

## A jogállamiságtól a jogi szingularitásig

Simon Deakin / Christopher Markou  
Jogi Kar / Cambridge-i Egyetem

# TERVEZET CSAK ÉSZREVÉTELEK MEGTÉTELÉRE

*A matematikát úgy határozhatjuk meg, mint azt a tárgyat,  
amelyben soha nem tudjuk, hogy miről beszélünk, és azt sem,  
hogy igaz-e, amit mondunk.*

*III Bertrand Russell  
II Újabb munkák a matematika alapelveiről  
I 1901*

*Önmagában semmi sem számít. Ami  
számít, az az, hogy semmi sem "önmagában és  
önmagától".*

*Chuck Klosterman III  
Szex, drogok és kakaós  
puffancsok II  
2003 I*

### 1. A Minden Új Minden Hajnala

Mielőtt a legtöbbszörnek fogalmuk sem volt arról, hogy mit jelent a negyedik ipari forradalom,<sup>1</sup> a 2019-es davosi Világgaazdasági Fórum a "Társadalom 5.0" hajnalát hirdette Japánban.<sup>2</sup> Célja: egy "emberközpontú társadalom megteremtése, amely a gazdasági fejlődést a társadalmi problémák megoldásával egyensúlyba hozza egy olyan rendszer segítségével, amely nagymértékben integrálja a kibertér és a fizikai teret".<sup>3</sup> A mesterséges intelligencia (AI) és a különböző digitális technológiák felhasználásával a "Társadalom 5.0" azt javasolja, hogy az embereket megszabadítsa a következőktől:

...a mindennapi nehézkes munkát és olyan feladatokat, amelyekhez nem értenek különösebben, és az új értékteremtés révén lehetővé teszik, hogy csak azokat a termékeket és szolgáltatásokat nyújtsák azoknak az embereknek, akiknek szükségük van rájuk, akkor, amikor szükségük van rájuk, optimalizálva ezzel az egész társadalmi és szervezeti rendszert.

A japán kormány elismeri, hogy e jövőkép megvalósítása "nem lesz nehézségektől mentes", de a terv világossá teszi, hogy "szembe kíván nézni ezekkel a nehézségekkel, azzal a céllal, hogy a világon elsőként, kihívásokkal küzdő országgént, modellértékű jövőbeli társadalmat mutasson be".

---

<sup>1</sup> K Schwab, *A negyedik ipari forradalom* (Crown Publishing 2017).

<sup>2</sup> H Nakanishi, "A modern társadalom elérte határait. A társadalom 5.0 felszabadít minket" (*Világgazdasági Fórum*, 2019. január 9.) <<https://www.weforum.org/agenda/2019/01/modern-society-has-reached-its-limits-society-5-0-will-liberate-us>>.

<sup>3</sup> Kabinetiroda, Japán kormánya, "Társadalom 5.0" <[https://www8.cao.go.jp/cstp/english/society5\\_0/index.html](https://www8.cao.go.jp/cstp/english/society5_0/index.html)>.

Bár a Társaság Japánon kívülről is élvez5.0 támogatást,<sup>4</sup> a japán optimizmus jellemzője. a technológiai fejlődés lehetőségeiről. <sup>5</sup> Japán azonban nincs egyedül azzal, hogy látja, hogy a technológiai fejlődés...

A negyedik ipari forradalom technológiai új kormányzási rendszereket és "algoritmikus szabályozást" tesznek lehetővé.<sup>6</sup> Ez különösen igaz a számítások egy bizonyos típusára, a statisztikai technikák egy családjára, a gépi tanulás (Machine Learning, ML) nevű technikára,<sup>7</sup> amely központi szerepet játszik az óriási technológiai lehetőségek mérnöki jövőjében, amelyet a Társadalom példáz5.0.

Általában az ML "algoritmusok fejlesztését foglalja magában nagy adathalmazok statisztikai elemzése révén, amelyekben nagyszámú példát tartalmaznak.<sup>8</sup> A matematikai paraméterek iteratív beállítása és az adatok megtartása lehetővé teszi, hogy az algoritmus az adatok ismételt megismerése révén automatikusan frissüljön (vagy "tanuljon"), és optimalizálja a feladatban nyújtott teljesítményt. Kezdetben a technikákat a társas objektumok azonosítására alkalmazták, mint például az arcfelismerés esetében. Az ML és a kapcsolódó mélytanulási technikák (Deep Learning, DL) egymást követő áttörései és teljesítményugrása,<sup>9</sup> a mesterséges intelligenciába, mint az összetett társadalmi-technikai problémák univerzális megoldójába vetett hitet erősítették. A döntéshozatal gyorsaságának és hatékonyságának ígéretes növekedése, valamint a költségek és a bürokratikus terhek csökkentése termékeny talajjá teszi a közszférát számos, a mesterséges intelligenciát hasznosító "technika" számára. Ezek közé tartozik a LegalTech,<sup>10</sup> GovTech,<sup>11</sup> és a Reg- Tech<sup>12</sup> (a jogi, kormányzati és szabályozási technológia rövidítése), amelyek "intelligens" szoftveralkalmazások fejlesztését foglalják magukban a jogi, politikai és emberi döntéshozatali kontextusokban történő alkalmazáshoz.

A Társadalom 5.0 terv azonban nem a japán kormány *ex nihilo* alkotása. Sokkal inkább egy kialakulóban lévő ortodoxiát fogalmaz meg - amelyet a "techek" most kihasználnak -, miszerint a jog, a politika és a gazdaság alapvető társadalmi rendszereinek alkalmazkodniuk kell a "lényegében digitális" új módozatokhoz, vagy el kell pusztulniuk.

---

<sup>4</sup> C Dube és M Minevich, "Mapping AI Solutions in Japan's Society 5.0" (2018) IPSoft/AI Pioneers White Paper <[https://www.ipsoft.com/wp-content/themes/ipsoft\\_v2/images/v2/pdf/jps-wtppr.pdf](https://www.ipsoft.com/wp-content/themes/ipsoft_v2/images/v2/pdf/jps-wtppr.pdf)>.

<sup>5</sup> J West "Utópia és nemzeti versenyképesség a technológiai retorikában: The Case of Japan's Informa- **ti** Infrastructure" (1995) 12 The Information Society 3.

<sup>6</sup> K Yeung és M Lodge "Algorithmic Regulation" in K Yeung and M Lodge (eds) *Algorithm Regulation* (Ox- ford University 2019).

<sup>7</sup> vö. E Alpaydin, *Machine Learning* (MIT Press 2016); G Marcus és E Davis, *Rebooting AI: Building Artificial Intelligence We Can Trust* (Pantheon 2019) 42-48; D Spiegelhalter, *The Art of Statistics: Learning from Data* (Pelican 2019) 143-187.

<sup>8</sup> Spiegelhalter, *A statisztika művészete*, 144.

<sup>9</sup> I Goodfellow, Y Bengio és A Courville, *Deep Learning* (MIT Press 2016) 1-8.

<sup>10</sup> R Dale, "Industry Watch: Törvény és szórend: (2019) Natural25 Language Engineer- ing 1.

<sup>11</sup> IBM, "Digitális átalakulás: Reinventing the Business of Government" (2018) <<https://www-07.ibm.com/sg/manufacturing/pdf/manufacturing/Digital-transformation.pdf>>.

<sup>12</sup> DW Arner, JN Barberis és RP Buckley, "FinTech, RegTech and the Reconceptualization of Financial Regu- lation" (2016) University of Hong Kong Faculty of Law Research Paper No. 2016/035 <<https://ssrn.com/ab- stract=2847806>>; G Roberts, "Fintech spawns regtech to automate compliance with

szabályozások" (*Bloomberg*, 2016. június28) <<https://www.bloomberg.com/professional/blog/fintech-spawns-regtech-automate-compliance-regulations>>.

kormányzás". Ezt gyakran úgy idealizálják, mint ami egy "hipotetikus új államhoz" vezet, "egy kis intelligens maggal, amelyet a nagy adatok tájékoztatnak... és amely a kormányt (végre) egy valóban posztburokratikus "formáció-államhoz" vezet".<sup>13</sup> Egyesek, mint például Tim O'Reilly, azzal érvelnek, hogy a "rég" állammodell lényegében egy "automata", ahol a pénz bemegy (adó), a közjavak és közszolgáltatások pedig kijönnek (utak, rendőrség, kórházak, iskolák).<sup>14</sup> Ő és mások szerint itt az ideje, hogy "újrarendeljük a kormányzást a mesterséges intelligenciával".<sup>15</sup> most, hogy a technológiai változás "ellaposította" a kormányzatot.<sup>16</sup> a világot, erodálta az államhatalmat, és modelleket adott az állampolgárság és a terület szétválasztására.<sup>17</sup> Csak ha a kormányzatot "platformként" fogjuk fel, akkor lesz lehetséges a kritikus hálózati externáliákat hasznosítani, és biztosítani azt, amit Jonathan Zittrain "generativitásnak" nevez - a nyílt végű platformok, mint a Facebook vagy a YouTube hátborzongató képességét, hogy olyan lehetőségeket teremtsenek, amelyek túlmutatnak az alkotóik által elképzeltéken.<sup>18</sup> O'Reilly szerint a Big Tech "azáltal lett sikeres, hogy megváltoztatta az összes szabályt, nem pedig azzal, hogy a meglévő rendszerben játszott".<sup>19</sup> Szerinte a kormányoknak világszerte követniük kell a példájukat. A Japánból, Szingapúrból, Észtországból és máshonnan származó bizonyítékok azt mutatják, hogy sokan már így is tesznek.<sup>20</sup>

Az egyre inkább "intelligens" és adatvezérelt kormányzás felé való elmozdulás már folyamatban van, és nem látszik, hogy ez a tendencia megállna.<sup>21</sup> A mámorító "új kormányzás illata" és az olyan programok techno-utópisztikus víziói, mint a Társadalom 5.0, azonban nem terelhetik el a figyelmet arról a kritikus kérdéstről, hogy mit is jelent pontosan a "techno-szabályozás" az emberi jogok, a méltóság és az emberi döntéshozók szerepe szempontjából a bonyolult társadalmi-technikai rendszerekben, amelyek többé-kevésbé önmagukat is működtetni ígérik.<sup>22</sup> Frank Pasquale megjegyzi, hogy a társadalmi "hatalom egyre inkább algoritmikusan fejeződik ki".<sup>23</sup> míg John Danaher

---

<sup>13</sup> P Dunleavy and H Margetts, "Design principles for essentially digital governance" (2015) 111th Annual Meeting of the American Political Science Association <<https://eprints.lse.ac.uk/64125/>> (Az Amerikai Politikatudományi Társaság 111. éves ülése). 26.

<sup>14</sup> T O'Reilly, "A kormányzat mint platform" (2010) Innovációk 1.

<sup>15</sup> H Margetts és C Dorobantu, "Rethink government with AI" (*Nature*, 92019. április) <<https://www.nature.com/articles/d41586-019-01099-5>>.

<sup>16</sup> G Hadfield, *Egy lapos világ szabályai: Why Humans Invented Law and How to Reinvent It For a Complex Global Economy* (Oxford University Press 2016).

<sup>17</sup> M Loughlin, "The erosion of sovereignty" (2016) *Netherlands45 Journal of Legal Philosophy* (A szuverenitás erodálódása) 2.

<sup>18</sup> J Zittrain, "The Generative Internet" (2006) *Harvard119 Law Review* (Harvard Jogi Szemle) 1974.

<sup>19</sup> O'Reilly, "A kormány mint platform", 38.

<sup>20</sup> M Goede, "E-Esztország: The e-government cases of Estonia, Singapore, and Curaçao" (2019) *Archives8 of Business Research* 2.

