hogyan lehet egy adott esetben meghatározni a rendezettség és a bonyolultság mértékét, ez a kísérlet, hogy a művészet "esztétikai minőségét" egy egyenletben fejezzük ki, ma már furcsának tűnik számunkra. Ennek ellenére, különösen az 1960-as években több kísérlet is történt a művészet és esztétikai kritériumainak formális megértésére, és ennek következtében a művészet szintetikus előállítására számítógépekkel (bár a számítógépek lassúak voltak, a kimeneti lehetőségek korlátozottak, és a számítógépes technológiák csak kutatóintézetekben és nagyvállalatoknál voltak elérhetőek).

6 Még több előfutár van, például a számítógépes klaszter, lamus, aki zenét komponál, és még egy albumot is kiadott (lásd: <http://melomics.uma.es/>). A szlogen - nem meglepő módon - "zene mindenkinek, mindennek", a művészet demokratizálásáért igérvé. Egy másik fontos példa lenne az Aaron, amely egy olyan szoftverrendszer, amely feltalálójával, Harold Co- hen-nel együttműködve festményeket készít (lásd: <https://www.computerhistory.org/atchm/harold-cohen-and-aar-on-a-40-year-collaboration/>).

7 Elég problematikus, hogy a "Computers and Creativity" (McCormack/d'Iverno 2012) című, egyébként kiváló esszégyűjteményben egyetlen olyan hozzászólás sincs, amely a kötet központi kérdését a művészeti rendszer szociológiai kérdéseivel próbálja összekapcsolni. A "művészettörténet" csak Frieder Nake kiemelkedő hozzájárulásában kerül néhányszor említésre -

Nake tanulmányát leszámítva a kötet legtöbb hozzászólása de-szocializálja és de-történetesíti a gépek lehetséges kreativitásának kérdését.



4. ábra: Piet Mondrian, *Kompozíció vonalakkal* (1917)

Michael 1967, Noll "A számítógép mint kreatív médium" című esszéjében leír egy információs esztétikai kísérletet Mondrian festményével (4. ábra).⁸ Noll azt írja: "Az alkotás egy képet ábrázol, amiben az alkotó a festményt ábrázolja:

[Egy] kísérletet végeztek Piet Mondrian "Kompozíció vonalakkal" (1917) című művével és egy számítógép által generált képpel, amely pszeudorandom elemekből állt, de összkompozíciójában hasonlított a Mondrian-festményhez. Bár Mondrian nyilvánvalóan gondosan és rendezetten helyezte el a függőleges és vízszintes sávokat a festményén, a számítógép által generált képen a sávokat egy pszeudorandom számgenerátor alapján helyezték el, amelynek statisztikáit úgy választották ki, hogy megközelítsék a Mondrian-festmény sűrűségét, hosszát és szélességét. A két kép xerográfiai másolatát egymás mellett mutatták be a középiskolától a posztdoktori képzésig terjedő végzettségű alanyoknak 100; az alanyok egy nagy tudományos kutatólaboratórium populációjának meglehetősen jó mintavételét képviselték. Megkérdezték tőlük, hogy melyik képet részesítik előnyben, és azt is, hogy a pár közül melyik képet tartják Mondrian alkotásának. Az alanyok 59 százaléka a számítógép által generált képet részesítette előnyben; csak 28 százalékuk tudta helyesen azonosítani a Mondrian által készített képet. Általánosságban úgy tűnt, hogy ezek az emberek

8 Egyébként ez az egyik legkorábbi szöveg, amelyben a számítógépet médiumként írják le - a számítógép médiummá válása tehát visszanyúlik a művészet és az esztétika kérdéseihez.

a számítógép által generált kép véletlenszerűségét emberi kreativitással, míg a Mondrian-festmény rendezett sávok elhelyezése számukra gépiesnek tűnt. Ez a megállapítás természetesen nem von le Mondrian művészi képességeiből. Végül is az ő festménye adta az ihletet a számítógép által generált kép előállításához használt algoritmusokhoz, és mivel számítógépek évekkel50 ezelőtt még nem léteztek, Mondrian nem rendelkezhetett számítógéppel. (1967: 92)

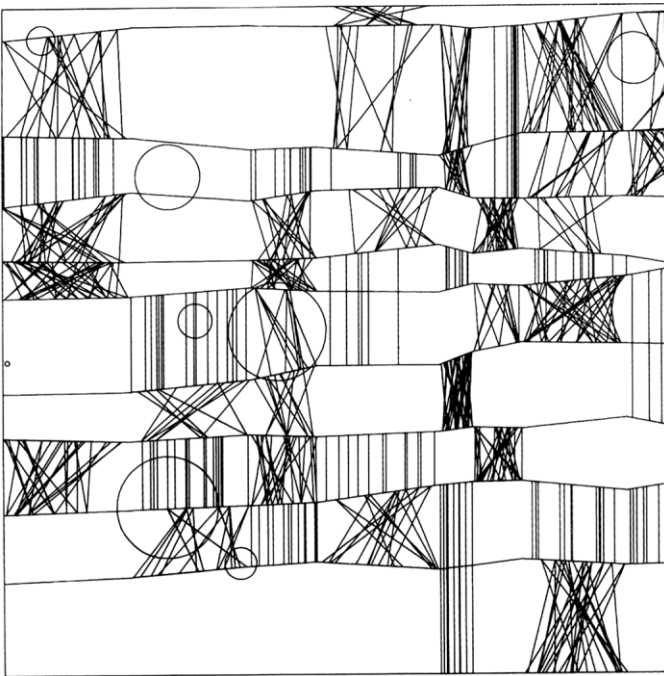


5. ábra: Michael Noll szimulált Mondrianja

Noll tehát egy Mondrian-képet szimulál (5. ábra) egy olyan statisztikai eloszlás alapján, amely feltételezhetően leírja a Mondrian-mű vonalainak elrendezését. És mivel ez az eloszlás "rendezettebbnek" tűnik, a megfigyelők egy - feltehetően nem reprezentatívnak tekinthető - csoportja a szimulált Mondriant azonosítja a valódival, míg a valódi túl szabályosnak és mechanikusnak tűnik.⁹ Ez a preferencia azt mutatja, hogy a rendezetlen képet magasabb esztétikai minőségben értelmezik? Ez ellentmondana Birkhoffnak, de háborzongató módon emlékeztet bennünket a disz-

⁹ Hasonló diskurzust találunk a legújabb mesterséges intelligencia művészetében is: "Elgammal szerint a hétköznapi szemlélők nem tudnak különbséget tenni egy AI által generált kép és egy "normális" kép között egy galéria vagy egy művészeti vásár kontextusában" (Bogost 2019).

a legújabb mesterséges intelligencia által generált portrék torz, elmosódott és "művészi" tulajdonságait. Vagy azért, mert a rendetlenséget egyszerűen az emberi szerzőségere való utalásként értelmezik? Az is kiderülhet, hogy egy adott kép művészi mivolta nem csak az érzékelési jellemzőktől függ... Megjegyzendő, hogy Noll olyan formai, algoritmikus szabályokat próbál találni, amelyek lehetővé teszik egy olyan műalkotás létrehozását, amely műalkotásként *azonosítható*. Nem Noll az egyetlen, aki ezt a megközelítést követi. Frieder Nake, az információs esztétika egyik legjelentősebb teoretikusa és művelője is megpróbálta nyomon követni a képképzés mintáját Paul Klee munkásságában, és így létrehozni egy számítógép által generált Klee-t (6. ábra).



6. ábra: Frieder Nake Klee-szimulációja

Igaz, ezek a példák kiragadva vannak a kontextusból, és az információs esztétika megfelelő megvitatásához legalább Max Bense, Rul Günzenhäuser és Abraham Moles álláspontját figyelembe kellene venni, amire ebben az esszében nem térhetünk ki. Mindenesetre mai szemmel nézve ezek a kísérletek meglehetősen furesának tűnnek, mert a műalkotásokat kivonják történelmi kontextusukból, és formalizálható absztrakt struktúrákká redukálják őket - hasonlóan a közelmúlt mesterséges intelligenciával előállított művészetéhez. Hangsúlyozni kell tehát, hogy Mondrian és Klee esztétikai stratégiái bizonyos történelmi kérdésekre reagáltak, nem utolsósorban magára a művészettörténetre. Az információs esztétika és a mesterséges intelligencia művészete úgy tűnik, hogy a

műalkotásokat ahistorikus, formális struktúrákként értelmezi - bár ezek a műalkotások is tekinthetők válaszoknak az alábbiakra

a művészet mint olyan bizonyos történelmi, ebben az esetben számítástörténeti kontextusát (vö. a forma és a történelem bonyolult kapcsolatára többek között Buchloh 2015). Ez a közelmúltbeli mesterséges intelligencia művészetében is újra megjelenni látszik: "Ez talán a mesterséges intelligencia művészet elkerülhetlensége: A művészettörténeti kontextus széles rétegei absztrahálódnak általános, vizuális mintázatokká". (Bo- gost 2019)¹⁰

Ezek a kísérletek azonban viszonylag jellemzőek egy nagyobb fejlődésre, nevezetesen a kognitív munka formalizálására és lehetőség szerint részleges vagy teljes gépekre való átvitelére irányuló kísérletekre, amelyre a számítógépek mint szimbólumfeldolgozó gépek alkalmasak.¹¹ A számítástechnika területéről származó korai szövegek, mint Douglas Engelbart "Program on Human Effectiveness" vagy J. C. R. Licklider "Man-Machine Symbiosis" című műve az 1960-as évekből, programadóak erre. (vö. Licklider 1960: 4-11; Engelbart 1991: 235-244). Noll, Nake és mások az esztétikai munkát próbálják formalizálhatóvá és elvileg gépek által végrehajthatóvá tenni. Próbálkozásaik - akár szándékosan, akár nem - az esztétikai munka automatizálásához való hozzájárulásként értelmezhetők.

Az áruterelés más formáiban ezek a folyamatok általában hozzájárultak az áruk olcsóbbá válásához - és így a "demokratizálódáshoz" abban az értelemben, hogy az áruk egyre több ember számára válnak megfizethetővé. Az információs esztétikának nyilvánvalóan nem ez volt a célja, hiszen - például - Nake a mű authoraként működött (és nem Klee); ha pedig Klee-nek nevezte volna a művét, az komoly jogi problémákhoz vezethetett volna. De még ha a művészetnek ez a demokratizálása lett volna is az információs esztétika célja - mondjuk Mondrianok és Klee-k ipari, olcsó tömeggyártásának beindítása -, ez nem működött volna a művészeti rendszerben: A művészet nemcsak formai struktúrák és stratégiák kérdése, hanem történelmi és különösen társadalmi helyek és szerepek kérdése is. Most erre térünk rá.

III. Művészet, tudás és munka

Úgy tűnik, a művészetet nem fenyegeti közvetlenül a számítógépesítés. A művészet nem szerepel a sokat vitatott oxfordi kutatási jelentésben, amely ideges vitát indított el a munka eltűnéséről: az egyetlen tevékenység, amely hasonlít a művészeti gyakorlathoz, a "művészeti igazgató" tevékenysége, aki meglehetősen könnyen megússza a 95. helyet a számítógépesítés valószínűségi listáján. (vö. Frey/Osborne 2013: 59.) A művészek is-

10 Ugyanez igaz Schmidhubers hozzájárulására McCormack/d'Iverno (2012). Ez ismét egy olyan megközelítés, amely egy tárgy esztétikai értékét formalizálja - anélkül, hogy igazán feltenné a kérdést, hogy ez az érték nem a tárgy relatív történelmi helyzetéből és nem (csak) formai, belső struktúrájából származik.