<sup>21</sup> Obama Fehér Ház, "Digitális kormányzat: Building a 21<sup>st</sup> Century Platform to Better Serve the American People" (2012), <https://obamawhitehouse.archives.gov/sites/default/files/omb/egov/digital-government/digital-government.html>; UK Cabinet Office, "Government Transformation Strategy to2017 2020", február h92017, <https://www.gov.uk/government/publications/government-transformation-strategy-2017-to-2020>; Treasury Board of Canada Secretariat, "Digital Operations Strategic Plan: 2018-2022", 2018. december 27., <https://www.canada.ca/en/government/system/digital-government/digital-operations-strategic->

terv-2018-2022.html.

<sup>22</sup> E Medina, "Rethinking Algorithmic Regulation" (2015) *Kybernetes*44 6/7.

<sup>23</sup> F Pasquale, *The Black Box Society* (Harvard University Press 2015) 8.

óva int az "algokrácia fenyegetésétől" - azzal érvelve, hogy "nehéz elviselni az algokrácia fenyegetését, azaz megtalálni a módját annak, hogy az emberek "a hurokban maradjanak" és érdemben részt vegyenek a döntéshozatali folyamatban, miközben megtartják az algokratikus rendszerek előnyeit".<sup>24</sup> Mindkettő kulcsfontosságú megfigyelés az 5.0 társadalom számára, ahol "az emberek, a dolgok és a rendszerek... mind a kibertérben kapcsolódnak egymáshoz, és az emberek képességeit meghaladó mesterséges intelligencia által elért optimális eredmények visszakerülnek a fizikai térbe". Az emberi képességeket "meghaladó" mesterséges intelligencia gondolata azonban az, ahol a Társadalom 5.0 víziója még élesebb fókuszba kerül. Az "emberközpontú társadalom" törekvő retorikáján túllépve ez végső soron egy olyan jövő, ahol a mesterséges általános intelligencia (AGI) már nem hipotetikus. Az AGI hipotézise szerint egy gépet úgy lehet megtervezni, hogy bármilyen "általános intelligens cselekvést" elvégezzen.<sup>25</sup> amire az ember képes - ez az elképzelés Japánban régóta intézményes támogatásra talál.<sup>26</sup> De az AGI megidézése az, ami miatt a "Társadalom 5.0" nehezen egyeztethető össze azzal, hogy pontosan mit is jelent ez az emberi döntéshozók központiségére és szükségletére nézve.<sup>27</sup>

## 2. A jogállamiságtól a jogi szingularitásig

Bár a Társadalom 5.0 talán jól példázza azt, amit Jevgenyij Morozov a "megoldáselvűség" ostobaságának nevez, ez nem kizárólag japán jelenség.<sup>28</sup> Valójában az ilyen techno-megoldáshit már régóta része a Richard Barbrook és Andy Cameron által "kaliforniai ideológiának" nevezett "dotcom neoliberalizmusnak".<sup>29</sup> Ez az ideológia azonban mostanra bekúszott a jogi technológia fejlesztőinek retorikájába, akik az adatintenzív - és így célzottan gazdag - jogi környezetet tartják szem előtt. A befektetések, a hatékonyabb és olcsóbb szolgáltatások ígérete, valamint az emberi ügyvédek és bírák fölötti döntési képességek állítása révén a LegalTech-et most arra használják, hogy a mesterséges intelligenciára és a Big Data-ra épülő "intelligens" jog új korszakát vezesse be.<sup>30</sup> Egyesek, például Max Tegmark fizikus számára a felhasználási lehetőség egyértelmű:

Mivel a jogi folyamat absztrakt módon számításnak tekinthető, amely a bizonyítékokra és a jogszabályokra vonatkozó információkat írja be, majd döntést hoz, egyes tudósok arról álmodnak, hogy robotbírókkal teljesen automatizálják azt: Olyan mesterséges intelligencia-rendszerek, amelyek fáradhatatlanul ugyanazokat a magas jogi normákat alkalmazzák minden ítéletre anélkül, hogy az emberi hibáknak, például az elfogultságnak, a fáradtságnak vagy a legújabb ismeretek hiányának engednének.<sup>31</sup>

---

<sup>24</sup> J Danaher, "Az algokrácia fenyegetése: Reality, Resistance and Accommodation (2016) *Philosophy & Technology* 3, 266.

<sup>25</sup> A Newell és HA Simon, "Computer science as empirical inquiry: symbols and search" (1976) *Communications of the ACM* 3, 116.

<sup>26</sup> EY Shapiro, "The fifth generation project-a trip report" (1983) *Communications of the ACM* (Az ACM közleményei) 9.

<sup>27</sup> DJ Chalmers, "A szingularitás: *Journal of Consciousness Studies* 7; N Bostrom, *Superintelligence: Paths, Dangers, and Strategies* (Oxford University Press 2014).

<sup>28</sup> E Morozov, *Hogy mindent megmentsen, kattintson ide: Technology, Solutionism, and the Urge to Fix Problems That Don't Exist* (Penguin 2014).

<sup>29</sup> R Barbrook és A Cameron, "The Californian Ideology" (1996) *Science & Culture* (A kaliforniai ideológia). 1.

<sup>30</sup> MB Chalmers, "SmartLaw 2.0: (*Lexology*, 222018. augusztus) <<https://www.lexology.com/library/detail.aspx?g=c7605940-ef62-4c49-9451-3799083beb60>>.

<sup>31</sup> M Tegmark, *Élet 3.0: Embernek lenni a mesterséges intelligencia korában* (Allen Lane 2017) 105.



Mások, például Richard Posner bíró, óvatosak, de nem kevésbé szimpatikusak az elképzeléssel:

A bírói mentalitásnak nem lenne nagy jelentősége, ha a bírák nem tennének mást, mint a jogalkotók, közigazgatási szervek, alkotmányalkotók és egyéb bíróságon kívüli források (beleértve a kereskedelmi szokásokat) által létrehozott egyértelmű jogi szabályokat alkalmaznák azokra a tényekre, amelyeket a bírák és esküdtek elfogultság és előítéletek nélkül határoztak meg. A bírák jó úton haladnának afelé, hogy digitalizált mesterséges intelligenciaprogramok által szuperáljanak... Nem tudom, hogy az originalisták és más jogászok miért nem lelkesednek az AI-ért.<sup>32</sup>

Eugene Volokh jogtudós még egy jogi Turing-tesztet is javasol annak megállapítására, hogy egy "mesterséges intelligenciával rendelkező bíró" érvényes jogi döntéseket hoz-e. Volokh szerint a kimenet meggyőző ereje számít:

Ha egy entitás elég megbízhatóan végez orvosi diagnózisokat, akkor elég intelligens ahhoz, hogy jó diagnosztika legyen, legyen szó akár emberről, akár számítógépről. Nevezhetjük "intelligensnek", de lehet, hogy nem. De így vagy úgy, de használnunk kell. Hasonlóképpen, ha egy entitás elég jól ír bírói véleményeket... akkor elég intelligens ahhoz, hogy jó mesterséges intelligenciájú bíró legyen... Ha egy rendszer megbízhatóan olyan véleményeket ad ki, amelyeket megalapozottnak tartunk, akkor el kell fogadnunk, anélkül, hogy ragaszkodnánk valamilyen előre meghatározott struktúrához a vélemények előállításának módját illetően.<sup>33</sup>

E nézetek lényege, hogy az emberi bírák nem egyszerűen helyettesíthetők a mesterséges intelligenciával, hanem hogy a "mesterséges intelligenciával rendelkező bírák" előnyben részesítendőek, mivel feltételezhető, hogy nem öröklik az emberi döntéshozatal elfogultságait és korlátait.<sup>34</sup> Más tudósok, például Giovanni Sartor és Karl Branting azonban továbbra is szkeptikusak:

Egyetlen egyszerű szabálylanc- vagy mintaillesztési algoritmus sem képes pontosan modellezni a bírói döntéshozatalt, mivel a bírói testület feladata, hogy ésszerű és elfogadható döntéseket hozzon pontosan azokban az esetekben, amelyekben a tények, a szabályok vagy azok illeszkedése ellentmondásos.<sup>35</sup>

A legmerészebb elképzelés azonban Ben Alarie jogtudóstól és jogi technológiai vállalkozótól származik:

Annak ellenére, hogy általános bizonytalanság uralkodik a jog és a jogi intézmények előtt álló út sajátosságait illetően, és hogy mit várhatunk el gépeinktől ahhoz, hogy fontos hozzájárulást nyújtsanak a joghoz, az évszázad során biztosak lehetünk abban, hogy a technológiai fejlődés (1) a világban megfigyelhető jelenségek lényegesen nagyobb mértékű számszerűsítéséhez ("több adat"); és (2) pontosabb mintafelismeréshez vezet az új technológiák és módszerek segítségével.

---

<sup>32</sup> RA Posner, *How Judges Think* (Harvard University Press 2008) 5.

<sup>33</sup> E Volokh, "Chief Justice Robots" (2019) *Duke Law Journal* 1135,1138.

<sup>34</sup> EM Harley, "Hindsight Bias in Legal Decision Making" (2007) *Social Cognition: Special Issue: The Hind-sight Bias 1*; KL Mosier és LJ Skitka, "Human Decision Makers and Automated Decision Aids: Made

for Each Other?" in R Parasuraman and M Mouloua (szerk.) *Automation and Human Performance: Theory and Applications* (CRC Press 2009).

<sup>35</sup> G Sartor és LK Branting, "Bevezetés: Judicial Applications of Artificial Intelligence" in G Sartor and LK Branting (eds) *Judicial Applications of Artificial Intelligence* (Springer 1998). 1.

ods ("jobb következtetés"). Ebben a hozzászólásomban amellet érvelek, hogy az ellenzőknek mindaddig igazuk lesz, amíg empirikusan be nem bizonyítják, hogy tévednek. E tendenciák csúcspontja lesz az, amit én "jogi szingularitásnak" fogok nevezni.<sup>36</sup>

Bár az AGI-t nem tekintik szükséges feltételnek, Alarie jogi szingularitását úgy írja le, mint egy olyan pontot, ahol az AI egy olyan jogrendszert vezetett be, amely "minden ember teljes megértését meghaladja".<sup>37</sup> A jogi szingularitás, amely a jog hiányosságára és esetlegességére adott válaszként fogalmazódott meg, implicit módon a jogi érvelésnek mint a vitarendezés, valamint a hatáskörök, jogok és felelősségek elosztásának alapjának megszüntetésére irányuló javaslat. Bár a jognak a társadalomban betöltött szerepéről nem esik szó, még 5.0,kevésbé az emberi jogászokról és bírákról, Alarie jogi szingularitása His interpolációnak tekinthető.

Még ha a mesterséges intelligencia kutatásában használt matematikai vagy szimbolikus logika legalább elméletben képes is lenne lemásolni a jogi érvelés szerkezetét, ez nem feltétlenül venné figyelembe azokat a politikai, gazdasági és szociokulturális tényezőket, amelyek befolyásolják a jogi diskurzust és a jogrendszer fejlődését.<sup>38</sup> Jürgen Habermas szerint ezek a tényezők azért fontosak, mert:

Az egységes tudomány pozitivista tézise, amely az összes tudományt egy természettudományos modellbe olvasztja, kudarcot vall a társadalomtudományok és a történelem közötti szoros kapcsolat miatt, valamint amiatt, hogy ezek a tudományok a jelentés szituáció-specifikus megértésén alapulnak, amelyet csak hermeneutikusan lehet magyarázni... a szimbolikus előstrukturált valósághoz nem lehet csak megfigyeléssel hozzáférni.<sup>39</sup>

De ha a matematikai logika nem képes megragadni a jogi érvelés "szituáció-specifikus megértését" és a társadalmi világ komplexitását - legalábbis nem olyan mértékben, amely megfelelne annak, ahogyan a természetes nyelvi kategóriák a társadalmi referenciákat és a jelentés jellegét megismerik -, akkor a jogi szingularitás által feltételezett "mesterséges intelligencia bírák" hipotetikus létrejötte a jog egy sajátos felfogását eredményezné: olyan felfogást, amelyben a jogi ítéletek lényegében determinisztikus vagy valószínűségi eredmények, amelyeket statikus vagy egyértelmű jogi szabályok alapján, társadalmi vákuumban hoznak létre. Ez tagadná, vagy irrelevánsnak tekintené a jogról alkotott konkurens felfogásokat, különösen azt az elképzelést, hogy a jog társadalmi intézmény, amely társadalmilag konstruált tevékenységeket, kapcsolatokat és normákat tartalmaz, amelyeket nem könnyű számszerű funkciókba átültetni.<sup>40</sup> Ez azt is figyelmen kívül hagyná, hogy a jogi döntéshozatal a hatalom olyan gyakorlását jelenti, amely egyszerre anyagi és - Pierre Bourdieu értelmében - "szimbolikus".<sup>41</sup> E hatalom rekurzív gyakorlása csak akkor legitim, ha az a folyamat, amelyen keresztül minden egyes hatalomgyakorlás történik, megfelel az uralkodó eljárási elvárásoknak, amelyek rendkívül esetlegesek és társadalmilag konstruáltak.

---

<sup>36</sup> Alarie, "A törvény útja", 3.

<sup>37</sup> Alarie, "A törvény útja", 11.

<sup>38</sup> P Stein, *Jogi evolúció*: (Cambridge University Press 2009); S Deakin, "Legal Evolution: *University of Cambridge Faculty of Law Research Paper No. 41/2011* (2011).

<sup>39</sup> J Habermas, *On the Logic of the Social Sciences* (1967) idézi: W Outhwaite, *Habermas: Key Contemporary Thinkers* (2. kiadás, Polity Press 2009). 22.

<sup>40</sup> H Ross, *Law as a Social Institution* (Hart 2001).

<sup>41</sup> P Bourdieu, *Language and Symbolic Power* (Harvard University Press 1991).

Ahogy HLA Hart megjegyezte, kritikus fontosságú, hogy "megőrizzük azt az érzést, hogy valaminek a jogilag érvényesnek való minősítése nem döntő az engedelmesség kérdésében, és hogy bármilyen nagy is a hivatalos rendszer fenséges vagy tekintélyes aurája, a követeléseit végül is alá kell vetni...".

erkölcsi vizsgálatnak vetik alá.<sup>42</sup> Röviden, a jogrendszer mint teljesség, és annak eredményei, beleértve a jogi

ítéletek, ellenőrzésnek és egyensúlynak kell alávetni, vagy algoritmikus kifejezéssel élve "visszacsatolásnak", hogy meghatározzák belső műveleteinek "jogi érvényességét", valamint a kibocsátási és végrehajtási folyamatot a társadalommal szemben. A "mesterséges intelligenciával rendelkező bírák" egyik különös veszélye tehát az, hogy az erkölcsi ellenőrzés nem jelentene értelmes megelőzést a hatalommal való visszaélésekkel szemben.

Ez különösen igaz az úgynevezett "fekete doboz" algoritmusokra, amelyek technikai szinten kifürkészhetetlenek, vagy ahol az üzleti titok védelme megakadályozza a diagnosztikai vizsgálatot.<sup>43</sup> Az absztrakció egy másik szintjén azonban az ADM és a "mesterséges intelligenciával foglalkozó bírák" elleni kifogások, még mielőtt azokat elfogadnák és megismernék, szintén egy módja annak, hogy a jogrendszert erkölcsi szempontból is kritikusán vizsgálják. Bár ez a nézőpont nem tagadja, hogy a jogi döntéshozatal bizonyos értelemben strukturálisan algoritmikus természetű - amennyiben meghatározott lépéseket követ egy adott kimenet eléréséhez -, téved, ahogy Mireille Hildebrandt megjegyzi:

...a jogi ítélet matematikai szimulációját összetéveszteni magával a jogi ítélettel. Míg a gépek nagyon jók lehetnek az ilyen szimulációban, maga az ítéletalkotás a jogbiztonság bármely konkrét értelmezésének vitathatóságán alapul a jogrendszer in- tegritásának fényében - ami messze túlmutat a kvázi matematikai konzisztencián.<sup>44</sup>

Ha figyelmen kívül hagyjuk a hatalom és a legitimitás átadásának és a társadalmi rend generációkon belüli és generációk közötti fenntartásának gyakran finom módjait, akkor a jogrendszer "fekete dobozolásán" keresztül a szabadelvű-demokratikus rend egyik fő intézményének aláásását kockáztatjuk. Ezért a "mesterséges intelligenciával rendelkező bírák" gondolatát a legnagyobb óvatossággal és szkepticizmussal kell kezelni, különösen akkor, ha eredményei továbbra is ellenállnak mind az erkölcsi, mind a technikai ellenőrzésnek.<sup>45</sup>

---

<sup>42</sup> HLA Hart, *The Concept of Law* (3. kiadás, Oxford University Press 2012) 210.

<sup>43</sup> *State of Wisconsin v Loomis* N881.W.2d (749Wis. 2016); vö: Wisconsin Supreme Court Requires Warning Before Using Risk Assessments in Sentencing" (2017) Harvard Law Review 1530<<https://harvardlawreview.org/2017/03/state-v-loomis/>>; H-W Liu, C Fu és Y-J Chen, "Beyond State v Loomis: Artificial Intelligence, Government Algorithmization and Accountability" (2019) International Journal of Law and Information Technology 2; R Yu és RS Alo, "What's Inside the Black Box? AI Challenges for Lawyers and Researchers" (2019) Legal Information Management (2019). 1.

<sup>44</sup> M Hildebrandt, "A jog mint számítás a mesterséges jogi intelligencia korában: (2017) <<https://dx.doi.org/10.2139/ssrn.2983045>>.

<sup>45</sup> M Annany és K Crawford, "Látni tudás nélkül: (2018) New Media & Society 3; J Pearl és D MacKenzie,

*The Book of Why: The New Science of Cause and Effect* (Az ok és az okozat új tudománya) (Penguin 2018).  
358.

7

### 3. A digitális számítás eredete

A matematikai és statisztikai számításokat régóta a legjobb módszereknek tekintik arra, hogy a jelenből következtetni lehessen arra, hogy mi fog történni a jövőben.<sup>46</sup> Míg a mesterséges intelligencia és a különböző statisztikai technikák - legalábbis bizonyos mértékig - egyre inkább lehetővé teszik az előrejelzést, a jog kiszámíthatóbbá tételének gondolata összefonódik a számítógépek általánosabb fejlődésével. De melyek ennek a digitális számítógépek által létrehozott számítási univerzumnak az alapvető tulajdonságai?

A számítási univerzum legalapvetőbb elemei a bitek, az információelmélet és a digitális kommunikáció alapegységei. A "bit" szó - a "bináris számjegy" szűkítése - a következő volt John Tukey amerikai matematikus alkotta meg.<sup>47</sup> Akár egy kilobájtos univerzum, akár zettabájt, csak kétféle bit van, amely különbséget tesz benne: a térbeli és az időbeli különbségek. A digitális számítógép úgy választ az információ e két formája között, hogy keményen kódolt szabályok szerint elemzi szerkezetüket és sorrendjüket. A térben változó, de időben nem változó szerkezetként megtestesülő bitek képezik a számítási memória alapját. Míg a szekvenciában megtestesülő bitek, amelyek időben változnak és térben változatlanok maradnak, az úgynevezett számítógépes kód alapját képezik.

A kommunikáció alapvető egységének azonosítása, amelyet a bináris alternatívák (pl. 0/1, ki/be) közötti egyetlen megkülönböztetés képvisel, volt a központi gondolata Claude Shannon információelméletírő akkoriban titkos matematikai elmélete a kriptográfiának<sup>1945</sup>,<sup>48</sup> amelyet a kommunikáció matematikai elméletében (Mathematical Theory of Communication) bővített ki<sup>1948</sup>.<sup>49</sup> A kibernetikus Gregory Bateson később újrafogalmazta Shannon elméletét a "minden különbség, ami különbséget tesz" nyitott kifejezéssel.<sup>50</sup> A számítások digitális univerzumában azonban az egyetlen különbség, ami számít, az a nulla és az egy közötti különbség.

Az az elképzelés, hogy két szimbólum elegendő minden kommunikáció kódolásához, nem volt új. Inkább Francis Bacon (1561-162) állította fel. A *De augmentis scientiarum* című traktátusában<sup>1623</sup> Bacon így érvelt:

Két betű öt helyen történő áthelyezése elegendő lesz a különbségekhez<sup>32</sup> [és] ezzel a Művészettel megnyílik egy út, amellyel az ember kifejezheti és jelezheti elméjének szándékait, bármilyen távolságban, tárgyakkal... amelyek csak kétszeres különbségre képesek.

---

<sup>46</sup> Spiegelhalter, *A statisztika művészete*, 143-187.

<sup>47</sup> CE Shannon, "A Mathematical Theory of Communication" (1948) 47 The Bell System Technical Journal 3 ("A logaritmusos bázis kiválasztása megfelel az információ mérésére szolgáló egység kiválasztásának. Ha a bázist 2 használnjuk, az így kapott egységeket bináris számjegyeknek, vagy rövidebben biteknek nevezhetjük, a J. W. Tukey által javasolt szóval.").

<sup>48</sup> CE Shannon, "A Mathematical Theory of Cryptography-Case 20878" (1945) MM-45-110-92 <<https://www.iacr.org/museum/shannon/shannon45.pdf>>.

<sup>49</sup> Shannon, "A kommunikáció matematikai elmélete".

<sup>50</sup> GA Bateson, "Re-examination of 'Bateson's Rule' (1971) 60 Journal of Genetics 230; vö. MH Schroeder, "The Difference that Makes a Difference for the Conceptualization of Information" (2017) Proceedings 1221.