11 Vö. az automatizálás történetéről Noble (1984).

nem válnak a racionalizálás tárgyává, és műveiket továbbra is "műveknek" fogják nevezni. A művészi munka olyan műtípusnak tűnik, amelyet nem lehet vagy nem szabad formalizálni, algoritmizálni, és következésképpen nem is lehet minden további nélkül mechanikusan reprodukálni. Ez azt a gyanút sugallja, hogy a művészi munka valójában nem is munka, hanem a tevékenység egy másik formája vagy - legalábbis - a munka egy másik formája. De miért is?

Legalábbis első pillantásra a művészeti piac pontosan úgy néz ki, mint bármely más piac: a művészek tisztában vannak azzal, hogy munkájukkal pénzt kell keresniük. Néhányan szinte elképzelhetetlen mértékben sikerrel járnak, de a legtöbbjüknek komoly nehézségekbe ütközik a művészeti piacon való versenyben. A művészek ráadásul ki vannak téve annak, hogy a munka "kiadásaik refinanszírozásának szükségszerűsége" (Luhmann 1994: 191). Munka, pénz a saját munkáért, piac, verseny, gazdagok kontra szegények - ez az első pillantás azt sugallja, hogy a művészeti vállalkozás semmiben sem különbözik más termelési formáktól. Nem a szabadság birodalma, hanem csak egyfajta szolgáltatás vagy fogyasztási cikkek ipara, amely egy speciális piacot szolgál ki.

Ezért találhatunk olyan eseteket, amikor a munkát technológiailag szuperfluxussá teszik a művészeti szakmában. Ha megnézünk egy nagy stúdiót, például az Olafur Eliasson Stúdiót, megfigyelhetjük, hogy az új, számítógéppel támogatott technológiák bevezetése közvetlenül a munka eltűnéséhez vezet. Egy sor munkakör, például a weboldal karbantartása, a menedzsment, a PR, a logisztika, egészen a stúdiót takarító emberekig, helyettesíthető. Mivel az Olafur Eliasson Stúdió is kapitalista körülmények között működik, a munka racionalizálásával valószínűleg költségeket takarít meg. Ez az aspektus azonban külsődleges; az információ és a generatív esztétika által jelzett problémát kezeli, elsősorban azt, hogy *maga* a művészi munka, a művészetcsinálás racionalizálható. Bár Olafur Eliasson egyetlen kidolgozott projektjét sem tudná megvalósítani csapata nélkül, úgy tűnik, hogy a csapatmunka eltűnése a *szervó* nevének fekete dobozában nem érinti a művészetet "önmagában". És amennyiben a csapatmunka feketedobozba kerül, megváltoztathatja összetételét anélkül, hogy bármit is változtatna Eliasson művészetének "művészi mivoltán".

Ezzel végül elérkeztünk az információs esztétika által felvetett probléma lényegéhez. Nyilvánvaló, hogy az emberi szerző vagy okozó közreműködése nélküli művészet gondolata - még akkor is, ha szerepe éppen abban áll, hogy demonstratívan visszavonul - nem tűnik számunkra hihetőnek. Nem látjuk a természetben megjelenő művészetet.¹² Az információs esztétikában és a mesterséges intelligencia művészetében implikált kérdés eltér Walter Benjamin híres esszéjének, a *Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit*-nak a megfigyeléseitől, ahol minden a *termék*, a *műalkotás* technikai reprodukciójának kérdése körül forog. Az információs esztétikában és a mesterséges intelligencia művészetében azonban a *műalkotást létrehozó mű reprodukálhatóságának* kérdése a központi kérdés. Az a tény, hogy a mű technikai reprodukálása nem tűnik lehetségesnek.

- 12 Bár a művészetet történelmileg néha a természethez közel állónak értelmezték ("Kunstschönes" vs. "Naturschönes"), vö. erről Kant (1914 [1790]).

nem kell a "zsenialitásnak" tulajdonítani - ez egy nagyon hagyományos válasz lenne -, amely végső soron isteni eredetű, és így önmagában véve nem technikai jellegű. Luhmann megjegyzi: "A művész zsenije elsősorban a teste" (2000: 38). Egyszerűen azt mondhatnánk tehát, hogy a tudásnak a dolgozó testtől való elválasztása, amely a kapitalizmus fejlődésére jellemző, és amelyet talán Marx fedezett fel először, a művészetben nem valósul meg, illetve nem valósulhat meg. De miért is?

Az egyik lehetséges ok az lehet, hogy a műalkotás - az információs esztétika (legalábbis Birkhoff szerint) a komplexitás pontos formalizálására tett kísérletei ellenére - *túlságosan összetett*, és így az azt létrehozó mű nem érthető meg kellőképpen. Észrevehető tehát, hogy Noll és Nake a festészet egy bizonyos típusára összpontosít, amelyet az alapvető geometriai formák széleskörű igénybevétele határoz meg. Az ilyen formák elég egyszerűnek tűnnek ahhoz, hogy formalizálásukat sugallják, míg más, összetettebb formák elkerülnék őket. Nyilvánvaló, hogy a technológia fejlődött: Ma már nem elegendők a geometriai minták a komputáló rendszerek kreativitásának demonstrálására, hanem (bár elmosódott) portrékra van szükség. A portrék történelmileg a "zsenialitás" történetéhez kötődnek. Ahogy az elején megjegyeztük, az ilyen portrék aligha jelentik a mai¹³ művészet status quo-ját - a gépek figuratív művészet létrehozására való felhasználása kevésbé arról szól, hogy a mesterséges intelligencia kritikusan reflektál a kortárs művészetre (legalábbis amikor Nake és Noll művei készültek, az absztrakt festészet eléggé központi szerepet játszott), sokkal inkább a számítógépes grafika új képességeiről.

De visszatérve az auctorialitás kínos kérdésére: Még geometriai formák, igaz, hogy nem igazán tudjuk elképzelni a művészetként való létezésüket művészet nélkül. Hiszen még ha egy művész - mint például Nake - pontosan úgy határozná is meg magát, hogy minden munkát gépekre bíz, az eredményt akkor is "Nake művének" neveznénk, ami hasonló ahhoz a már említett megállapításhoz, hogy a munkamegosztással működő műteremben folyó termelés "feketedobozba" kerül egy *szelő* neve alá. Hasonló folyamat látszik lejátszódnia az *Edmond de Belamy portréjával* kapcsolatos vitában is, hogy ki a szerző végső soron.

A műalkotás egy olyan munka eredménye, amelyben a test és a tudás, vagyis az a tudás, *hogyan kell ezt a konkrét művet létrehozni*, nem választható szét - és ez azt jelenti, hogy a "szerző" funkciója központi jelentőségű. (vö. Graw 2012: 43-45.) Legalábbis ez az az ideológiai figura, amely történelmileg a művészeti rendszerre jellemzőként jelent meg. Ezért a műalkotások nem lehetnek a tudomány és *egy hamis test* közötti kapcsolat - *ezt* neveznénk hamisításnak. Még ha például Donald Judd egyik művét vennénk is, amely Sebastian Egenhofer (2008: 214) szerint "feloldódik az ipari diszpozíció anonimitásában", akkor is értelmetlen lenne, ha egy másik személy vagy egyszerűen egy vállalat a készítés módjának ismeretében újra legyártaná ugyanazt a tárgyat, ahogy ez elvileg a reprodukciók ipari gyártásában is

történik - nem lenne lehetséges, hogy

13 Talán a "posztmodern" neo-figuratív festészet néhány típusának kivételével.

hogy ezt a reprodukciót (Donald Judd) műalkotásként ismerje el. Vagy vegyünk egy másik példát: Elaine Sturtevant kölcsönkérte Warholtól a *Virágok* szitanyomó matricáit, és újra kinyomtatta a *Virágokat*, sőt 1991, egy egész kiállítást készített a *Warhol Virágokból - és Warholt, a Virágok gyártási folyamatára utalva, gyakran idézik, amikor azt mondta: "Nem tudom. Kérdezd meg Elaine-t."* (Idézi: Arning 1989: 44.) Mindazonáltal Sturtevant Warhol tudásának kisajátítása nem Warhol munkájának racionalizálása abban az értelemben, hogy Sturtevant most egyszerűen "olcsóbb Warholokat" készít, hanem inkább "Sturtevanokat". Diederich Diederichsen rámutat arra, hogy a kurátorok, kritikusok, közönség stb. által végzett relevanciakijelölő munka is a műalkotást létrehozó munkához és annak piaci értékéhez tartozik. (vö. Diederichsen 2012: 99) Ennek a munkának az "elosztott" jellege lehetetlenné teszi a racionalizálását - mivel folyamatosan kíséri a művet, azaz soha nem ér véget, és a jövőben is kiszámíthatatlan fordulatokat vehet. A "Sturtevanok" a jövőben fontosabbak és drágábbak lehetnek, mint a "Warholok". Azt is mondhatnánk, hogy Sturtevant a kisajátításával eleve láthatóvá teszi Warhol szerzői funkcióját.

Természetesen lehet próbálkozni az előállítási költségek megtakarításával, de nincs értelme olcsóbban kínálni a Warhol's *Flowers-t*, mert csak a Warhol's Factory *virágait* fogadják el eredetinek, ami persze nem zárja ki, hogy a *Flowers* olcsó reprodukcióit poszterként (amelyek nem számítanak műalkotásnak, csak reprodukciónak) állítsák elő. Warhol élete 1987-ben ért véget, és ezzel megszűnt az eredeti "Warholok" gyártása - és ez szükségszerű: Hosszú távon a művészek halandósága miatt a műalkotások szűkössé válnak, és ezért van piaci értékük (vö. a szűkösség fogalmáról kritikusán Panayotakis 2012). Akapitalizmusban senkinek sem lehet érdeke a műalkotások "demokratikus" előállítására.

Nyilvánvalóan a döntő különbség az, hogy a művészetben a reprodukciók, mint Sherrie Levine újrafényképezései vagy Sturtevant ismétlései, mindig eredetiek (amelyek az ismétlés révén az eredeti diskurzusát mutatják). Ezzel szemben a "hagyományos" árutermelésben nincsenek eredetiek,¹⁴ hanem csak újratermelések sorozatai, pl. *Virágok* plakátjai, amelyek mind az eredeti műalkotásokra utalnak, de sorozatban mind azonos rangúak. És még egyszer: eredetinek lenni annyit jelent, mint a művész testéhez kapcsolódni. *Ez a mű nem választható el a testtől, ezért nem lehet formalizálni és racionalizálni.* Az információs esztétika próbálkozásai értelmetlennek tűnhetnek, mert itt az ipari termelésből származó racionalizációs diszpozíciót egy olyan területre helyezik át, amely ezt blokkolja. Vagy ahogy Adorno fogalmaz:

14 Nehéz esetnek tűnik a termékhamisítás, mivel a kifejezés már magában foglalja az "eredeti terméket" - de még az eredeti termékek is azonos termékek sorozataként léteznek. A termékhamisítás azt jelenti, hogy egy termék tévesen azt állítja, hogy az adott sorozathoz tartozik. Ez eltér az eredeti és a másolat viszonyától a művészetben.

Másfelől azonban, amikor az autonóm művészet komolyan nekilátott az ipari folyamatok befogadásának, azok mindig kívül maradtak az ipari folyamatokon. [...] A művészet radikális industrializálódása, a megvalósult technikai normákhoz való töretlen alkalmazkodása ütközik azzal, ami a művészetben ellenáll az integrációnak. (1997 [1970]: 217).