A továbbiakban példákat hozott arra, hogyan lehet ezt a bináris kódolást papír-, hang- vagy fénysebességgel továbbítani.<sup>51</sup> Bacon felismerését, miszerint a nulla és az egy elégséges az aritmetikához, Thomas Hobbes (1588-1679) továbbfejlesztette, leginkább a *De Corpore* (1655) és a *Computation* (vagy "Logique") című művében (1656). Hobbes ebben a korábbi művében fejti ki először azt az elképzelését, hogy az emberi gondolkodás a számítás egy formája. "Az érvelés alatt - mondja - a számítást értem":

Számolni pedig azt jelenti, hogy egyszerre több dolog összegét összegyűjteni, vagy a maradékot megismerni, ha az egyik dolgot elvettük a másiktól. Számolni tehát ugyanaz, mint összeadni vagy kivonni.<sup>52</sup>

A következő szakaszban Hobbes néhány példát hoz arra, hogyan működik az összeadás az emberi gondolkodásban, és arra a következtetésre jut, hogy az ötletek összeadása lehetővé teszi az összetettebb ötletek megfogalmazását. Hobbes szerint "egy négyszögletes alakzat, egy egyenlő oldalú alakzat és egy négyszögletes alakzat fogalmaiból áll össze a négyzet fogalma".<sup>53</sup> Bár elismeri, hogy ez az akció az emberi gondolkodásnak csak egy töredékét teszi ki, a továbbiakban a tételeket és a szillogizmusokat az összeadás szempontjából írja le:

A szillogizmus nem más, mint egy összegzés, amely két (egy közös terminuson keresztül, amelyet középső terminusnak nevezünk) egymáshoz kapcsolt tételekből áll össze; és így a szillogizmus három névből álló összeadás, ahogyan egy tétel is kettőből áll.<sup>54</sup>

Ily módon Hobbes az elme számítógépes elméletének (Computational Theory of Mind, CTM) egyik előfutárának tekinthető. A CTM központi gondolata az, hogy az emberi elme egyfajta számítógép. Ahogy Fodor fogalmaz: "a szándékos törvények közvetlen megvalósító mechanizmusai komputációsak... [A komputációk] az intencióban szemlélve, a szintaktikai leírás alatt álló szimbólumok leképezései a szintaktikai leíráson kívüli szimbólumokra."<sup>55</sup> Bár van bizonyos fokú összhang Hobbes munkája és a későbbi elmeelméletek között, az az állítás, hogy Hobbes "prófétai módon indította el a mesterséges intelligenciát", túlzásnak tűnik.<sup>56</sup>

#### 4. A Leibniz-álmom és a jog matematizálása

Sokkal kevésbé túlzás azonban a mesterséges intelligencia fogalmi szerveit a német polihisztor Gottfried Wilhelm Leibniznek (1646-1716) tulajdonítani. Mindazonáltal egyértelmű Hobbes hatása Leibniz gondolkodására:

---

<sup>51</sup> F Bacon, *De Augmentis Scientiarum* (1623), G Wats fordítása: *Of the Advancement and Proficiency of Learning, or the Partitions of Sciences* (1640) 265-266.

<sup>52</sup> T Hobbes, *De Corpore* a *De Corpore* I. részében (AP Martinich (trans) Abaris Books 1981) 1.2.

<sup>53</sup> Hobbes, *De Corpore*, 1.3.

<sup>54</sup> Hobbes, *De Corpore*, 4.6.

<sup>55</sup> JA Fodor, *The Elm and the Expert* (MIT Press 1994) 8.

<sup>56</sup> B Haugeland, *Artificial Intelligence; The Very Idea* (MIT Press 1985) 23.



Thomas Hobbes, mindenütt az elvek mélyreható vizsgálója, helyesen állapította meg, hogy minden, amit elménk tesz, számítás, amely alatt vagy egy összeg hozzáadását, vagy egy különbség kivonását kell érteni... Tehát ahogyan az algebrának és az analitikának két elsődleges jele van, a + és a -, ugyanúgy van két kopulája, a "van" és a "nincs".<sup>57</sup>

Hobbes után, és David Hilbert német matematikus előtt,<sup>58</sup> Leibniz úgy vélte, hogy lehetséges a logika, a nyelv és a matematika következetes rendszerének kidolgozása, amelyhez mechanikus szabályok szerint manipulálható, egyértelmű szimbólumokból álló alfabet-et használ. Egy 1675-ös, a Royal Society titkárának és az Isaac Newtonnal folytatott levelezés közvetítőjének írt levelében Leibniz azt írta, hogy "eljön az idő, és hamarosan eljön, amikor Isten és az elme ismerete nem kevésbé biztos, mint a számoké és a számoké, és amikor a gépek feltalálása nem lesz nehezebb, mint a geometriai problémák megoldása".<sup>59</sup> Leibniz, megelőlegezve azt, amit ma szoftverként emlegetünk, kétirányú kapcsolatot látott a logika és a mechanizmusok között. Egy 1679-ben Christiaan Huygens holland matematikusnak küldött levelében Leibniz azt a megfigyelést fűzte hozzá, hogy "egy gép leírását, legyen az bármilyen bonyolult, olyan karakterekkel is el lehet végezni, amelyek csupán az ábécé betűi, és így az elme számára módszert biztosítunk a gép és minden részének megismerésére".<sup>60</sup>

Leibniz elégedetlen volt az általános tízes számrendszer által lehetővé tett fáradságos aritmetikával, és úgy vélekedett, hogy "[a] kiváló emberekhez méltatlan, hogy rabszolgaként órákat vesztegessenek a számolással járó munkára, amelyet biztonságosan át lehetne adni bárki másnak, ha gépeket használnának. <sup>61</sup> A teher enyhítésére Leibniz azt javasolta, hogy "dolgozzanak ki egy általános szimbolikus nyelvet és egy hozzá tartozó algebrát, hogy az emberi kutatás bármely területén bármely tétel igazsága egyszerű számítással meghatározható legyen".<sup>62</sup> A felfedezés művészete című művében (1685) így állította:

---

<sup>57</sup> GW Leibniz, "A kombináció művészetéről" in GHR Parkinson (szerk.) *Leibniz: Logical Papers* (Clarendon 1966). 3.

<sup>58</sup> D Hilbert, "Die Grundlegung der elementaren Zahlenlehre" 104 *Mathematische Annalen* 485; fordította W Ewald: "The Grounding of Elementary Number Theory" in P Manscou (szerk.) *From Brouwer to Hilbert: The Debate on the Foundations of Mathematics in the 1920s* (Oxford University Press 1998) 266-273.

<sup>59</sup> "GW Leibniz to Henry Oldenburg, December 1675/18," in HW Turnbull (szerk.) *The Correspondence of Isaac Newton, Vol. 1* (Cambridge University Press 1959) 401.

<sup>60</sup> "GW Leibniz-kiegészítés egy Christiaan Huygenshez írt levélhez, 16798, szeptembere" in LE Loemeker (szerk.) *Philosophical Papers and Letters, Vol. 1* (University of Chicago Press 1956) 384-385.

<sup>61</sup> G Leibniz idézi: M Nadin, 'Predictive and Anticipatory Computing' in PA Laplante (szerk.) *Encyclopaedia of Computer Science and Technology: Kötet* (CRC Press, 2017).

<sup>62</sup> FC Kreiling, "Leibniz" (1968) *Scientific American* 5,95.

Az egyetlen módja annak, hogy kijavítsuk érvelésünket, ha olyan kézzelfoghatóvá tesszük, mint a matematikusoké, hogy egy pillantással megtaláljuk tévedésünket, és amikor viták vannak a személyek között, egyszerűen azt mondhatjuk: Számoljuk ki, minden további nélkül, hogy kiderüljön, kinek van igaza.<sup>63</sup>

A logikai kalkulus segítségével Leibniz nagyszabású projektbe kezdett, amelynek célja egy olyan "univerzális szimbolizmusról" szóló elképzelés kidolgozása volt, amelyben az ész minden igazsága egyfajta kalkulusra lenne redukálva. Projektjének középpontjában az a gondolat állt, hogy: "az emberi gondolatok egyfajta ábécéje kidolgozható, és hogy minden felfedezhető és megítélhető ezen ábécé betűinek összehasonlításával és a belőlük alkotott szavak elemzésével".<sup>64</sup> Leibniz kidolgozott egy olyan univerzális kódolási rendszert, amelyben az elsődleges fogalmak prím számokkal ábrázolhatók, így átfogó keretet biztosítva a számok és a gondolatok leképezéséhez. Miután maga is elvégezte az alapmunkát, Leibniz úgy vélte, hogy a számok és az ideák teljes hálózatba rendezése nemcsak megvalósítható, hanem "néhány kiválasztott ember öt év alatt be is fejezhetné a dolgot... Mindössze két évükbe telne azonban, hogy tévedhetetlen számítással kidolgozzák az élet számára leghasznosabb tanokat, vagyis az erkölcs és a metafizika tanításait". Ez így volt: "az emberi fajnak lesz egy újfajta eszköze, amely sokkal jobban növeli az elme erejét, mint az optikai lencsék erősítik a szemet, és amely éppúgy fölötte áll majd a mikroszkópnak vagy a távcsőnek, mint ahogy az ész fölötte áll a látásnak".<sup>65</sup>

Bár Leibniz úgy vélte, hogy a bináris kódolás az egyetemes nyelv alapja,<sup>66</sup> a kínaiaknak tulajdonította a nyelv megalkotását. Konkrétan a Yi-Jing - egy ősi kínai filozófiai könyv, amelyet a Kr. e. század második felében állítottak össze - hexagrammaiban látta Leibniz "a bináris aritmetika elemeit... amelyet néhány ezer évvel később fedeztem fel újra".<sup>67</sup> Bár Leibniz ontológiája különbözött az egyes írások között,<sup>67</sup> egyesek szerint az Álma a Yi-Jing ontológiai valóságának félreértésén alapult.<sup>68</sup> Röviden, a Yijing a valóságot nem teljesen valóságosnak, hanem inkább álmhoz vagy illúzióhoz hasonlónak kezeli.<sup>69</sup> Az emberek által megtapasztalt valóság "álma" állítólag a Yin és a Yang bináris ellentétéből ered, ahogyan azok végtelen kombinációikban játszódnak le. Smith megjegyzi, hogy "[a]z Yijing bináris szerkezete elbűvölte és inspirálta Leibnizt, [bár] a Yijing számszimbolikája numerológiai maradt, és... soha nem volt igazán matematikai." <sup>70</sup> Mint kon- szekvencia, Leibniz rendszere a valóság dualista szemléletét integrálta, ahol minden reprezentálható volt.

---

<sup>63</sup> GW Leibniz idézi: M Gelford és Y Kahl, *Knowledge Representation, Reasoning, and the Design of Intelligent Agents* (Cambridge University Press 2014) 7.

<sup>64</sup> GW Leibniz, "Az általános jellemzőről" in *Leibniz: Philosophical Papers and Letters*, 224.

<sup>65</sup> GW Leibniz, "kb. 1679" in *Leibniz: Philosophical Papers and Letters*, 344.

<sup>66</sup> G Leibniz, "A bináris aritmetika magyarázata" < <https://www.leibniz-translations.com/binary.htm>>.

<sup>67</sup> T Tho, 'Mi Leibniz ontológiája? Rethinking the Role of Hylomorphism in Leibniz's Metaphysical Development' (2015) *Journal of Early Modern Studies* 4; J Whipple, "Leibniz on Fundamental Ontology: Idealism and Pedagogical Exoteric Writing" (2017) 4 *Ergo* 11.

<sup>68</sup> ES Nelson, "A Peking és a filozófia: From Leibniz to Derrida" (2011) *Journal of Chinese Philosophy* 38.

<sup>69</sup> C-Y Cheng, "The Yi-Jing and Yin-Yang Way of Thinking" in B Mou (szerk.) *The Routledge History of Chinese Philosophy* (Routledge, 2008).

<sup>70</sup> RJ Smith, *Jósok és filozófusok*: (Westview Press 1991) 205.

küldött és megértett 1-es és 0-s számokkal, vagy: Yin és Yang. A bináris kódot később George Boole angol matematikus finomította, akitől a Boole-logika ered,<sup>71</sup> a modern számítás Leibniz dualista ontológiáját örökölte.<sup>72</sup>

Leibniz fennmaradt feljegyzéseiből kiderül, hogy iteratív módon fejlesztett ki egyszerű algoritmusokat a decimális és a bináris jelölés közötti választásra, valamint az alapvető aritmetikai függvények elvégzésére az 1-es és 0-s karakterláncok segítségével. "A bináris aritmetikában", jegyezte meg, "csak két jel van, és 0a1, amivel minden számot felírhatunk."<sup>73</sup> Ezzel a rendszerrel és egy mechanikus számológéppel ő Leibniz meggyőződött arról, hogy az emberi gondolkodás logikai-matematikai számításokkal való formalizálása nemcsak lehetséges, hanem a matematikai szigor és pontosság bevezetése az összes humán tudományba. Ez az általában Leibniz-álomként (*characteristica universalis*) emlegetett elképzelés fontos előfutára volt a számítástechnika fejlődésének, és előrevetítette a kognitív képességek javítására és kiterjesztésére irányuló későbbi kutatásokat.<sup>74</sup>

De a félreértés Leibniz gondolkodását is befolyásolta egy számára különösen érdekes dologról: a jogról.<sup>75</sup> Leibniz számára a jog azt példázta, hogyan kell a társadalomnak megoldania "az életről és egészségről, az államról, a háborúról és békéről, a lelkiismeret mértékletességéről, [és] az örökkévalóságról való gondoskodásról szóló legsúlyosabb tanácskozásokat".<sup>76</sup> A jogot az emberi racionalitás legfejlettebb eszközeként dicsérte, különösen az okok "mérlegét" vagy "logometrikus mérlegét", amelyet a bírák arra használtak, hogy értékeljék "a vitázók érveinek, a szerzők véleményének [és] a tanácsadók hangjának] relatív súlyát".<sup>77</sup> Ahhoz, hogy a jog olyan pontos legyen, mint a matematika, az igazságos jogi szabály olyan volt, amely egyesítette az absztrakt jogi axiómákat - például azt az elvet, hogy nem szabad kárt okozni másoknak - a természettudományok empirikus felismeréseivel. *Doctrina Conditionum* című művében (1669) ezt a hipotézist a római joggal tesztelte.<sup>78</sup> Bár ambíciói kétségtelenül nagyszabásúak voltak, Leibniz nem egy monolitikus "tudományos modellt" akart ráerőltetni a

---

<sup>71</sup> G Boole, *The Mathematical Analysis of Logic* (MacMillan, Barclay & MacMillan 1847).

<sup>72</sup> J Teixeira, "Computational Complexity and Philosophical Dualism" (1998), <https://www.bu.edu/wcp/Papers/Cogn/CognTeix.htm>; D King, "Cartesian Dualism, and the Universe as Turing Machine" (2003) *Phi-47 losophy Today*. 2.

<sup>73</sup> GW Leibniz, "Discourse on the Natural Theology of the Chinese", fordítás a "Lettre sur la philosophie chinoise à Nicolas de Remond"-ból, in: H Rosemont és DJ Crook (ford. és szerk.) *Monographs of the Society for Asian and Comparative Philosophy*, No. (4) University of Hawaii Press 1977). 158.

<sup>74</sup> S Toulmin, "The Dream of an Exact Language", in B Göranson és M Florin (szerk.) *Dialogue and Technology: Art and Language* (Springer, 1991); vö. N Bostrom és A Sandberg, "Cognitive Enhancement: Methods, Ethics, Regulatory Challenges" (2009) *5 Science and Engineering* 3; M Dresler, A Sandberg, C Bublitz és mások, "Hacking the Brain: Dimensions of Cognitive Enhancement" (2019) *ACS10 Chemical Neuroscience*. 3.

<sup>75</sup> vö. M Armgardt, "Leibniz mint jogtudós" (2014) *20 Fundamina* 1.

<sup>76</sup> GW Leibniz idézi M Dascal "Introductory Essay" in M Dascal (szerk. és ford.) *GW Leibniz: The Art of Controversies* (Springer 2006). 38.

<sup>77</sup> GW Leibniz idézve: *GW Leibniz: A viták művészete*, 36.

<sup>78</sup> P Boucher, "Leibniz: Milyen jogi racionalizmus?" in M Dascal (szerk.) *Leibniz: Milyen racionalista?*



törvény. Inkább úgy vélte, hogy az axiomatikus módszer segítene pontosabbá, és ideális esetben előre megszabhatóvá tenni azt.

Ezek az ötletek azonban 1679-ben Leibniz képzeletében egy olyan digitális számítógéppé álltak össze, amelyet ma digitális számítógépnek neveznénk, ahol a bináris számokat golyókkal lehetne ábrázolni és mechanikus kapukkal irányítani. "Ez a [bináris] számítás" - írta Leibniz:

...egy gép (kerekek nélkül) a következő módon valósíthatná meg, méghozzá könnyen és erőfeszítés nélkül. A tartályt úgy kell ellátni lyukakkal, hogy azok nyithatók és zárhatók legyenek. Azokon a helyeken legyenek nyitva, amelyek a-1-nak felelnek meg, és 1 maradjon zárva azokon, amelyek a 0-nak felelnek meg. A nyitott kapukon keresztül kis kockák vagy golyók hulljanak a pályákra, a többin keresztül semmi. Ezt [a kaputáblát] szükség szerint oszlopról oszlopra kell eltolni.<sup>79</sup>

Bár nem nevezte így, Leibniz néhány évvel 270 azelőtt találta fel a váltóregisztert, hogy azt a Colossus Mark 2-ben - egy 1944-ben kifejlesztett brit kódtörő számítógépben - megvalósították volna, és a szövetségesek a normandiai partraszállás során használták volna.<sup>80</sup> A modern számítógépek középpontjában álló váltóregiszterekben a feszültséggradiensek és az elektronimpulzusok felváltották az eredeti tervekben szereplő golyókat és a gravitációt, de minden gyakorlati szempontból a modern számítógépek még mindig nagyjából ugyanúgy működnek, mint ahogy Leibniz elképzelte őket a következő években 1679.

## 5. *Calculus!* Leibniz hatása a jogra

Leibniz művének posztumusz angol fordításai hívták fel a figyelmet az angolul beszélő világra. Különösen a jogtudósokat vonzotta, például a formalista Christopher Columbus Langdell, aki a *Cases on Contracts* (1871) című művében a jog axiomatikus felfogását mutatta be.<sup>81</sup> Leibniz számos állítását visszhangozva Langdell azzal érvelt, hogy "a jog, tudománynak tekintve, bizonyos elvekből vagy tanokból áll. Ezek olyan szintű elsajátítása, hogy azokat állandó könnyedséggel és bizonyossággal tudja alkalmazni az emberi ügyek egyre kuszább szálaira, ez teszi az igazi jogászt...".<sup>82</sup> Eltért azonban a jogi axiómák forrását illetően. Míg Leibniz a létező hazai és a természetjog kombinációját tekintette forrásuknak, Langdell úgy vélte, hogy a jogi axiómák - vagy hogy az ő kifejezését használjam: "elvek" - empirikusan származtathatók, mint a természettudományokban. Langdell szerint az empirikus "jogtudomány" adatait a bíróságok által eldöntött és közzétett esetek tartalmazzák.<sup>83</sup> Úgy vélte, hogy ezek az esetek "adatai" lehetővé tennék a szerződéses jogviták kimenetelének előrejelzését - ez volt az elsődleges célja.<sup>84</sup>

---

<sup>79</sup> GW Leibniz, "De Progressione Dyadica-Pars I (1679. március 15.)" in E Hochstetter és H-J Greve (szerk.) *Her- rn von Leibniz' Rechnung mit Null und Einz* (Siemens Aktiengesellschaft 1966).

<sup>80</sup> J Copeland és mások (szerk.) *Colossus: The Secrets of Bletchley Park's Codebreaking Computers* (Oxford University Press 2006) 74-77, 100; vö. T Kidder, *The Soul of a New Machine* (Back Bay Books 2000).

<sup>81</sup> CP Wells, "Langdell and the Invention of Legal Doctrine" (2010) *Buffalo 28 Law Review* (2010). 3.

<sup>82</sup> CC Langdell, *Cases on Contracts* (Little Brown & Co. 1871), vi.

<sup>83</sup> MH Hoeflich, "Langdell and the Invention of Legal Doctrine", (1986) *The 30 American Journal of Legal History* 2.

<sup>84</sup> CP Wells, "Langdell and the Invention of Legal Doctrine", 599-605.

Oliver Wendell Holmes amerikai jogász, annak ellenére, hogy hevesen nem értett egyet Leibnizzel és Langdell-lel, lelkes csodálója volt a német polihisztornak. A Harvard jogi karának előadójaként, ahol Langdell egyidejűleg dékán volt, Holmes korai karrierjét a Leibnizi-langdelli axiomatika elutasításával töltötte, és a jogtudományban olyan fordulatot kezdeményezett, amely óvakodott minden formalizmustól vagy "logikától" a jogban. A *The Common Law* (1881) című művében Holmes híres módon kijelentette, hogy "A jog élete nem a logika volt: a tapasztalat volt".<sup>85</sup> Egy későbbi, *A jog útja* (1897) című cikkében, amelyet ravasz módon közvetlenül azután tett közzé, hogy Langdell lemondott a dékáni tisztségéről, Holmes elítélte a jogi axiomatikát, mint "olyan tévedést, amelyet fontosnak tartok leleplezni".<sup>86</sup>

Holmes ebben a későbbi cikkében fogalmazta meg "a jog előrejelzési elméletét".<sup>87</sup> amelynek középpontjában a retorikai "rossz ember" állt: egy olyan személy, aki mentes az etikától vagy a bíróságok jogászai szerepéről alkotott magasztos elképzelésektől, és akit csak a kártérítés elkerülése és a börtöntől való megmenekülés érdekel. Holmes maradandó hozzájárulása a jogtudományhoz az volt, hogy elutasította a jogi formalisták, például Langdell által képviselt determinisztikus és ~~okarets~~ jogfelfogást, akik a jogot szabályok és normák következetes halmazának tekintették, amelyből mindig "helyes" választ lehet levezetni. Scott Brewer szerint "...az a teoretikus, aki nagy sűrűségű axiomatizációra törekszik, mint Leibniz és Langdell, úgy tűnik, hogy meg kell próbálnia a Scylla között lavírozni, hogy olyan nyitott és gyakran olyan homályos axiómákat generáljon, hogy soha nem tudna megállapodni az axiómák végső halmazában, hanem örökké revidálnia kellene azokat."<sup>88</sup>

Holmes formalizmusának elutasítása adta az amerikai jogi reáliskola jogi axiomatikával szembeni kritikájának ideológiai alapját.<sup>89</sup> Roscoe Pound például kigúnyolta a "mechanikus jogtudományt".<sup>90</sup> mint a bírák lusta gyakorlatát, hogy a precedenseket formulaszerűen alkalmazzák az esetekre, a következményekkel való számottevő figyelmen kívül hagyásával. Pound szerint a precedensek logikája nem képes megoldani a jogászai problémákat; figyelmeztetett, hogy az axiomatika azzal a kockázattal jár, hogy a társadalmilag konstruált és politikailag befolyásolt jogi fogalmak magától értetődő igazságokká merevednek. Pontosan az a törekvés, hogy elkerüljék ezt a megcsontosodást - és ne hagyják, hogy a "rossz ember" győzzön - vezette Holmes-t és a realistákat arra, hogy a jogot eredendően határozatlannak és a szabályok és elvek koherens vagy teljes rendszereként való értelmezéssel szembeszegülőnek tekintsék.

ples, amelyek közül mindig volt egy "helyes" válasz.<sup>91</sup> *A Logical Method and Law* (1924) című könyvében John

Dewey kiterjesztette Holmes érvét, miszerint "az általános tételek nem döntenek konkrét ügyekben" - egy

---

<sup>85</sup> OW Holmes, *The Common Law* (Little Brown & Co. 1881) 1.

<sup>86</sup> OW Holmes, "The Path of the Law" (*A jog útja*) (1897) *Harvard Law Review* 5,997.

<sup>87</sup> MH Fisch, "Justice Holmes, the Prediction Theory of Law, and Pragmatism" (1942) *The Journal of Philosophy* 4; RA Posner, "The Path Away from the Law" (1996) *Harvard Law Review*. 1039.