De hogy tisztázzuk a dolgot: Ez nem azt jelenti, hogy a művészet önmagában a szabadság utópisztikus birodalma, amely megszabadult az ártermelés korlátaitól és szabályaitól. Ez csak azt jelenti, hogy az ártermelésnek egy olyan területe, amelynek más, történelmi, folyamatos szabályai vannak, és ezért a művészet bizonyos értelemben gazdaságilag kivételes (vö. Beech 2015).

IV. Következtetés

Úgy tűnik, hogy arra a kérdésre, hogy a gépek lehetnek-e kreatívak vagy sem, nem lehet ontológiai választ adni. Az, hogy mit jelent a "kreatív", túlságosan történelmileg folyamatosnak és alakíthatónak tűnik. Gondoljunk csak azokra az állításokra, amelyek szerint a gépi tanulás miért nem kreatív: Bogost (2019) kitarít emellett, hogy "minden gépi tanulási technikának egy meghatározott tréningkészletre kell alapoznia". Ez igaz - de nem vonatkozik ez az emberi művészekre is, akiknek az érzékelési és pl. festői képességeiket kell betanítaniuk? Vagy ismét Bogost (ibid.): "Egy neurális háló semmit sem tudna kikövetkeztetni a reneszánsz vagy az antikvitás sajátos szimbolikus jegyeiről - hacsak nem tanítják meg rá, és ez nem történne meg pusztán azzal, hogy sok portrét mutatunk neki.". Nem igaz ez az emberekre is? És még tovább: A "kreativitás" nem mindig oszlik meg az emberi és a nem emberi szereplők között? Nem mondják-e gyakran az emberi művészek, hogy a műveik "válaszolnak", miközben a festés folyamatában vannak? Lehet-e művésznek lenni mindenféle nem-emberi anyag és közvetítő nélkül, amelyek nem csupán átlátszó eszközök egy előre megadott "vízió" számára (vö. Hensel/Schröter 2012)? És nem lehetne-e akár fejlett mesterséges intelligenciákat is elképzelni, talán humanoid robotok formájában, amelyek művészek lehetnek (vö. Kjosen 2012 egy hasonló, az érték munkaelméletével kapcsolatos érvelésért)? Nem lehetne-e az általuk létrehozott műveket ugyanúgy a "testi" jelenlétükhöz kötni, mint az emberi művészek esetében? Gondoljunk csak az olyan virtuális popsztárookra, mint Miku Hatsune¹⁵, aki egyfajta hanghordozóként működhet, egyfajta "aláírással" (itt: a hangjával) rendelkezve - de természetesen a művész halandósága mint egyfajta természetes hiányosság, amely korlátozza a művet, ilyen esetben nem adott (a virtuális sztár fogalmáról lásd Schröter 2000).

Bár jelenleg úgy tűnik, hogy a gépek által előállított műalkotások gondolata történelmietlen és abszurd, mert a műalkotásoknak helyük van a történelemben, és a művész testéhez kell kötődniük, ezért ritkák és így tovább, lehet, hogy a távoli jövőben a dolgok megváltozhatnak: Talán a művészet

demokratizálódásához más társadalmi kontextusra van szükség - ahogyan a mesterséges intelligencia demokratizálódásához is, kiszakítva azt

15 https://en.wikipedia.org/wiki/Hatsune_Miku.

a nagy monopóliumok kezéből. Ezért is tűnik annyira helyénvalónak, hogy a sokat vitatott AI-festmény Edward Bellamy portréja - hiszen Bellamy írta 1888-híres regényében, a *Looking Backward 2000-1887*. Az AI úgyszólván visszatekintett egy történelmi festészeti módra, történelmi stílusban és kissé elavult gesztusokkal (az aláírás), de ez lehet egy lehetséges jövőből való visszatekintés metaforája is, amelyben az AI művész lehet. Bellamys regény központi témája a jövő egy egészen másfajta gazdasága, egyfajta posztkapitalizmus, amely valóban előfeltétele lehet az AI mint művészet demokratizálódásának. A regényben pedig több olyan futurisztikus médiumról is szó esik, amelyek manapság persze abszurdnak tűnnek - de amelyek egy olyan jövőbeli medialitás metaforáiként is olvashatók, amelyben még a "kreativitás" is legalább részben automatizálódhat (ezzel ellentétes nézetet lásd Kelly 2019).

Bibliográfia

- Adorno, Theodor W. (1997 [1970]): Adorno Adorno: Esztétikai elmélet. Újrafordítva, szerkesztve és Robert Hullot-Kentor fordítói bevezetőjével, London: Athlone Press.
- Arning, Bill (1989): "Sturtevant." In: Kortárs Művészeti folyóirat 2. kötet, 2. szám, 39-50. oldal.
- Beech, Dave (2015): Művészet és érték. Művészeti gazdasági kivételesség a klasszikus, neoklasszikus és marxista közgazdaságtanban, Leiden és Boston: Brill.
- Bogost, Ian (2019): "Itt van az AI-Art aranyláza. Egy mesterséges intelligencia 'művész' önálló kiállítást kapott egy chelsea-i galériában. Vajon újra feltalálja a művészetet, vagy tönkreteszi?" In: The Atlantic március 3. (<https://www.theatlantic.com/technology/archive/2019/03/ai-created-art-invades-chelsea-gallery-scene/584134/>).
- Buchloh, Benjamin H.D. (2015): In: "Formalizmus és történetiség": Ibid., Formalizmus és történetiség. Models and Methods in Twentieth Century Art, MA/London: Cambridge, pp. 1-87.
- Diederichsen, Diedrich (2012): Diederichsen: "Zeit, Objekt, Ware." In: In: Texte zur Kunst No. 88, pp. 95-101.
- Egenhofer, Sebastian (2008): Abstraktion Kapitalismus Subjektivität. Die Wahrheitsfunktion des Werks in der Moderne, München: Fink.
- Engelbart, Douglas C. (1991): "Bushnak és az Emberi hatékonyságról szóló programnak" In: James M. Nyce/Paul Kahn/et al. (szerk.), From Memex to Hypertext. Vannevar Bush and the Mind's Machine, Boston/MA: Academic Press, pp.235244.
- Frey, C. B./Osborne, M. A. (2013): "A foglalkoztatás jövője. Mennyire fogékonyak a munkahelyek a számítógépesítésre. Resource Document", Oxford Martin School, Oxfordi Egyetem.

(http://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf).

- Graw, Isabelle (2012): "Der Wert der Ware Kunst. Zwölf Thesen zu menschlicher Arbeit, mimetischem Begehren und Lebendigkeit". In: *Texte zur Kunst* No. 88, pp. 43-45.
- Hensel, T., & Schröter, Jens. (2012): "Die Akteur-Netzwerk-Theorie als Herausforderung der Kunstwissenschaft". In: *Ibid.* (szerk.), *Die Akteur-Netzwerk-Theorie als Herausforderung der Kunstwissenschaft*. Schwerpunktherausgeber-schaft der Zeitschrift für Ästhetik und Allgemeine Kunstwissenschaft No. 57, pp. 5-18.
- Kant, Immanuel (1914 [1790]): *Kant: Az ítélet kritikája*. Fordította bevezetéssel és jegyzetekkel J. H. Bernard, második kiadás, London: MacMillan.
- Kelly, Sean Dorrance (2019): "Egy filozófus azt állítja, hogy egy mesterséges intelligencia nem lehet művész. A kreativitás emberi törekvés, és mindig is az lesz.", MIT Technology Review, 2019.2.21. (<https://www.technologyreview.com/s/612913/a-philosopher-argues-that-an-ai-can-never-be-an-artist/>).
- Kjøsen, Atle (2012): "Álmodnak-e az androidok értéköbbletről?" (https://www.academia.edu/2455476/Do_Androids_Dream_of_Surplus_Value).
- Licklider, J.C.R. (1960): In: "Ember-Computer szimbiózis" (Man-Computer Symbiosis): *IRE Transactions on Human Factors in Electronics* Vol. HFE-1, No. pp1,411.
- Luhmann, Niklas (1994): "Luhmann, Luhmann: Kapitalismus und Utopie. In: *Merkur* 48. kötet, 540. szám, pp. 189-198.
- Luhmann, Niklas (2000): *Stanford: Luhmann Luhmann: A művészet mint társadalmi rendszer*: Stanford University Press.
- McCormack, Jon/d'Iverno, Mark (2012) (szerk.): *Heidelberg: Computers and Creativity*, Heidelberg: Springer.
- Mersch, Dieter (2019): "Kreativität und Künstliche Intelligenz. Einige Bemerkungen zu einer Kritik algorithmischer Rationalität". In: *Zeitschrift für Medienwissenschaft* 21, megjelenés alatt.
- Noble, David F. (1984): *F. Nobel F. Nobel, Nobel F. Nobel: A termelés erői. A Social History of Industrial Automation*, New York: Knopf.
- Noll, Michael (1967): Noll: "A számítógép mint kreatív médium". In: *IEEE Spectrum* Vol. 4, No. 10, pp. 89-95.
- Panayotakis, Costas (2012): "Theorizing Scarcity. Neoklasszikus közgazdaságtan és kritikái" In: *Szemle a radikális politikai közgazdaságtanról* 45. kötet, 2. szám, pp. 183-200.
- Payne, Stewart (2007): "A szürkehályog a kulcs Monet elmosódott stílusához". In: *The Telegraph* May (<https://www.telegraph.co.uk/news/uknews/1551703/Cataracts-the-key-to-Monets-blurry-style.html>)16.
- Reckwitz, Andreas (2012): *Die Erfindung der Kreativität. Zum Prozess gesellschaftlicher Ästhetisierung*, Berlin: Suhrkamp.

Schröter, Jens (2000): Lara Croft: "Lara Croft. Funktionen eines virtuellen Stars."

In: Ulrike Bergermann, Hartmut Winkler (szerk.), TV-Trash. The TV Show I Love to Hate, Marburg: The TV Show I Love to Hate: Schüren, pp. 123-144.

- Schröter, Jens (2019): Schröter: "Digitale Technologien und das Verschwinden der Arbeit". In: Thomas Bächle, Caja Thimm (szerk.), Die Maschine: Freund oder Feind?, Wiesbaden: Springer VS, pp. 183-210.
- Sternberg (1999) (szerk.): Handbook of Creativity, Cambridge: Cambridge University Press.
- Sudmann, Andreas (2019): "KI-Phantasien: Kommt jetzt der Terminator mit Pinsel?" In: CCB Magazin február 8. (<https://www.creative-city-berlin.de/de/ccb-magazin/2019/2/8/andreas-sudmann-ki-forschung/>).

"Ez egy orwelli 1984 jövő a küszöbünkön, igaz?"

Természetes nyelvfeldolgozás, mesterséges neurális hálózatok és a mesterséges intelligencia (demokratizálásának) politikája

Andreas Sudmann beszélgetése Alexander Waibelrel, a Karlsruhei Technológiai Intézet informatika professzorával, aki egyben a Carnegie Mellon Egyetem Számítástechnikai Iskolájának professzora is.

Andreas Sudmann: Alex, te vagy az egyik úttörő a mesterséges neurális hálózatok (ANN) és a természetes nyelvfeldolgozás (NLP) területén. Mi volt a kezdeti motivációd, hogy belépj erre a kutatási területre?