<sup>88</sup> S Brewer, "Law, Logic, and Leibniz: in A Artosi, B Pieri and G Sartor (szerk.) *Leibniz: Logikai-filozófiai rejtvények a jogban* (Springer 2013).

<sup>89</sup> SR Ratner, "Legal Realism School", in R Wolfrum (szerk.) *Max Planck Encyclopedia of Public International Law Online* (Oxford University Press, 2007).

<sup>90</sup> R Pound, "Mechanical Jurisprudence", *Columbia Law Review* no8. 605 (1908).

<sup>91</sup> C Gray, *The Nature and Sources of Law* (Columbia University Press 1909); R Pound, "Justice According to Law" (1914) 1 *The Midwest Quarterly* 3; KN Llewellyn, *The Bramble Bush: The Classic Lectures on the Law and Law School* (Oxford University Press, [1930] 2008).

a "szabályskeptizmus" néven ismert álláspont. Dewey számára egyetlen jogi érv sem ábrázolható érvényesen deduktív következtetésként, és a jog soha nem ábrázolható pontosan deduktív axiomatikus rendszerekkel, érveléssel:

Természetesen minden ok megvan arra, hogy a jogi szabályok a lehető leghatározottabbak és legáltalánosabbak legyenek. De a ténylegesen elérhető előzetes bizonyosság mennyisége és fajtája nem formai, hanem ténybeli kérdés. Nagy a biztonság ott, ahol a társadalmi viszonyok meglehetősen egységesek, és ahol az ipar, a kereskedelem, a szállítás stb. a régi szokások csatornáin mozog. Sokkal kisebb ott, ahol a találmányok aktívak, és ahol az új üzleti és kommunikációs eszközök az emberi kapcsolatok új formáit hozzák létre.<sup>92</sup>

A realisták általában szkeptikusan és lekezelően tekintettek a bírói véleményekre és a bírákra. Mivel az írott szabályok (törvények, esetek) nem határozták meg, hogy mi a jog, a jogi érvelés célja nem az volt, hogy elmagyarázza, hogyan jutott a bíróság egy-egy döntésre, és útmutatást nyújtson a jövőben hasonló helyzetekkel szembesülő bírák és peres felek számára. A valódi cél a döntések "arányosítása" és "legitimálása" volt, és a nyilvánosság és a többi bíról elől a valódi, gyakran visszatetsző indoklások elrejtése. Mint ilyen, semmilyen deduktív axiomatikus módszer nem tudott számot adni ezekről a hatásokról, és a társadalomtudományi felismerések fokozatos beépülése a jogtudományba segített a realisták szkeptizmusának elfojtásában.

## 6. *Characteristica Universalis Lex*

De ez korántsem volt az utolsó szó a jog mint axióma eszméjéről. A 20. században a Leibniz-álom inspirálta Alfred Tarskit<sup>93</sup> és másokat, például a mesterséges intelligencia úttörőjét, John McCarthyt, hogy megvizsgálják, vajon az axiomatikus módszer alkalmazható-e a matematikán túl is. Az 1950-es években McCarthy a mesterséges intelligencia (AI) akkoriban kialakulóban lévő területén belül egy olyan módszertant dolgozott ki, amelyet logikai alapú megközelítésnek neveztek el.<sup>94</sup> McCarthy elképzelése az volt, hogy egy számítógépes program nagy mennyiségű "emberi tudást" rögzít és tárol egy adott területről (például a jogról vagy az orvostudományról), és matematikai logikát használ annak ábrázolására, valamint logikai következtetéseket a tudásbázisból a kívánt "kimenet" eléréséhez szükséges "legjobb" cselekvések meghatározására.<sup>95</sup> Amióta McCarthy formalizálta a logikai alapú megközelítést, a mesterséges intelligencia kutatása továbbra is az AGI vagy "erős mesterséges intelligencia" hosszú távú célját követi.<sup>96</sup> Ezzel szemben a "gyenge mesterséges intelligencia" olyan gépeket foglal magában, amelyek speciális problémamegoldó vagy "érvelő" feladatokat látnak el "szűk területeken", amelyek nem igénylik az emberi intelligencia szélesebb aspektusait.

---

<sup>92</sup> J. Dewey, "Logical Method and Law" (1924) Cornell Law Review 1,25.

<sup>93</sup> AB Feferman és S Feferman, *Alfred Tarski: Life and Logic* (Cambridge University Press 2004).

<sup>94</sup> J McCarthy, "A szimbolikus kifejezések rekurzív függvényei és gépi számításuk", rész: "A szimbolikus kifejezések rekurzív függvényei és gépi számításuk".  
I" (1960), <http://www-formal.stanford.edu/jmc/recursive.pdf>; vö. SL Andersen, "John McCarthy: (2002) 17  
IEEE Intelligent Systems 5.

<sup>95</sup> J McCarthy, "A logikai mesterséges intelligencia filozófiai és tudományos előfeltételei", in HJ Levesque és F Pirri (szerk.)  
*Logical Foundations for Cognitive Agents* (Springer 1999).

<sup>96</sup> P Wang és B Goertzel (szerk.) *Theoretical Foundations of Artificial General Intelligence* (Atlantis Press,



2012); L Muehlhauser, "What is AGI" (2013) Machine Intelligence Research Institute <<https://intelligence.org/2013/08/11/what-is-agi/>>.

A logikai AI kezdetben ígéretes alkalmazása az úgynevezett jogi szakértői rendszerek (LES) volt, amely az 1970-es és 1980-as években élte fénykorát. Mivel a hasonló szakértői rendszerek olyan "bonyolultabb" területeken működtek, mint az orvostudomány, széles körben feltételezték, hogy a viszonylag "egyszerű" jogi területen is működni fognak.<sup>97</sup> A LES-t egy egyszerű gondolat alapozta meg: "...hogya a jogi szabályokat számítógépes formális rendszerbe lehet önteni, és a tanácsadás a másik végén kijön".<sup>98</sup> L Thorne McCarty informatikus például magabiztosan hirdette:

A [jog] ideális jelöltnek tűnik a mesterséges intelligencia megközelítéséhez: a "tényeket" egy alacsonyabb szintű szemantikai hálózatban reprezentálnánk, a "jogot" egy magasabb szintű szemantikai leírásban, a jogi elemzés folyamatát pedig egy mintaillesztési rutinban.<sup>99</sup>

Az akkori lelkes fejlesztők szerint egy hasznos LES létrehozásához mindössze arra volt szükség, hogy: 1) lefordítani a jogszabályokat vagy "szabályokat" gépileg olvasható formátumba; 2) írni egy olyan szoftvert, amely képes értelmezni azokat; és 3) összegyűjteni néhány jogi szakértőt, hogy magyarázzák el a jogi szabályokat, hogy aztán egy számítógépes programozó be tudja őket vetni. A LES fejlesztői tehát úgy értelmezték feladatukat, mint a Hart által "nyílt textúrának" nevezett formalizálást.<sup>100</sup> a jogi szabályok matematikai feltételrendszerekké amelyről, ahogyan Leibniz évszázadokkal korábban törekedett rá, nem lehetett nézeteltérés.<sup>101</sup> Azzal, hogy...

Ennek megkísérlésével a Leibniz-álom újjáélesztésére törekedtek, a számítások felhasználásával, hogy elérjék a "mechanikus jogtudományt", amelyet Pound és a jogi realisták a jog és a jogi érvelés bonyolult és meghatározhatatlan természetének radikális leegyszerűsítéseként kifogásoltak.

Néhány közepes siker ellenére a LES túlságosan ambiciózusnak és rövidlátónak bizonyult abban a kísérletében, hogy Occam borotváját a jogtudományra és a jog szociológiai kontextusára alkalmazza, hogy egy megtisztított matematikai "lényegét" nyerjen ki belőle. Bár a kudarcnak számos és összetett oka volt, a LES-projekt összeomlása elsősorban a technikai korlátoknak és az axiomatikus módszerek jogra való alkalmazásának bonyolultságának durva alábecsülésének volt köszönhető.<sup>102</sup> A kudarc magva azonban az volt, hogy eleve tagadták, hogy a jog társadalmi kontextusa, célja és antropológiai szerepe releváns lenne a jog ~~érvelés~~ szempontjából. Bár az LES a jog "szűk" területén működött, a jogi érvelés, a következtetés és a következtetés természete egyértelművé tette, hogy az emberi intelligencia szélesebb aspektusait igényli. Mivel az emberek úgy tanulják meg a feladatok elvégzését, hogy a megszerzett tudás egyszerre válik hallgatólagos tudássá

---

<sup>97</sup> R Susskind, *Transforming the Law* (Oxford University Press 2000), 162-206.

<sup>98</sup> P Leith, "The Rise and Fall of the Legal Expert System" (A jogi szakértői rendszer felemelkedése és bukása) (2010) *European Journal of Law and Technology* 1.

<sup>99</sup> LT McCarty, "Some Requirements for a Computer-based Legal Consultant", idézi: RN Moles, *Definition and Rule in Legal Theory: A Reassessment of HLA Hart and the Positivist Tradition* (Basil Blackwell Oxford 1987), 269.

<sup>100</sup> Hart, *A jog fogalma*, 124-154.

<sup>101</sup> WG Popp és B Schlink, "Judith, A Computer Program to Advise Lawyers in Reasoning a Case" (1974) *Jurimetrics Journal* 303.



és implicit,<sup>103</sup> A LES mérnökök nem tudták axiomatikát használni a jogi ismeretek és intuíciók kivonatolására és reprezentálására.<sup>104</sup>

A 20. század végére a logikai-intelligens megközelítéstől a konnekcionista modellek használata felé történő elmozdulás ismét megváltoztatta a dolgokat.<sup>105</sup> A valós feladatokban elért sikereik - például Gary Kasparov sakknagymester legyőzése az IBM Deep Blue gépével - újjáélesztették az AGI-kutatás iránti érdeklődést.<sup>106</sup> A század elején<sup>21</sup> az IBM Watsonjának szenzációs győzelme a Jeopardyban és a DeepMind AlphaGójának győzelme a Go társasjátékban úgy tűnt, hogy a "megoldó intelligencia" célja technikailag megvalósíthatóvá vált.

és kereskedelmi szempontból életképes.<sup>107</sup> A logikai-AI megközelítéssel ellentétben ezek az újabb rendszerek, amelyeket a

az agy neurotudományos modelljei, nem próbálták meg kimerítően formalizálni az emberi szakértelmet és tudást, hogy axiómákat hozzanak létre. Ehelyett az ML és DL technikák alkalmazásával hatalmas mennyiségű történeti példát (képzési adatokat) használtak, amelyek lehetővé tették számukra, hogy iteratív módon frissítsék matematikai paramétereiket, és optimalizálják a teljesítményt az irányított feladatokban. Bár eredményeik kétségkívül lenyűgözőek voltak, David Spiegelhalter statisztikus emlékeztet:

...hogyan ezek olyan technológiai rendszerek, amelyek múltbeli adatokat használnak fel azonnali kérdések megválaszolására, nem pedig olyan tudományos rendszerek, amelyek a világ működésének megértésére törekednek: kizárólag az alapján kell megítélni őket, hogy mennyire jól végzik el az adott korlátozott feladatot, és bár a megtanult algoritmusok formája nyújthat némi betekintést, nem várható el tőlük, hogy a mindennapi életben képzelőerővel vagy emberfeletti képességekkel rendelkezzenek.<sup>108</sup>

Még ezzel a fenntartással is biztonsággal megjósolhatjuk, hogy az ML jogi alkalmazása iránti érdeklődés továbbra is növekedni fog. Ez nem csak azért van így, mert ez egy új, bár még nagyrészt kipróbálatlan technológia. Ahogy rövid történeti áttekintésünkéből kiderült, a funkcionálisan teljes jogrendszer gondolata, amely a matematikai modellekhez kapcsolódó formális logika és axiomatikus érvelés segítségével leírható, nem új. A sorozatos cáfolatok és megfordítások ellenére mélyen megragadta a jogi képzeletet. Így a mai jog-AI vita, a jog és a technológia közötti korábbi vitákhoz hasonlóan, új formában teszi fel a jog mint kormányzási mód természetére vonatkozó alapvető kérdést: kiszámítható-e a jog? A gyűjteményben szereplő hozzászólások különböző módon erre a kérdésre keresik a választ.