Alexander Waibel: Waibel: Mintegy negyven éve dolgozom a tudományos életben. Németországban születtem, német szülőkkel, és az MIT-n, majd a Carnegie Mellonon tanultam, a számítástechnika és a mesterséges intelligencia vezető egyetemmein. Ezekben az intézményekben alakítottam ki munkám fő irányvonalát és inspirációját, vagyis azt a kérdést, hogy hogyan tanulunk és kommunikálunk emberi lényekként, és hogyan tudunk olyan technológiát építeni, amely segíthet az emberi kommunikáció javításában. Akkoriban, az 1970-es években, amikor én egyetemre jártam, az emberek már akkor is gondolkodtak a mesterséges intelligenciáról, hiszen a terület első definícióit az 1950-es években javasolták. A Nobel-díjas Herbert Simon, aki tagja volt a szakdolgozati bizottságnak, részt vett az intelligens gépek építéséről szóló híres dartmouthi konferencián 1956-ban, ahol ő és más kutatók meghatározták a mesterséges intelligencia korai vízióját.

Akkoriban mindenki intelligens gépeket próbált építeni kereső algoritmusok, szabályok és logikai formulák segítségével. De nekem, mint diáknak, ez egy kicsit olyan volt, mint "des Kaisers neue Kleider" (a császár új ruhái). Hallgattam ezeket a híres embereket, akik egy olyan problémáról beszéltek, amely számomra soha nem lenne megoldható a mesterséges intelligencia szabálya alapú megközelítésével. Számomra intuitíve világos volt, hogy az életünk során megtanulható tudás és tények mennyisége olyan hatalmas, hogy mindezt szabályokba foglalni lehetetlen lenne. És ami még rosszabb, ezeket állandóan módosítani kellene, mert a körülöttünk lévő világ állandóan változik. Valójában ez teljességgel lehetetlen, ezért már a kezdetektől fogva aggodalommal töltött el,

hogy azt mondjam: Soha nem fogjuk elérni ezeket a célokat, hacsak nem fejlesztünk ki olyan tanuló gépeket, amelyek képesek maguktól megszerezni ezt a tudást.

Ezen az alapvető tudományos törekvésen kívül azonban egy másik dimenzió is mindig is fontos volt számomra mint tudósnak: Még ha ez most talán közhelyesnek is hangzik - ez a folyamatos vágy volt, hogy hozzájáruljak a társadalomhoz és jobbá tegyem a világot, szemben a saját személyes törekvéseim vagy érdeklődési köröm követésével. Kutatóként nem egyszerűen a kíváncsiság vezérel, hanem gyakorlati célok vezérelnek. A gyakorlati célok módját adnak az előrehaladás értékelésére, és pozitív hatással lehetnek a társadalomra, ha sikeresek vagyunk. És ezek közül talán az egyik legkövetkezetesebb célom a munkám során, mint külföldiként, aki nyelveken nőtt fel, az volt, hogy olyan gépeket építsek, amelyek segíthetnek nekünk, embereknek a nyelvek közötti fordításban, szöveges vagy beszélt nyelvi fordításban. A történelem során számos kísérlet történt arra, hogy gépeket használjanak szövegek fordítására, ami önmagában is elég nehéz. De amikor az embereket nyelvi korlátok között próbáljuk összekötni, akkor a *beszélt* nyelvet is le kell fordítanunk. A 70's-ben ez abszurd célnak tűnt, sőt, megoldhatatlannak tűnt, mivel a beszéd a probléma összetettségének és bonyolultságának egy teljesen új dimenzióját adta, mivel a beszéd szöveggé alakítása (a fordítás előtt) önmagában is megoldatlan mesterséges intelligencia-probléma volt. Nem tudtuk, hogyan ismerjük fel a beszédet, és ami még rosszabb, az emberek soha nem beszélnek tiszta szöveget... beszéd közben hibáznak, dadognak, tévóváznak és javítgatják magukat beszéd közben. És hogyan kombinálnád ezt a másik nehéz AI-problémával, a fordítással? És mindez együtt olyan nehéz volt, hogy a mesterséges intelligencia korai időszakában elképzelhetetlen volt a megvalósítása. De számomra megszületett az álom, és fiatalos naivitással és optimizmussal belevágtunk. Mondanom sem kell, hogy ez egy nehéz probléma volt és maradt, egy olyan probléma, amelyen a mai napig dolgozunk. De az akadályok, nehézségek és késedelmek ellenére az út során láthattuk erőfeszítéseink gyümölcsét. Visszatekintve, valójában kiváltság, hogy az emberiségnek *abban a* generációjában élünk, amely a nyelvi korlátok eltűnését látja, és hogy lehetőségünk volt arra, hogy az ezt lehetővé tevő technológiákon dolgozzunk.

A tudományos siker kulcsa a gépi tanulásban elért fejlődésnek köszönhető. módszerek, valamint az ezeket támogató rendelkezésre álló számítási teljesítmény és adatok robbanásszerű növekedése. Ahhoz azonban, hogy az elképzelés valósággá váljon, az akadémiai fejlődést át kellett ültetni a társadalmi alkalmazásba. Ennek érdekében több olyan vállalatot indítottunk, amelyek viselhető beszéd- és nyelvtechnológiák, végül pedig mobil beszédfordítók létrehozására specializálódtak. Egyikük, a Jibbigó, megépítette és eladta az első telefonra szerelt Dialog fordítót. Ezt az App Store-on keresztül értékesítették, és segített a turisták és az egészségügyi dolgozók kommunikációjában. A céget később felvásárolta a Facebook, és mi tovább dolgoztunk a még fejlettebb telepítéseken. Most például olyan új tolmácseszközöket fejlesztünk, amelyek segítenek a migránsoknak Németországban kommunikálni az orvosokkal, ha nem beszélnek a nyelvet. Egyetemi környezetben, a KIT-en automatikus

szinkrontolmács-szolgáltatást telepítettünk, hogy a külföldi diákok Németországban tanulhassanak, és az előadás alatt szinkrontolmácsolás segítségével követhessenek egy német nyelvű előadást.

A karlsruhei csapatom és én is végeztünk korai kísérleteket az Európai Parlamentben, hogy kiderítsük, vajon egy ilyen technológia segítséget nyújthat-e ebben a legnehezebb nyelvi környezetben. Tehát végső soron nem csak a fordításról van szó, hanem arról, hogy hogyan tudunk olyan technológiát létrehozni, amely képes áthidalni a korlátokat, amely összehozza a világot, és lehetővé teszi, hogy az emberek jobban megértsék egymást. És ahhoz, hogy ezt a nehéz feladatot megoldjuk, olyan gépeket kell építenünk, amelyek több különböző szinten képesek hatékonyan tanulni.

A szükséges alapokat pedig már az 1980-as és 1990-es években kidolgozta, különösen az úgynevezett TDNN modellekkel kapcsolatos kutatásai révén. Talán mesélne egy kicsit arról, hogyan jutott el ehhez a bizonyos megközelítéshez, és elmagyarázná, hogyan működik?

Amikor a Carnegie Mellonon írtam a doktori disszertációm, lenyűgözött az ötlet, hogy olyan tanulási algoritmusokat építsek, amelyek jobban utánozzák a tömegesen párhuzamos, holisztikus tanulást, amelyet emberként az agyban végzünk. Felfedeztem, hogy a kutatók már a 1950/80-as években javasoltak úgynevezett "per- ceptronokat", amelyek ugyan tanultak, de csak nagyon egyszerű osztályozási feladatokat tudtak megoldani. Mégis, az a tény, hogy ezeket a funkciókat valóban meg lehet tanulni, nemcsak izgalmas volt, hanem közvetlenül alkalmazhatónak tűnt azokra a homályos és többértelmű nyelvi és észlelési problémákra, amelyeken dolgoztam. Ismétlem, ez egy olyan időszak volt, amikor az emberek azt hitték, hogy a beszéd felismerést és a nyelvi fordítást elsősorban szabályok segítségével lehet megoldani, ami számomra abszurdnak tűnt, tekintve a tények és részletek óriási mennyiségét, amelyeket össze kellett volna rakni. Mindazonáltal az egyszerű perceptronok és hasonló módszerek a beszéd felismerés esetében is komoly korlátokat jelentettek, mivel egyszerre csak egyetlen neuront lehetett betanítani. Ezzel szemben az agy egész varázsa abban rejlik, hogy nem egyetlen neuront, hanem neuronok egész hálózatait képzi ki, és hogy a neuronok egy *együttese* sokkal nagyobb teljesítményű feladatokat képes elvégezni. De hogyan képeznénk ki egy teljes hálózatot?

Éppen ebben az időben történt, hogy egy fiatal professzorasszisztens véletlenül Geoffrey Hinton a Carnegie Mellonra jött, és elkezdett dolgozni a Boltzmann-gépeknek nevezett valamin. Amíg ott volt, sok hasznos beszélgetést folytattunk, és ő mutatott be nekem valamit, amivel a USC San Diegóban dolgoztak, egy backpropagation nevű algoritmust. Ez sokkal közelebb állt ahhoz, amit én kerestem, és azonnal ráugrottam. A backpropagation egy egyszerű algoritmus volt, az egyszerű perceptron kiterjesztése - azzal a különbséggel, hogy ez az algo- ritmus most a perceptronok egész hálózatát optimalizálná, és gondoskodna arról, hogy az optimálisan működjön. Ha megmondjuk az egész hálózatnak, hogy mit kell tennie, akkor valójában minden egyes belső neuront úgy tud beállítani, hogy mindegyikük megpróbálja hozzájárulni ahhoz, ami a neuronok egész együttese számára a legjobb. Ez nagy lépésnek tűnt a jó irányba, nagy előrelépésnek a minták osztályozása felé, de a beszéd és a nyelv

esetében ez nem volt elég. Mivel a legtöbb valós világbeli problémában nem a minták felismerése az egyetlen probléma, hanem a megtalálni a

mintázat, amelyet elsősorban osztályozni kell. Ez mindig azt jelentette, hogy először szegmentálni kellett a jelet, hogy megtaláljuk az érdekes mintákat (hangokat, képeket), mielőtt azokat osztályozni lehetett volna. A beszédben a beszédet úgy kellene vágni, hogy azonosítani lehessen egy adott fonéma elejét és végét, és ha ez megvan, akkor megpróbálhatunk egy neurális hálózatot alkalmazni a hangok osztályozására és beszédsequenciává való összeállítására. Ez azonban több különálló, nehéz probléma megoldását jelentette. És a kemény tanulság a beszédben az volt, hogy ez halálos, mivel mindegyik hibázik. Ezért világossá vált számomra, hogy olyan neurális hálózatra van szükségünk, amely nemcsak csodálatos osztályozó, hanem a helyzettől független mintákat is felismeri, ezt a tulajdonságot nevezzük "shift invariance"-nak. Mit is jelent ez? Azt jelenti, hogy olyan neurális hálózatot építünk, amelyet nem csak egy adott mintára alkalmazunk, hanem az összes neuront a bemenet egy tartományában mozgatjuk, és hagyjuk, hogy lényegében addig pásztázzák a bemenetet, amíg fel nem világít, amikor hasznos, hasznos mintát talál. Az ilyen egységekből álló hálózatok így megtanulhatnak, hogy a jelben bekövetkező kis elmozdulásoktól függetlenül összeállítsák az összes hasznos evidenciát. Az ilyen eltolódás-invariancia természetesen szükséges a beszédhez, mivel a beszéd folyamatosan áramlik és változik, de mint kiderült, a mesterséges intelligencia számos más problémájához is szükséges, beleértve a képeket, a zenét, a játékokat, a nyelvet és még sok más problémát.

Mindezekben a helyzetekben az első kihívás az, hogy tudd, *hol van* a hasznos pat-
ternek, mielőtt helyesen osztályozhatná őket. Ezért a dolgok osztályozása a
dolgok eltolódás-invarianciával történő de- te- rálásával volt az a kulcsprobléma,
amit meg kellett oldanunk. Ezt a célt szem előtt tartva mentem Japánba
poszt doktorandusznak, ahol az akkori legerősebb szuperszámítógépekhez volt
hozzáférésem. És ezzel a számítási kapacitással lehetőségem nyílt egy új modell
kifejlesztésére, amely az időkéleltetéses neurális hálózat (TDNN) néven vált
ismertté. Ez még mindig egy többrétegű ("mély") neurális hálózat volt, de most
már kifejezetten váltásinvariáns osztályozásra volt kiképezve. És mint kiderült,
fantasztikusan jól működött; jobban működött, mint az összes többi, akkoriban
létező módszer.

Ez az új TDNN modell tehát felváltotta a többi módszert?

Sajnos még mindig nem volt meg a szükséges számítási teljesítmény ahhoz, hogy
elég nagy hálózatokat építsünk ki. Akkoriban Japánban a rendelkezésre álló
legnagyobb szuperszámítógépeket használtuk, és összehasonlítottuk őket más
statisztikai vagy szabályalapú módszerekkel összehasonlító adatokon - és azt
tapasztaltuk, hogy sokkal, de sokkal jobbak. Amikor azonban megpróbáltunk
nagyobb hálózatokat és gyakorlati beszéd felismerő rendszereket építeni, még
mindig számítási korlátokba ütköztünk, és sok kompromisszumot kellett
kötnünk, ami ártott a teljesítménynek, így más kutatók egyszerűbb módszerekkel
fokozatosan felzárkózhattak, és hasonló vagy akár valamivel jobb teljesítményt
értek el, mint mi. Ennek következtében - és ez az 1990-es évek végén, a 2000-es

évek elején történt - az emberek elvesztették érdeklődésüket a neurális hálózatok iránt, és egyszerűen más statisztikai módszereket használtak.

Tíz év tel el, és csak kevesen foglalkoztak tovább a neurális hálózatokkal, egészen 2008 körül, amikor - inkább véletlenül - az Egyesült Államokban különböző emberek újra megpróbálkoztak ezekkel a régi neurális hálózatokkal, de sokkal nagyobb számítási kapacitással és sokkal nagyobb mennyiségű, az interneten keresztül elérhető tanulási adatokkal. És mint kiderült, ezek a már az 1980-as években kifejlesztett neurális hálózati módszerek hirtelen elképesztően jobban működtek, mint bármely más megközelítés a mesterséges intelligencia területén. És nem csak egy *kicsit* működtek jobban, hanem 30 százalékkal jobban. A mi területünkön, tudják, egész PhD-dolgozatokat írnak, ha fél százalékos előrelépést érnek el; tehát ha valami 30 százalékkal jobb, az egyszerűen forradalmi. Ennek eredményeként az egész közösség két éven belül átállt a neurális megközelítésekre. A másik dolog, ami meglepetésként ért, az az, hogy mi történik, ha több réteget adunk hozzá a hálózatához. Az 1980-as években egy vagy két úgynevezett rejtett réteggel rendelkezünk, és ez volt minden, amit ki tudunk számítani. De most, az új számítási teljesítménynek köszönhetően három, négy, öt és még több réteget is létrehozhatunk. Senki sem számított arra, hogy ez tovább javítja majd a teljesítményt, de így történt. Ma már 40, 80 vagy akár több száz rétegből álló hálózatokat használunk beszéd felismerésre a laboratóriumunkban. És azok az izgalmas neurális modellek, amelyek 20 évvel ezelőtt olyan jól működtek, ma még jobban működnek. A TDNN-ek továbbhaladtak, és képfeldolgozásra, játékokra, beszédre és más problémákra is alkalmazták őket, és általánosabb néven: "konvolúciós neurális hálók" néven váltak ismertté. Ma már a legtöbb modern mesterséges intelligenciamotor szívében megtalálhatók.

A nagy teljesítményű hardverekhez és nagy mennyiségű adathoz való hozzáférés szempontjából bizonyára hasznos volt, hogy egy ideig a Facebooknak dolgoztam.

Igaz, két évig a Facebooknál voltam igazgató, de természetesen sok barátom dolgozik a Google-nél, az Amazonnál és a Microsoftnál is. Sok diákunk dolgozik most a Google-nél, az Amazonnál, a Microsoftnál és így tovább. És sokan közülük a mi laborjainkban végeztek. Az a hatalmas adatmennyiség, amelyet ezek a vállalatok ellenőriznek, természetesen kincsesbánya a tanulási programok számára. Ezekben a nagy internetes cégekben hatalmas neurális hálózatokat képeznek hatalmas mennyiségű adaton hatalmas mennyiségű számítási teljesítményt használva, és a teljesítménynövekedés még mindig növekszik. És ez meglepő és lenyűgöző. De ha megkérdezné tőlem, hogy mi az új áttörés a mesterséges intelligenciában ma az évekkel 2000-évihez képest, azt kell mondanom: nem sok minden. Ezek megint csak nagyjából ugyanazok a hálózati technikák és képzési algoritmusok, amelyek az 1980-as években dolgoztunk, csak most nagyságrendekkel több adatot és nagyobb számítási teljesítményt használunk, és ezek valójában sokkal, sokkal jobban működnek, mint azt valaha is elképzeltük.

Eddig elsősorban a beszédfelismerés, a gépi fordítás és az ANN technológiai aspektusairól beszéltünk. Talán most már beszélhetnénk ezeknek a modelleknek és a mesterséges intelligencia alkalmazásainak politikai dimenziójáról is. Ön szerint melyek a leglényegesebb politikai szempontok a természetes nyelvfeldolgozás területén általában, illetve a beszédfelismerés és a gépi fordítás területén különösen?

Erről persze sok mindent lehetne mondani. Az egyik fontos szempont a kutatásfinanszírozás politikája lenne, amely közvetlenül érint minket a tudományos tevékenységünk szempontjából. A tudományos támogatás politikai tényezőktől függ, és ezek néha működnek, néha nem. És valóban lenyűgöző különbségek vannak az Egyesült Államok és Németország, vagy az európai és ázsiai országok között, mert minden ilyen ország vagy kultúra másképp közelíti meg a tudományos támogatást, és ezért sajátos erősségei és gyengeségei vannak. Ezért ezen a szinten politikailag aktív vagyok, hogy jobb európai kutatástámogatási mechanizmusokat vezessek be és javítsak. A másik politikai dimenzió azonban az, hogy mit teszünk a kutatásainkkal. Mint már említettem, a világról alkotott képem szerint szeretek olyan projekteket végezni, amelyekkel a társadalom valamely aspektusát próbálom javítani. És mint már mondtam, a kutatás számomra mindig is azt jelentette, hogy az emberek jobban megértik egymást. Ha a kutatásainkkal olyan gépeket tudunk építeni, amelyekkel jobban tudunk kommunikálni, akkor ez azt jelenti, hogy kevesebb félreértés lesz.

Pályafutásom során több céget is alapítottam, és az egyik ilyen cég egy kézi beszédfordítót épített telefonra. Ez volt az első mobil telefonos beszédfordító rendszer. Ezt az 2009. induló céget "Mobile Technologies"-nak hívták, a termék neve pedig "Jibbigo" volt. Beszélni lehetett a telefonba, és a rendszer lefordította a bemenetet egy másik nyelvre. Hatalmas siker volt. Az Apple például reklámokat készített vele. Mindenhol használták, és az emberek visszajöttek hozzánk, és azt mondták: "Végre megértem az apósomat!", és "Most már tényleg megértem a többi embert!". És humanitárius projekteket is indítottunk, például rendszereket építettünk, mondjuk thai és khmer nyelven, hogy az amerikai, európai vagy japán orvosok segíthessenek a vidéki embereknek az egészségügyi ellátásban, és hasonló dolgokat telepítettünk Dél-Amerikában is.

Az ilyen típusú technológia iránti széleskörű érdeklődés miatt a vállalatot a Facebook (a világ "nyitottá és összekapcsoltá" tétele) vette fel, és 2013 két évig vezettem egy tudósokból álló csapatot, amely a fordítási technológiát fejlesztette. A Facebooknál azonban kiderült, hogy a technológia felhasználása a bejegyzések fordítására és más vállalati felhasználási esetekre nagyobb prioritást élvez, mint a beszédfordítóink interaktív kommunikációs aspektusa, amelyet szerettem volna előmozdítani, ezért visszatértem az egyetemre, hogy folytassam a technológia oktatási és humanitárius vonatkozásain végzett munkánkat.

A Cambridge Analytica és más botrányok miatt a Facebookot egyre gyakrabban éri masszív kritika, ezért a technológiai óriásvállalatra még nagyobb nyomás nehezedik, hogy megfeleljen idealista programjának. Ugyanakkor az olyan vállalatok, mint a Microsoft vagy az OpenAI a mesterséges intelligencia demokratizálását követelik. Mi az általános véleménye erről a koncepcióról?

Ismétlem, a világ sokkal összetettebb, mint azt egy egyszerű szlogen sugallja. A Facebook jó példa erre. Biztos vagyok benne, hogy a kezdeti céljuk a hírek demokratizálása volt. Ha bárki posztolhat híreket, milyen csodálatos lehet ez! Ha bárki tényeket közölhet a Wikipédián, milyen csodálatos lenne az! Nincs többé szakértő, aki diktálja a véleményét, igaz? De ha igazán végiggondoljuk a dolgot - ha bárki bárkiről bárkit közzétehet, és világszerte elérheti a közönséget -, akkor az előny nem feltétlenül az, hogy olyan emberek is hallatják a hangjukat, akiket korábban nem hallottak, hanem az egész világra kiterjedő visszaélési és manipulációs lehetőséget is nyitnak. És pontosan ez az, amire most rájövünk. A demokratizálódás tehát rendben van. De ugyanebben a folyamatban ugyanúgy ott van a tömeges manipuláció és visszaélés lehetősége, és ezért nagyon óvatosnak kell lennünk.

Más szóval: Többé-kevésbé szkeptikus vagy ezzel a koncepcióval kapcsolatban?

Ismétlem, óvatosnak kell lennünk, és folyamatosan át kell gondolnunk, hogy mit teszünk, mert az olyan eszközök, mint az internet, olyan erősek. Néha olyan dolgokat hozunk létre, amelyeknek nem szándékolt következményei vannak. És újra kell értékelnünk a technológiát, és törekedni kell arra, hogy jó irányba mozdítsuk el. Miközben az internet az információ demokratizálódásához vezetett, most ismét az információ és a hatalom masszív koncentrációját látjuk. Inkább egy olyan világot szeretne, amelyben csak a Google, a Microsoft, a Facebook, az Amazon vagy az Apple rendelkezhet intelligens rendszerekkel, és mindenki más ki van szolgáltatva nekik e technológia használata tekintetében? Szeretne egy olyan világot, amelyben csak az egyik nagy techóriás képes gépi úton felismerni bárki arcát, és senki más nem képes rá? Ezek a technológiák gyakorlatilag monopóliumokat ösztönöznek, amelyek hihetetlen mennyiségű adatot hordoznak és rengeteg tudást generálnak, ugyanakkor - a legjobb szándék ellenére is - rengeteg lehetőséget biztosítanak a manipulációra. A közelmúltban láthattuk, hogy ez a Cambridge Analytica-botrány esetében is így van. A kérdés tehát ismét az, hogy szeretné-e, ha az összes adat és a mesterséges intelligencia a világban mindössze három-négy vállalatnál összpontosulna?