---

<sup>103</sup> M Polanyi, *The Tacit Dimension* (Rev ed, University of Chicago Press 2009) 3-25.

<sup>104</sup> P Leith, "Fundamental Errors in Legal Logic Programming" (1986) 29 *The Computer Journal* 6.

<sup>105</sup> SI Gallant, "Connectionist expert systems" (1988) *Communications* 32 of the ACM (Az ACM közleményei) 2.

<sup>106</sup> Y Seirawan, HA Simon és T Munakata, "The Implications of Kasparov Vs. Deep Blue" (1997) *Communications of the ACM* 8; M Newborn, "Deep Blue's contribution to AI" (2000) *Annals of Mathematics and Artificial Intelligence* 1-4.

<sup>107</sup> D Hassabis D Kumaran, C Summerfield és M Botvinick, "Neuroscience-inspired Artificial Intelligence," (2017) 95 *Cell* 2; JE Laird, C Lebiere és PS Rosenbloom, "A Standard Model of the Mind: Toward a Common Computational Framework Across Artificial Intelligence, Cognitive Science, Neuroscience, and Robotics" (2017) 38 *AI Magazine* 4.



## 7. A komputacionalizmus és a valóság matematizálása

Egyre inkább elfogadják, hogy a biológiai és a mesterséges intelligencia csak a szubsztrátokban különbözik,<sup>109</sup> egyesek szerint a mesterséges intelligencia és az adattudomány egy napon lehetővé teszi a társadalmi valóság rendszerszintű modellezését és az emberi viselkedés teljes körű előrejelzését.<sup>110</sup> Bár ez a hosszú távú cél, az a közvetlenebb cél, hogy a mesterséges intelligenciát az ügyvédek és bírák kognitív tartományának leképezésére használják, bizakodásra ad okot, hogy a mesterséges intelligenciát alkalmazó jogi technológiák és ADM-rendszerek új generációja sikerrel jár majd ott, ahol logikai alapú elődeik nem tudtak: a jogi érvelés, következtetés és következtetés leképezésében. Velük együtt jár egy ismerős refrén: A LegalTech javítani fogja az igazságszolgáltatáshoz való hozzáférést, csökkenteni fogja a költségeket, és számos más gyakorlati előny mellett javítani fogja a jogi adminisztráció hatékonyságát.<sup>111</sup> Bár az új generáció funkcionális képességei kétségtelenül javultak, a LegalTech vállalkozás még mindig azon a Leibniz-Langdelli-féle feltételezésen nyugszik, hogy a jognak és a jogi érvelésnek van egy tisztított lényege, amelyet matematizálni kell. Ez Mag Tegmark fizikust arra készíti, hogy elgondolkodjon:

...miért mutatott fizikai világunk olyan rendkívüli matematikai szabályszerűséget, hogy a csillagászat superhőse, Galileo Galilei azt hirdette, hogy a természet "egy matematika nyelvén írt könyv", a Nobel-díjas Eugene Wigner pedig a "matematika ésszerűtlen hatékonyságát a fizikai tudományokban" mint magyarázatot követelő rejtélyt emelte ki?<sup>112</sup>

Végső soron a "matematika ésszerűtlen hatékonyságába" vetett hit mozgatja a jelenlegi mesterséges intelligencia-vállalkozás nagy részét, beleértve a jogi területen való alkalmazását is. A szerződéselemzés például a jogi technológiai alkalmazások népszerű fókuszában áll, és számos startup cég verseng azért, hogy a szerződési feltételek saját értelmezései végleges referenciaként érvényesüljenek. Miközben az üzleti felhasználási eset egyértelmű, a sok dolog közül azt hagyja figyelmen kívül, hogy az olyan széles körben megfogalmazott általános kikötések, mint az "ésszerűség" vagy a "jóhiszeműség" szimbiózisban működnek alacsonyabb szintű fogalmakkal és végső soron az egyedi jogviták tényspecifikus eseteivel. Egy jogi szabály alkalmazása a társadalmi tények egy halmazára ebben az értelemben algoritmikus, amely magában foglalja a fogalmak hozzáadását összetettebb fogalmak kialakításához, de az általánosság különböző szintjein kifejezett fogalmak és szabályok közötti kölcsönhatások függvénye. A determinisztikus vagy "mechanikus jogtudományi" megközelítés, amely figyelmen kívül hagyja az általánosság különböző szintjein lévő fogalmak közötti kölcsönhatást, nem ér el többet, mint hogy a jogi koncepciókat magától értetődő számítógépes "igazságokká" merevíti. Következésképpen egy radikálisan leegyszerűsített jogi ontológiát terjeszt elő, amely feltételezi, hogy ezek a fogalmak stabil referenciák, amelyek egy Turing-gép által megoldható számítási problémák részét képezik.<sup>113</sup> Így a jogi elméletet alátámasztó implicit feltevés, hogy a jogi

---

<sup>109</sup> PS Churchland, "Taníthat-e nekünk valamit a neurobiológia az öntudatról? (1994) 42 *Proceedings and Addresses of the American Philosophical Association* 3; M Allen és KJ Friston, "From cognitivism to autopoiesis: towards a computational framework for the embodied mind" (2018) *Synthese* (19562018); vö. M Velmans, "Is human information processing conscious?". *Behavioral and Brain Sciences* no14. 4 (1991).

<sup>110</sup> M Morrison, *A valóság rekonstruálása: Models, Mathematics, and Simulations* (Oxford University Press, 2015); Tegmark, *Our Mathematical Universe*.

<sup>111</sup> Susskind, *Transforming the Law*, 237-243; R Susskind, *Tomorrow's Lawyers* (2<sup>nd</sup> ed, Oxford University Press 2017) 93-121.

<sup>112</sup> Tegmark, *Matematikai univerzumunk*, 6.

<sup>113</sup> MacCormick, *What Can Be Computed*, 317-330.

A technológiai vállalkozás szerint a "jog" végső soron Turing Complete, és a "jogi problémák" végső soron "helyes" vagy "helytelen" - vagy "1" vagy "0" -ként határozhatók meg.<sup>114</sup>

A LegalTech fejlesztői látszólag közömbösek azzal kapcsolatban, hogy a LES miért bukott eleve, és továbbra is ragaszkodnak a jog axiomatikus felfogásához, de a valóságnak egy implicit felfogásához is, amelyet pánszámításnak neveznek - ez a nézet szerint a fizikai valóság információval leírható, az egység pedig egy számítógépes program vagy számítógépes folyamatok hálózatának determinisztikus vagy valószínűségi kimenete.<sup>115</sup> Legismertebb formációja Tegmark matematikai univerzum-hipotézise (MUH) szerint: "fizikai világunkat nemcsak hogy matematikával írjuk le, hanem az maga a matematika, ami minket egy óriási matematikai objektum öntudatos részeivé tesz...[és] arra kényszerít, hogy lemondjunk a valóságról alkotott legmélyebben gyökerező elképzeléseink közül sokról."<sup>116</sup> A megfigyelők, beleértve az embereket is, akik ebben a matematikai objektumban léteznek, "öntudatos struktúrák", amelyek "szubjektíven úgy érzékelik magukat, mintha egy fizikailag "valós" világban léteznének".<sup>117</sup>

Mivel a MUH azt sugallja, hogy a matematikai létezés egyenlő a fizikai létezéssel - és minden matematikailag létező struktúra fizikailag is létezik -, a valóságot végső soron a kiszámítható függvények határozzák meg. Azáltal, hogy a jogot, a jogi fogalmakat, a normákat és az általuk érintett emberi és nem emberi viselkedésformák körét végső soron "kiszámítható" funkcióként határozzák meg, a LegalTech alkalmazások és a "mesterséges intelligencia bírák" iránti felhívások a társadalmi valóság és magának az információnak a természetéről alkotott sajátos felfogást támogatják. Egyrészt a bináris kódban kifejezett Yi-jingből levezetett Leibniz-féle dualizmust, másrészt a MUH által felállított tisztán digitális ontológiát, ahol minden kiszámítható, mert minden számítás. Ez nemcsak a számítás megoldatlan elméleti és gyakorlati korlátai miatt problematikus.<sup>118</sup>-ami felveti a tágabb kérdéseket: mi számítható? és számítható-e a jog? -, hanem azért is, mert a számítás bináris természete azt jelenti, hogy minden jogi problémának végső soron bináris logika segítségével kell eldönthetőnek lennie.

Míg Tegmark hipotézise jelentős figyelmet kapott a mesterséges intelligencia közösségen belül, a MUH továbbra is heves viták tárgyát képezi.<sup>119</sup> Piccini például megjegyzi, hogy a pánszámítástechnika mélyen vitatott: "egyes filozófusok szerint nyilvánvalóan hamis, túl ostoba ahhoz, hogy érdemes legyen megcáfolni; [míg] mások szerint nyilvánvalóan igaz, túl triviális ahhoz, hogy védelmet igényeljen".<sup>120</sup> McQuillan eközben összekapcsolja az adattudomány tágabb értelemben vett állításaival, amelyeket így jellemez:

---

<sup>114</sup> MacCormick, *What Can Be Computed*, 103-113.

<sup>115</sup> G Piccinini, *Fizikai számítás: A Mechanistic Account* (Oxford University Press, 2015) 51-73; M Miłkowski, "From Computer Metaphor to Computational Modelling: The Evolution of Computationalism" (2018) 28 *Minds and Machines*. 3.

<sup>116</sup> Tegmark, *Matematikai univerzumunk*, 17.

<sup>117</sup> M Tegmark, "Vajon a "Mindenek elmélete" csupán a végső együttes elmélet? (1988) *Annals* 270 of Physics 1.

<sup>118</sup> MacCormick, *What Can Be Computed*, 3-11, 294-311.

<sup>119</sup> J Schmidhuber, "Algorithmic Theories of Everything" (2000) <https://arxiv.org/abs/quant-ph/0011122>; P Hut, M Alford és M Tegmark, "On Math, Matter and Mind" (2006) *Foundations* 36 of Physics. 6.

<sup>120</sup> Piccinini, *Fizikai számítás*, 51.