Ezek az aggodalmak geopolitikai szinten is megjelennek. Miközben az internetet nagy globális egyesítő erőnek tervezték, most azzal fenyeget, hogy különböző erkölcsi és társadalmi hozzáállású, az elsőségért versengő nagy regionális szférákra szakad. Kínában, ahol kevesebb törvény vagy korlátozás vonatkozik az adatgyűjtésre és -kezelésre, azt látjuk, hogy az AI táplálja a kormány által felügyelt, automatizált tömeges megfigyelő állam kialakulását. Vajon ez - a verseny révén - aláássa a magánélet, a szabadság és a függetlenség

nyugati értékeit? Ez egy orwelli 1984 jövő a küszöbünkön, nem igaz? A demokratizálás legalábbis szélesebb körben terjeszti el a technológiát, és ezért van az, hogy a trösztellenes

erőfeszítések olyan fontosak, belföldön. De ha a globális egyensúlyról beszélünk - például Európa és Amerika között: Európának nincs nagy internetes vállalata, és ez aszimmetriákat teremt, ahol az egyik kontinens információ- és adatkezelésében döntően függ a másik kontinens mesterséges intelligencia-rendszereitől. Ez még mindig működik, mert Amerika és Európa között a kapcsolatok barátságosak és támogatóak, mivel mindkettő nyugati demokrácia. De mi a helyzet Kínával, Oroszországgal, Indiával? Kína drámai előrelépéseket tesz a mesterséges intelligencia terén, és dollármilliárdokat költ a mesterséges intelligenciára. Tehát csak idő kérdése, hogy Kína és mások is egyenrangúak, ha nem fejlettebbek lesznek az USA-nál. Európa pedig saját világos technológiai bázis és jövőkép nélkül kiszolgáltatottá válhat annak, amit más szereplők csinálnak, és sokkal sebezhetőbbé válhat a külső beavatkozással és manipulációval szemben. Számomra ezek aggasztó fejlemények.

Mindazonáltal nem mindegy, hogy amerikai vagy kínai technológiai monopóliumokról beszélünk.

Igen, az úgynevezett GAFA-t [Google, Apple, Facebook, Amazon] bizonyos értelemben másként kell kezelni. Kínában az olyan vállalatok, mint a Baidu, az Alibaba vagy a Ten- cent nagyon erős kormányzati kapcsolatokkal rendelkeznek. Az Egyesült Államokban más a játéktér, mert ott a nyilvánosság figyeli ezeket a vállalatokat, és ha valamelyikük visszaél az adathatalmával, az botrányba fullad, és azonnal bejárja a híreket, ami nagyon rossz a vállalatnak. Amíg a Facebooknál voltam, azt kell mondanom, hogy - minden botrány ellenére - lenyűgözött, hogy a vállalat mennyire igyekezett felelősségteljesen kezelni az adataikat. És az a tény, hogy még mindig botrányokat produkálnak, egyszerűen azt mutatja, hogy milyen nehéz ezt megtenni, és hogy ez mennyire érzékeny kérdés. De úgy gondolom, hogy az amerikai vállalatok is elfogadják a mesterséges intelligencia demokratizálásának gondolatát, mert ez az üzleti modelljük része. Természetesen a vállalatok végső soron nagyon önzőek, és azt próbálják tenni, ami a legjobb számukra. De az amerikai kontextusban az adatok védelme jót tesz az üzletnek, mivel a botrányok félelmetesek. Ezért a Microsoft és a Google azért foglalkozik a mesterséges intelligencia demokratizálásával, mert ez támogatja üzleti stratégiájukat. Vegyük például az Amazont: Az egyik legnagyobb üzletága az Amazon Web Services (AWS), amely többek között magában foglalja a csomópontok bérlését, így egy kis cég az Amazon szerverein végezheti a számításokat. A Microsoft vagy a Google hasonló számítási erőforrásokat biztosít, és ha ezen felül még AI-szolgáltatásokat is tudnak nyújtani, akkor az nyilván az ő üzletüknek is jót tesz.

Mit tartana a közeli jövő legfontosabb politikai kihívásainak a mesterséges intelligencia területén, az Ön, mint informatikus szemszögéből?

Úgy gondolom, hogy mindenekelőtt nekünk, tudósoknak kötelességünk ébernek lenni. De optimizmusra ad okot, hogy valójában sok idealista ember dolgozik ezekben a nagyvállalatokban. Így az, hogy vannak botrányok, jó hír, mert ez azt jelenti, hogy nem lehet titokban tartani az ilyen dolgokat, hogy a társadalmat arra kényszeríti, hogy továbbgondolja ezeket a dolgokat, ami jó. És ami azt illeti, hogy a politikának hogyan kellene reagálnia: Nos, ha megnézzük néhány szenátori vitát, rájövünk, hogy a politikusok nem lehetnek mélyen benne minden technológia minden aspektusában, és ezért nem biztos, hogy megérzik, hová vezethet, és hogyan kell reagálni. Emiatt úgy gondolom, hogy nagyon fontos, hogy a mesterséges intelligenciával foglalkozó tudósok hangosan és aktívan részt vegyenek a nyilvános párbeszédben, hogy mi (a tudomány, a közvélemény és a kormányok) biztosíthassuk, hogy ezeket a technológiákat az emberiség szolgálatába állítsuk, és ne a saját politikai vagy pénzügyi érdekeinket szolgáljuk.

Ami azonban ebben az összefüggésben aggaszt, az az a tény, hogy a nagyvállalatok mohón vesznek fel tudósokat, és az egyetemek nehezen tartják meg a tehetséges embereket. És nem szabad elfelejtenünk, hogy az egyetemek a nyílt viták és viták színterei (vagy azoknak kellene lenniük), hogy ne manipuláljanak bennünket gazdasági vagy politikai érdekek. A lényeg tehát az, hogy erős akadémiai környezetet kell fenntartanunk minden olyan fontos területen, ahol a mesterséges intelligenciát alkalmazzák. Ez pedig Európában különösen nagy kihívás: Nagy internetes cégek nélkül természetesen a tehetségek folyamatos elvesztésétől szenved. A saját németországi AI-laboratóriumomra hivatkozva elmondhatom, hogy a legjobb tudósok közül sokan, amint végeznek a pályázatukkal vagy a diplomájukkal, olyan ajánlatokat kapnak, amelyek körülbelül hétszer magasabbak (vagy még magasabbak), mint amit mi megengedhetünk magunknak egy egyetemen. És amikor a fiataloknak ilyen összegű pénzt ajánlanak, plusz egy esélyt, hogy egy olyan nagy cégnél, mint az Amazon, az Apple vagy a Facebook építhessenek valamit, akkor ugranak a lehetőségre. Más szóval, az agyelszívás óriási, és nem csak az akadémiairól az iparba, hanem az országok között is. Ezért a mesterséges intelligencia korában Európának sokkal agresszívebben kell fellépnie, hogy gondoskodjon a jövőjéről.

Mit tehet Európa, hogy ez megváltozzon? És honnan tudják a politikusok, az emberek vagy a nyilvánosság, hogy mely szakértőkben bízhatnak?

Nos, ami az utolsó kérdését illeti, úgy gondolom, hogy kevésbé kockázatkerülő magatartással és a technológiai zavarok ösztönzésének fokozásával: Európának kiváló tudósai és mérnökei vannak. Európában az innovatív és alapvető tudományok végzéséhez is kiemelkedő támogatás és szabadság áll rendelkezésre. Európának nagyon okos, jól képzett és idealista tudósai vannak. Azt is állíthatnám, hogy az amerikai élvonalbeli tudósok közül sokan Európában kaptak képzést, vagy itt kezdték pályafutásukat. Ez nem csak egy üres frázis, számos híres példát tudnék mondani. De az a terület, ahol rosszul teljesítünk,

az az, hogy hagyjuk, hogy a tudományos fejlődés megkérdőjelezze a társadalom status quóját. Mire van szükség

a gyors, gyakorlatias és bomlasztó holdudvari projektek. Az amerikai DARPA például nagyon sikeresen végzi ezt a kormányzat számára. És olyan vállalatok is, mint a Google, a Tesla, az Amazon, amelyek 30 évvel ezelőtt még nem léteztek.

A másik dolog, amit javítani kellene, az a technológia átadása az ipari hasznosításba. Európában valójában sok vállalkozó indít vállalkozásokat. Ott vannak a kockázatvállalók, ott vannak a fiatalok, ott vannak a ragyogó ötletek, és ott van az izgalom és a buzgalom, hogy ezt megtegyük. Ami hiányzik, az az ilyen tevékenységek támogatásához szükséges tőke, valamint a nagyobb hajlandóság és gyorsaság a fúziók és felvásárlások terén. Az Egyesült Államokban például a kis cégeket nagyon gyorsan felvásárolják. Néhány vállalat csak tizenkét vagy tizennyolc hónapig létezik, mielőtt egy nagyobb vállalat bekebelezi őket. Ez egy egészséges folyamat, mivel felgyorsítja az átmenetet az ötlettől a koncepciótól a termékig és az iparáig. De Európában ez nagyon ritka. Itt nagyon nehéz a vállalatok felvásárlása. Hosszú időbe telik a tőzsdére lépés, a tőzsdére lépés és így tovább. A kis, sikeres startupból nagyvállalattá válás Európában olyan nagy súrlódásokkal jár, hogy sok lehetőség kimarad; a gyorsaság a lényeg az ilyen típusú játékban. És ez végső soron sok kisvállalkozást és fiatal vállalkozót az USA-ba és Kínába sodor.

Egy másik politikai-etikai aggály, amelyről manapság sokan beszélnek, az algoritmikus elfoglaltságok problémája. Hogyan kapcsolódnak ezek a problémák a természetes nyelvfeldolgozással és a beszélt nyelvek fordításával kapcsolatos kutatásaihoz?

Örülök, hogy visszavezet minket ehhez a témához. Eddig sokat beszéltünk arról, hogy a mesterséges intelligencia hogyan hat a társadalomra. Egy másik fontos politikai dimenzió az, hogy megvitassuk, hogyan választjuk ki azokat a projekteket, amelyek hozzájárulnak ahhoz a társadalomhoz, amelyben élni szeretnénk. Számomra ez a beszédfordítást jelenti, mert szerintem ez az egyik legnagyobb probléma az Európa-szerte. Az európaiak különböző 23 nyelveket beszélnek, és ezek csak a hivatalos nyelvek. Valójában sokkal több nyelv van Európában. És ez a helyzet szeparációt, félreértéseket és jelentős veszteséget okoz az üzleti lehetőségek terén. Az egyik fő ok, amiért az e-kereskedelem Európában nagyobb kihívást jelent, az a széttagoltság. Minden európai országnak más a jogrendszere, más a szállítási rendszere, és ezek nagy része természetesen nyelvileg is megkövesedett, mert ha mindent több nyelven kell intézni, az megnehezíti a transznacionális üzleti cseréket. De azt mondani, hogy mindenkinek meg kellene tanulnia angolul, eszperantóul vagy valami hasonlót, nevetséges lenne, és nem is kívánatos. Valójában a nyelvek sokfélesége és változatossága olyasmiről szól, amire az európaiak joggal büszkék.

Ennek fényében a technológiát nem akadálynak, hanem óriási problémamegoldónak kell tekintenünk, ha olyan szöveg- és beszéd fordításra képes technológiát akarunk kifejleszteni, amely minden fronton áthidalja ezeket a nyelvi akadályokat, hogy valóban egy nyelvileg átlátható világot teremtsünk. Képzeld el, hogy Kínába, Oroszországba vagy Spanyolországba megyünk, és

úgy szeretnénk ezekben az országokban tevékenykedni, mintha otthon lennénk, a-

a nyelvi hátrányok kiküszöbölése. De ha végiggondoljuk, hogy mit jelentene egy ilyen scenárió, ha az adott országban az adott nyelv ismerete nélkül tartózkodnánk, milyen segítségre lenne szükség ahhoz, hogy ne vegyük észre a nyelvi korlátokat. És valóban, ennek elérése az a jövőkép, amin dolgozunk.

Életrajzok

Alexander, V. N. tudományfilozófus, *A biológus szeretője* című könyv szerzője: *Re- thinking Self-Organization in Art, Literature and Nature*. A Rockefeller Alapítvány Bellagio Centerének öregdiákja, a NY Council for the Humanities korábbi nyilvános ösztöndíjasa, a Third Way of Evolution csoport tagja, és a Dactyl Foundation igazgatója. Tagja a Biosemiotics folyóirat (Spring- er Publishing) és a Meaning Systems könyvsorozat (Fordham University Press) szerkesztőbizottságának. A sós evolúcióelmélettel kapcsolatos munkái a Fine Lines című folyóiratban jelennek meg: Vladimir Nabokov tudományos művészetek című könyvében, amelyet a Yale University Press adott ki. Alexander 2020-ban digitális humán tudományok Fulbright-ösztöndíjasaként az oroszországi Szentpéterváron fog dolgozni.

Beverungen, Armin (2019 októbere óta) a digitális kultúrák szervezésének junior professzora a lüneburgi Leuphana Egyetemen, és korábban a University of the West of England, a Leuphana Egyetem és a Siegeni Egyetem tanára volt. A *Spheres* című folyóirat alapító társszerkesztője: *Journal for Digital Cultures* (www.spheres-journal.org) és a Digital Cultures című könyvsorozat (meson press) társszerkesztője. 2019 nyarán a bochumi Center for Advanced Internet Studies ösztöndíjasa volt. Kutatásai a média- és a szervezéstudományok határterületein zajlanak, és jelenleg az algoritmikus menedzsment jelenségére összpontosít. Legutóbbi publikációi közé tartozik a *Markets* (Jens Schröterrel/Phil Mirowskival/Edward Nik-Khakkal, meson press és University of Minnesota Press), valamint az *Organization* szerkesztett száma "a digitális média szervezeti erői" témában (Lisa Conrad/Timon Beyes-szel).

Burkhardt, Marcus posztdoktori munkatárs a Siegeni Egyetem Digitális Me- dia és Módszerek Tanszékén. Kutatásai a digitális média történetével és elméletével, különösen az adatbázis-technológiák, a big data és az algoritmikus környezetek kogi(sti)csájával, valamint a tudásteremtés és -terjesztés médiumaival, a médiafilozófiával és a médiaelmélettel foglalkoznak.

Dippel, Anne a Friedrich-Schiller-Universitát Jena Kulturális Antropológiai és Kultúrtörténeti Tanszékének tudományos kutatója és előadója. Dolgozott a berlini Humboldt Egyetem Kép-Tudás-Gestalt Kiválósági Kustériumában, volt a berlini Institute for the Advanced Studies kutatója.

Médiakultúrák és számítógépes szimulációk (mecs) tanszékén, a Lüneburgi Leuphana Egyetemen, valamint az MIT Science, Technology and Society Programme (Tudomány, technológia és társadalom program) vendégprofesszora, és a terepmunkája idején a CERN együttműködés társult tagja. Dippel, A. (2017). *Das Big Data Game. Zur spielerischen Konstitution kollaborativer Wissensproduktion in der Hochenergiephysik am CERN*. In: *NTM 4/2017*, - 517485. Dippel, A. & Warnke M. (szerk.) (2017). *Interferenciák és események. Eπιστημικος elmozdulások a fizikában számítógépes szimulációk segítségével*. Lüneburg: Me-son Press. Dippel, A. Arbeit (2018). In: Markus Rautzenberg, Daniel Martin Feige, Michael Ostritsch (Hrsg.): *Philosophie des Computerspiels*. Stuttgart: Metzler Verlag, 123-148.

Djeffal, Christian a Münchener Műszaki Egyetem jog, tudomány és technológia professzora. A jog, a technológia és a társadalom kapcsolatával foglalkozik. Különösen az olyan új technológiák érdeklik, mint a mesterséges intelligencia és a dolgok internete. Interdiszciplináris kutatásai az alkotmányjoggal, a szabályozással és a szabványokkal foglalkoznak. Christian a berlini Humboldt Egyetemen doktorált "Static and Evolutive Treaty Interpretation: A Functional Reconstruction" címmel, amelyet a Cambridge University Press adott ki. Post-doc-ként az Alexander-von-Humboldt-Institute for Internet and Society "Global Constitutionalism and the Internet" nevű kutatócsoportját koordinálta.

Förster, Yvonne kutatóprofesszorként dolgozik a Shanxi Egyetemen, Kína, és filozófiát tanít a Lüneburgi Leuphana Egyetemen, Németország. Doktori fokozatát a jénai Friedrich-Schiller-Universität Jénában szerezte az *idő tapasztalatáról és ontológiájáról (Zeiterfahrung und Ontologie*, München: Fink 2012). Vendégprofesszorként esztétikát tanított a weimari Bauhaus Egyetemen, és a közelmúltban vezető kutatói ösztöndíjat kapott két felsőoktatási intézetben (*Media Cultures of Computer Simulation* a Leuphanán és *Cultural Sciences a Konstanzi Egyetemen*). Kutatási területe a technológia filozófiája, az esztétika, a fenomenológia és a divat mint művészet. További információért lásd: www.yvonnefoerster.com

Großmann, Jürgen, Dr.-Ing. a modellalapú fejlesztés, a modellvezérelt tesztelés, valamint a biztonságtechnika és a biztonságtesztelés szakértője. Jürgen Großmann emellett tapasztalatokkal rendelkezik az autóiipari szoftverrendszerek és alkalmazások, különösen az ITS-rendszerek tesztelése és modellezése terén. Széleskörű tapasztalattal rendelkezik az ezen a területen a megfelelőségi és interoperabilitási teszteléshez szükséges tesztpadok és tesztlaboratóriumok fejlesztésében. Jürgen Großmann számos szabványosítási tevékenységben szerzett tapasztalatot különböző szabványosítási testületek, többek között az OMG, az ETSI, az ASAM és az AUTOSAR számára.

Matzner, Tobias a "Média, algoritmusok és társadalom" professzora a németországi Paderborni Egyetemen. Kutatásainak középpontjában a digitális technológia, a kultúra, a politika és az etika metszéspontjai állnak. Legutóbbi publikációi közé tartozik: "A digitális technológia és a digitális társadalom: Az ember halott - éljen az algoritmus! Emberi-algoritmikus együttesek és liberális szubjektivitás. In: *Theory, Culture & Society* online először 2019. "Az algoritmusok etikájának és politikájának megragadása". In *A nagy adatok politikája és politikája*. Szerkesztette: Ann Rudinow Sætnan, Ingrid Schneider és Nicola Green. London: Routledge 2018.

McQuillan, Dan a londoni Goldsmiths Egyetem kreatív és szociális számítástechnika tanára. Kísérleti részecskefizikából szerzett doktori címét követően a tanulási zavarok és a mentális egészség, majd az emberi jogok területén dolgozott. Kutatási területe a mesterséges intelligencia társadalmi hatása és a radikális polgári tudomány gyakorlata. Legutóbbi publikációi között szerepel a "Népi tanácsok az etikus gépi tanulásért", az "Adattudomány mint gépi neoplatonizmus" és az "Algorithmic States of Exception". Nyilvános írásaiból válogatva <https://www.opendemocracy.net/en/author/dan-mcquillan/> oldalon olvasható.

Monea, Alexander a George Mason Egyetem adjunktusa, az angol tanszék és a kultúratudományi doktori program közös munkatársa. A számítás, az algoritmusok és a big data történetét és kulturális hatásait kutatja. Jelenlegi könyvprojektje azt vizsgálja, hogy a heteronormatív előítéletek hogyan épülnek be az internetes kommunikáció áramlását irányító képpalkotó algoritmusokba és tartalomszűrőkbe. Korábban a *Computational Culture and Cultural Studies ↔ Critical Methodologies* című folyóiratokban publikált tanulmányokat, amelyek a gráf-adatbázisok hatalmát vizsgálják, és amelyek a Google Knowledge Graph megvalósítását vizsgálják a webes szemantika leképezésére. Jeremy Packerrel közösen írt egy cikket a *The International Journal of Communication* című folyóiratban a "média genealógiáról", amely egy olyan módszertan, amely a technológiák megjelenését vizsgálja politikai és társadalmi hatásaikra összpontosítva. Legutóbbi publikációi: Monea, A. (2016). Graph Force: Retorikai gépek és a tudás *N-rizálása*. *Computational Culture*, 5. Monea, A. (2016). A különbség gráfolása: Numerikus közvetítés és a Google tudásgráfiájának esete. *Cultural Studies ↔ Critical Methodologies*, 16 (5), 452-461. Monea, A. & Packer, J. (2016). Média genealógia és a régészet politikája. *The International Journal of Communication*, 10 , 3141-3159.

Morin, Kevin a kanadai Nemzeti Tudományos Kutatóintézet doktorjelöltje. Munkája az innovációs ökoszisztémával foglalkozik a digitális területen, mint például a gépi tanulás és a mesterséges intelligencia. Az STS és a kritikai elmélet megközelítéseit vegyítve dolgozatában az igazolási rendszereket, a szereplők hálózatait, a diskurzusokat és a technológiai rendszerek digitális kultúráját írja le,

328 A mesterséges intelligencia demokratizálódása
mint például a mesterséges intelligencia fejlesztése.

Pasquale, Frank a big data, a mesterséges intelligencia és az algoritmusok jogát kutatja. Tanúskodott vagy tanácsot adott az Egyesült Államok Egészségügyi és Emberi Szolgálatok Minisztériumának, a képviselőház igazságügyi bizottságának és a Szövetségi Kereskedelmi Bizottságnak, valamint az Európai Bizottság főigazgatóságainak. A *The Black Box Society* (Harvard University Press, 2015) című könyv szerzője, amelyet kínai, koreai, francia és szerb nyelvre is lefordítottak. A könyv a hírnév, a keresés és a pénzügyek társadalmi elméletét dolgozta ki. Tagja volt az NSF által szponzorált Nagy Adat, Etika és Társadalom Tanácsának (Council on Big Data, Ethics, & Society). Számos egyetemen és nyilvános előadáson beszélt adatpolitikáról. Frank társszerzője volt egy közigazgatási jogi esetkönyvnek, és több mint 50 tudományos cikk társszerzője vagy szerzője. Társszervezője volt a "Unlocking the Black Box: Az algoritmikus elszámoltathatóság ígérete és korlátai a szakmákban" című konferenciát a Yale Egyetemen. Jelenleg a Robotika törvényei című könyvén dolgozik: *The Future of Professionalism in an Era of Automation* (szerződésben a Harvard University Press-szel).

Reutter, Lisa a trondheimi NTNU szociológia és politikatudományi tanszékének PhD-hallgatója, jelenleg a gépi tanulás alkalmazását és használatát kutatja a norvég közsférában. Emellett részt vesz a "Digitális infrastruktúrák és az állampolgárok szerepvállalása (DICE)" kutatási projektben.

Rieger, Stefan német nyelvet és irodalmat, valamint filozófiát tanult. Disszertációját a barokk adatfeldolgozásról és a mnemotechnika témakörében írta, habilitációs disszertációját pedig a média és az antropológia kapcsolatáról (*Die Individualität der Medien. Eine Geschichte der Wissenschaften vom Menschen*, Frankfurt/M. 2001). Jelenlegi kutatási területei a következők: Tudománytörténet, médiaelmélet és kulturális technikák. 2007 óta a bochumi Ruhr-Universität Bochum médiatörténetének professzora. Legutóbbi könyvkiadványai: *Bunte Steine. Ein Lapidarium des Wissens* (Berlin: Suhrkamp mit 2014, Benjamin Bühler) és *Die Enden des Körpers. Versuch einer negativen Prothetik*, Wiesbaden: Springer.

Roberge, Jonathan a kulturális és városszociológia docense az Institut National de la Recherche Scientifique-ben, ahol a digitális kultúra kanadai kutatói székét is betölti. Észak-Amerikában az első tudósok között van, akik kritikusan foglalkoznak az algoritmusok termelésével, ez a kutatási program *Algorithmic Cultures* című könyvében csúcsosodott ki (Routledge, 2016, németre fordította a transcript Verlag, 2017). Jelenleg a *The Cultural Life of Machine Learning* című kéziratot dolgozik, amely 2020 elején jelenik meg a Palgrave MacMillan kiadónál (Micheal Castelle-lel együtt). Legutóbbi publikációi: SEYFERT, R. et ROBERGE, J. (szerk.), *Algorithmic Cultures. Essays on Meaning, Performance and New Technologies*, London, Routledge, 2016.; ROBERGE, J. et MELANÇON, L., "Az algoritmikus kultúra King Kong-jának lenni mégiscsak kemény munka. Google's Re-

gimes of Justification and the Meanings of Glass", *Konvergencia. The International Journal of Research into New Media Technologies*, vol. no23, pp3,2017., 306-324.

Schieferdecker, Ina, Prof. Dr.-Ing., a berlini Fraunhofer FOKUS társigazgatója és a berlini Műszaki Egyetem nyílt elosztott rendszerek minőségmérnöki professzora. Emellett a berlini Weizenbaum Institute for the Networked Society, a német Internet-Institute igazgatója. Kutatási területei közé tartoznak a (városi) adatplatformok, a kritikus infrastruktúrák, valamint a szoftveralapú rendszerek megfelelése, in- teroperabilitása, biztonsága, megbízhatósága és tanúsítása. Az Association for Software Quality and Education (ASQF) elnöke, a Német Nemzeti Tudományos és Műszaki Akadémia (acatech), a Német Globális Változás Tanácsadó Testület (WBGU), valamint a Szövetségi Oktatási és Kutatási Minisztérium "Hightech Strategy 2025" című projektjének tagja.

Schneider, Martin A. fejlett tesztelési módszerekkel és technikákkal foglalkozik, és különböző fuzzing technikákon, biztonsági tesztelési mintákon és biztonsági tesztelési metrikákon alapuló modellalapú tesztelési technikákon dolgozik biztonsági szempontok tekintetében, valamint más módszerek, például a kockázatelemzés integrálásán a fuzz tesztelés hatékonyságának növelése érdekében. Ennek során innovatív biztonsági tesztelési módszereket fejleszt és értékkel különböző ipari területek keretében, különös tekintettel a komplex rendszerekre, a kiber-fizikai rendszerekre és a szolgáltatásorientált architektúrákra.

Schröter, Jens, Prof. Dr., 2015 óta a Bonni Egyetem médiatudományi tanszékvezetője. A Siegeni Egyetemen 2008-2015 között a multimédiás rendszerek professzora volt. 2008-2012 között a Siegeni Egyetem "Locating Media" elnevezésű doktori iskolájának igazgatója volt. 2012 óta a Siegeni Egyetem "Locating Media" DFG-graduális kutatóközpontjának tagja. 2010-2014 között (Prof. Dr. Lorenz Engell professzorral (Weimar) együtt) a "TV Series as Reflection and Projection of Change" című DFG-kutatási projekt igazgatója volt. A "Society after Money-A Dialogue" (Társadalom a pénz után - Egy párbeszéd) című kutatási projekt előadója volt (VW alapítvány; Dr. Stefan Meretz, Dr. Hanno Pahl-lal és Dr. Manuel Scholz-Wäckerle-vel együtt), 2016-2018. 4/2018 óta igazgatója (Anja Stöfflerrel együtt, Mainz) a DFG-kutatási projektnek "Van Gogh TV. Kritikai kiadás, multimédia-dokumentáció és hagyatékek elemzése" (3 év). 10/2018 óta a "Society after Money-A Simulation" kutatási projekt (VW alapítvány; Prof. Dr. Gabriele Gramelsbergerrel, Dr. Stefan Meretz, Dr. Hanno Pahl-lal és Dr. Manuel Scholz-Wäckerle-lel együtt) előadója (4 év). 2014. április/május: "John von Neumann"-ösztöndíj a Szegedi Tudományegyetemen, Hungary. 2014. szeptember: Vendégprofesszor, Guangdong University of Foreign Studies, Guangzhou, Kínai Népköztársaság.

2014/15 tél: Senior-ösztöndíj a "Media Cultures of Computer Simulation" kutatócsoportnál, 2017 nyár: IFK Bécs, Ausztria. 2018. tél: IKKM, Bécs: Senior-ösztöndíj.

Weimar. Legutóbbi publikációi: Post- monetär denken, Wiesbaden: (a "Project Society after Money" című projekttel együtt): Springer 2018; (a "Project Society after Money"-val együtt): Society after Money. A Dialogue, London/New York: Bloomsbury 2019; (Armin Beverungen, Philip Mirowski, Edward Nik-Khah-val együtt): Markets, Minneapolis/London: University of Minnesota Press és Lüneburg: Meson (sorozat: In Search of Media). Látogasson el a www.medienkulturwissenschaft-bonn.de / www.theorie-der-medien.de / www.fanhsiu-kadesch.de weboldalra.

Senneville, Marius az INRS urbanisztika mesterszakos hallgatója, ahol Montreal és Toronto mesterséges intelligencia ökoszisztémáival foglalkozik. Az STS és a vállalatirányítás tanulmányozása által egyaránt inspirált megközelítéssel az ipari és akadémiai tudósok közötti együttműködés különböző konfigurációira és azok hatására a kutatásra.

Spilker, Hendrik Storstein a média- és technológiaszociológia professzora a trondheimi NTNU szociológiai és politikatudományi tanszékén. Számos könyvet és cikket írt az internet és a digitalizáció kulturális politikájáról, többek között a "Digital Music Distribution: The Sociology of Online Music Streams" (Routledge, 2018), és jelenleg egy nagyszabású kutatási projektet vezet a "Digitális infrastruktúrák és az állampolgárok felhatalmazása (DICE)" témakörében.

Sudmann, Andreas dolgozott Göttingenben, Regensburgban, Bécsben és Berlinben, jelenleg pedig adjunktusként (*Privatdozent*) tanít és kutat a bochumi Ruhr Egyetem médiatudományi karán, ahol 2016-ban habilitált is. Legutóbb a marburgi Philipps Egyetem médiatudományi vendégprofesszora volt, majd a bécsi IFK kutatói ösztöndíját töltötte be. Jelenlegi kutatási területei közé tartoznak a mesterséges intelligencia (AI) médiaelméleti és történeti problémái, különösen a gépi tanulási módszerek, a populáris és a technikai/digitális vizuális média esztétikája és politikája, a szerialitás és a dokumentumfilm formái és folyamatai, a médiaelmélet és a médiakritika. Sudmann számos könyv és szerkesztett gyűjtemény szerzője a médiatudomány és a digitális kultúratudomány területén, köztük a *Computer Games as a Sociocultural Phenomenon* (a Palgrave Macmillan kiadónál jelent meg 2008-ban) és a *Digital Seriality* (az *Eludamos. Journal for Computer Game Culture* különszáma, 2014). Új könyve *Serielle Überbietung. Zur televisuellen Ästhetik und Philosophie exponierter Steigerungen* (*Serial Outbidding: The Televisual Aesthetics and Philosophy of Ostentatious Escalation*) 2017 októberében jelent meg a Metzler Kiadónál (Stuttgart). Christoph Engemannal együtt társszerkesztője a *Machine Learning* című antológiának. *Medien, Infrastrukturen und Technologien der Künstlichen Intelligenz* (Gépi tanulás. A mesterséges intelligencia médiumai, infrastruktúrái és technológiái) című kötetben.

Volmar, Axel a Siegeni Egyetem "Az együttműködés médiumai" (SFB 1187 Medien der Kooperation) nevű együttműködési kutatóközpontjának posztdoktori ösztöndíjasa. Mellon posztdoktori ösztöndíjas volt a McGill Egyetem Művészettörténeti és Kommunikációtudományi Tanszékén (2014-2016). Jelenlegi kutatási érdeklődése a médiatudományok, a tudománytörténet és a hangtudományok metszéspontjain helyezkedik el, és az audiovizuális távközlés történetére összpontosít. A *Klang-Experimente* című könyv szerzője. *Die auditive Kultur der Naturwissenschaften 1761- 1961* (Frankfurt/M: Campus, 2015) című könyvének társszerkesztője, valamint számos média- és hangtudományi különszám és gyűjteményes kötet társszerkesztője. Legutóbbi publikációi közé tartozik a *Format Matters. Theories, Histories, Practices* című kötet, amelyet Marek Jancoviccsal és Alexandra Schneiderrel közösen szerkesztett (Lüneburg: meson press, 2019).

Köszönetnyilvánítás

Ez a kötet az újonnan alapított "AI Critique" című könyvsorozat első kiadványa. Részben egy Bochumban tartott nemzetközi konferencia hozzászólásain alapul, 2018amelyet a Center for Advanced Internet Studies (CAIS) nagylelkűen támogatott.

Számos csodálatos kolléga vett részt ezen a konferencián, köztük Friedrich Balke, Michael Baurmann, Armin Beverungen, Gioele Barabucci, Marcus Burkhardt, Anne Dippel, Christian Djeffal, Till Heilmann, Timo Honkela, Lisa Gotto, Tobias Matzner, Christoph von der Malsburg, Roland Memisevic, Daniel Neyland, Ina Schieferdecker és Claus Pias.

E tudósok közül többen is írtak egy-egy esszét a kötethez. Később további kollégák is csatlakoztak a projekthez, köztük Dan McQuillan, Alexander Monea, Frank Pasquale, Stefan Rieger, Jonathan Roberge, VN Alexander és Jens Schröter.

Ez a projekt nem utolsósorban két ösztöndíjnak köszönhetően valósult meg, az egyiket a bochumi CAIS, a másikat a bécsi IFK adta. Nagyon hálás vagyok mindkét intézménynek, amiért lehetővé tették, hogy kiváló körülmények között kutathassak.

Őszinte köszönet Anna Tuschlingnak és Bernhard Dotzlernek a könyvsorozat elindításában tanúsított érdeklődésükért és elkötelezettségükért, valamint annak a sok kollégának és barátnak, akik az évek során segítettek nekem a mesterséges intelligencia kutatásában vagy az arról való különböző gondolkodásban: Friedrich Balke, Michael Baurmann, Regina Bendix, Ulrike Bergermann, Armin Beverungen, Shane Denson, Julia Eckel, Lorenz Engell, Christoph Ernst, Jens Eder, Oliver Fahle, Yvonne Förster, Lisa Gotto, Irina Kaldrack, Frank Kelleter, Peter Klimczak, Petra Löffler, Thomas Macho, Tobias Matzner, Alexander Monea, Roland Memisevic, Moritz Müller-Freitag, Till Heilmann, Kathrin Peters, Ben Peters, John Durham Peters, Claus Pias, Stefan Rieger, Simon Rothöhler, Jens Ruchatz, Gabriele Schabacher, Jens Schröter, Philipp Schweighauser, Florian Sprenger, Daniela Wentz és Sabine Wirth.

Legmélyebb hálámát családomnak, különösen feleségemnek, Anne-nek köszönöm támogatásukért és szeretetükért. Ezt a könyvet nekik ajánlom minden szeretetemmel.