...a kora újkori tudományt megalapozó neoplatonizmus visszhangja, amely Kopernikusz és Galilei munkásságában jelent meg. Vagyis egy olyan rejtett matematikai rendbe vetett hittel rezonál, amely on- tologiailag magasabb rendű, mint a mindennapi érzékszerveink számára elérhető.<sup>121</sup>

A MUH "digitális ontológiájával" szemben megfogalmazott erőteljes kritikák ellenére<sup>122</sup> és a strukturális realizmus információs megközelítésének széles körű támogatása mellett, amely szerint "a világról való tudás a világ struktúráinak ismerete",<sup>123</sup> a jelenlegi mesterséges intelligencia-kutatásokat nagyrészt az intelligencia pánkomputalista és materialista felfogása vezérli.<sup>124</sup>

A jogtudósok, látva a mesterséges intelligenciában rejlő lehetőségeket számos kontextusban, némileg beleszerettek az AGI-ba. A legtöbbször azonban a mesterséges intelligenciában rejlő határtalan lehetőségekkel kapcsolatos lelkesedésük nem több, mint a hétköznapi jogi kérdések "mágikus problémamegoldójának" karikírozása (pl. vállalatirányítási kérdések<sup>125</sup>). Ez általában azt jelenti, hogy figyelmen kívül hagyják az AGI által okozott szélesebb körű társadalmi átalakulást, és azt feltételezik, hogy a jogrendszer többé-kevésbé stabil marad az AGI utáni paradigmában. Egy másfajta megközelítés azonban megtalálható Ben Alarie jogtudós és LegalTech vállalkozó előrejelzésében, aki a közelgő "jogi szingularitást" jósolja.<sup>126</sup> A Ray Kurzweil által feltételezett techno- logikai szingularitásból eredeztethető,<sup>127</sup> Alarie "jogi szingularitás" olyan forgatókönyvet vizionál, amelyben az egyre nagyobb képességű és teljesítményű gépeket tervező gépek "intelligencia-robbanást" és az emberi megértést és ellenőrzést messze meghaladó "szuperintelligencia" kialakulását idézik elő.<sup>128</sup> Ha ez a pont elérkezik, "az elfogadott tények jogi jelentőségével kapcsolatos viták ritkák lesznek. Lehetnek viták a tényekről, de a [sic] egy megállapított tények egyértelmű jogi következményekkel járnak majd. A jog funkcionálisan teljes lesz".<sup>129</sup>

Bár az eredményeket nem szabad túlbecsülni, egyes rendszerek már "felülmúlták" az emberi jogi szakértőket az ügyek kimenetelének előrejelzésében.<sup>130</sup> Szakadék tátong azonban aközött, hogy statisztikai technikákat használnak-e az ügyek kimenetelének előrejelzésére olyan felszínes kritériumok alapján, mint például a joghatóság és a bírósági eljárás.

---

<sup>121</sup> D McQuillan, "Data Science as Machinic Neoplatonism" (2018) 31 *Philosophy & Technology* 2, 254.

<sup>122</sup> L Floridi, "A digitális ontológia ellen" (2009) *Synthese* 1681.

<sup>123</sup> J Smithies, *The Digital Humanities and the Digital Modern* (Palgrave 2017) 55; L Floridi, "A Defence of Informational Structural Realism" (2008) *Synthese*. 1612.

<sup>124</sup> Hassabis, Kumaran, Summerfield és Botvinck, "Idegtudományok által inspirált mesterséges intelligencia".

<sup>125</sup> M Petrins, "Corporate Management in the Age of AI" (2019) UCL Working Paper Series No3 <<https://dx.doi.org/10.2139/ssrn.3346722>>.

<sup>126</sup> Alarie, "A törvény útja", 1.

<sup>127</sup> R Kurzweil, *A szingularitás közel van: When Humans Transcend Biology* (Gerald Duckworth & Co. 2006).

<sup>128</sup> Bostrom, *Szuperintelligencia*.

<sup>129</sup> Alarie, "A törvény útja", 3.

<sup>130</sup> J Goodman-Delahunty, PA Granhag, M Hartwig és EF Lotus, "Insightful or Wishful: Lawyer's Ability to Predict Case Outcomes" (2010) *Psychology* 16, *Public Policy and Law* 2; B Alarie, A Niblett és A Yoon, 'Using Machine Learning to Predict Outcomes in Tax Law' (2017) <<http://dx.doi.org/10.2139/ssrn.2855977>>;

DM Katz, MJ Bommarito I és J Blackman, 'A general approach for predicting the behaviour of the Supreme Court of the United States' (2017) PLoS12 One. 4.

20

a bíró politikai hovatartozása, és a bírák tevékenységének helyettesítése mesterséges intelligenciával. Mindazonáltal a "jogi szingularitáriusoknak" nevezhető törekvések nem állnak meg az ügyek kimenetelének előrejelzésénél. Céljuk a jogi érvelésnek mint a vitarendezés alapjának a felváltása, és a jogkörök, jogok és felelőségek valamilyen proteikus triumvirátusának a legitim jogi tekintély helyébe való állítása.

## 8. Fejezet áttekintése

Az ebben a kötetben összegyűjtött előadások először a "Lex Ex Machina: A Law's Computability" című konferencián hangzott el, amelyet a Cambridge-i Egyetem Jesus College-jában tartottak 2019. december 19-én. Ez a konferencia abból a vágyból született, hogy összegyűjtsünk néhányat a jog és a technológia metszéspontján dolgozó legbefolyásosabb tudósokból, és megvizsgáljuk, hogy mit jelent a "számítógépes jog", a "ro-robot bírák" és a "jogi szingularitás" a jog mint társadalmi intézmény jövője szempontjából, valamint hogy visszavágjunk néhány szenzációhajhász állításnak, amelyek a jogi tudományosságból származnak a mesterséges intelligencia jogban való alkalmazásáról.

### 8.1 Simon Deakin + Christopher Markou - *Ex Machina Lex*

Az "Ex Machina Lex" című fejezetükben Christopher Markou és Simon Deakin egy határkérdést vet fel: milyen mértékben alkalmasak az ML/DL statisztikai technikái a jogi érvelés formalizálására? Miközben a mesterséges intelligencia jogban való felhasználásának új esetei jelennek meg, cikkünkben azt vizsgáljuk, ~~hogy~~ ezek sikeressége megvalósítható-e, és hogy mit jelentene a "siker" jogi kontextusban. Bár a számításnak és az adattárolásnak vannak anyagi és gyakorlati korlátai, és minden probléma megoldhatóságának elméleti korlátai, ez nem jelenti azt, hogy a jogi problémák szükségszerűen nem számíthatóak. Egyesek lehetnek, de nem feltétlenül az összes. A szerzők tehát azt vizsgálják, hogy a jogrendszer - a társadalmi rendszerek elmélete által implikált értelemben vett rendszer - milyen mértékben alkalmas a számítás és az automatizálás alkalmazására, és hogy a jogi érvelés stratégiai és számítógépes érveléssel való felváltása hogyan befolyásolhatja a jogrendszer autonómiáját, hogyan erodálhatja a jogállamiságot, és hogyan csökkentheti az állami tekintélyt a jogi kapcsolatok strukturálásában és közvetítésében. Megközelítésünk a jog rendszerszintű-evolúciós felfogását alkalmazza, hogy olyan egységesítő elveket határozzon meg, amelyek segítenek megmagyarázni a jogrendszer működését más társadalmi alrendszerekkel, köztük az öko- nomiaival, a politikával és magával a technológiával kapcsolatban, és amelyek segítenek tisztázni a jogi érvelés szerepét a jogi evolúció elősegítésében.

A fejezet azt sugallja, hogy a feltételezett "jogi szingularitás" - amely feltételezi az összes jogi bizonytalanság megszüntetését - a szimulációt, valamint az ML és a Big Data valószínűségi képességeit a jogi ítélezési folyamathoz kapcsolja. A szerzők Mireille Hildebrandt megfigyelésére építenek, miszerint "[a]mialatt a gépek nagyon jóvá válhatnak az ilyen szimulációban, maga az ítékezés a jogbiztonság bármely konkrét értelmezésének vitathatóságán alapul a jogrendszer integritásának fényében - ami túlmutat a kvázi matematikai konzisztencián -,".<sup>131</sup> és Alain Su- piot-t követve, aki a jog "antropológiai funkcióját" a "technológia humanizálásának technikájaként" azonosította, a következők szerint<sup>132</sup> azt állítják, hogy a jogi érvelésnek a számítással való felváltása végső soron a "jog uralmának" egy új "technológiai uralomnak" való alárendelését eredményezné.

---

<sup>131</sup> M Hildebrandt, "Law as computation in the era of artificial legal intelligence: speaking law to the power of statistics" (2017) <<https://www.ssrn.com/abstract=2983045>>.



A fejezet ezt a kritikát a számításlapú megközelítés mellett érvelők egyik kiemelt kérdésének vizsgálatán keresztül fejti ki, nevezetesen a munkavállalók és a független vállalkozók közötti különbségtétel adó- és munkaügyi célú meghatározásának vizsgálatán keresztül. Ez a különbségtétel történelmileg kontingensnek bizonyul, és számos ökológiai és politikai tényező alakította. A munka minősítésének jogi feladatát egy automatizált folyamatra redukálni elfedné azokat a politikai döntéseket, amelyek elkerülhetetlenül jelen vannak a jog területén.

## 8.2 Mireille Hildebrandt - Kódvezérelt jog: A múlt befagyasztása és a jövő skálázása

A "Kódvezérelt jog: Mireille Hildebrandt a "kriptográfiai jogot" vagy "intelligens szabályozást" (azaz az önvégrehajtó kódot mint a jogi szabályozás új típusát) vizsgálja. Először is fontos különbséget tesz az adatvezérelt jog és a kódvezérelt jog között. Az előbbi a jogi döntések előrejelzésére vagy a jogi érvek bányászatára vonatkozik, amelyet gyakran tárgyalnak a jogi mesterséges intelligencia vagy a LegalTech szakirodalomban, valamint Hildebrandt tágabb értelemben vett "jog mint számítás" című munkáiban.<sup>133</sup> Az utóbbi viszont a "jogi normákra vagy irányelvekre vonatkozik, amelyeket számítógépes kódban artikuláltak, akár egy szerződő fél, akár a jogalkalmazó hatóságok, akár a közigazgatás, akár a jogalkotó". Arra összpontosít, hogy ezek a kódvezérelt jogszabályok hogyan állapítanak meg előre meghatározott küszöbértékeket, amelyek "intelligens"/automatizált szabályozásokat váltanak ki (pl. társadalombiztosítási csalások felderítése).

Dolgozatának első része leírja, hogy mi a kódvezérelt jog, és elmagyarázza, hogy mit "csinál", azaz a szövegvezérelt jog. Példákat hoz a kódvezérelt jogra az intelligens szerződésekben (amelyek "nemcsak számítógépes kódban vannak megfogalmazva, hanem önvégrehajthatók is"), a közigazgatásban ("számítógépes kódban megfogalmazott döntéstámogató vagy döntéshozatali rendszer") és a törvényhozásban. Hildebrandt úgy véli, hogy a kódvezérelt törvények e példái kérdéseket vetnek fel a magánjog, a közjog, a büntetőjog és az alkotmányjog elveivel kapcsolatban. Arra a következtetésre jut, hogy a kódvezérelt jog "a törvényhozás, az értelmezés és az alkalmazás egy csapásra összefedi a jogalkotó, a végrehajtó hatalom és a bíróság közötti távolságot". A kódvezérelt jog implicit feltételezése az, hogy minden lehetséges forgatókönyvet előre lát, hogy lefedjen minden jövőbeli interakciót, amelynek szükségszerűen "rendkívül dinamikusnak és adaptívnek kell lennie, hogy kezelni tudja és szembe tudjon nézni azzal, amit nem lehet könnyen előre látni egyértelmű szabályokkal".

Hildebrandt ezután rátér a kódvezérelt jog kapacitás-feltevésére - vagyis arra, hogy a jogi normák jogba kódolásának pillanatában képesek vagyunk-e eléggé előre látni a jövőt. Azzal érvel, hogy a szövegvezérelt joggal szemben - amely a természetes nyelvre és a beszédaktusokra épül - a kódvezérelt jognak bizonyos korlátok vannak a számítógépes felépítésében és tervezésében. Ezek közé tartozik a formális dedukció ("ha ez, akkor az", IFTTT), a kifejezések és szabályok egyértelművé tétele, valamint a számítások hiányossága és következtelensége.<sup>134</sup> Ezek a korlátozások mind "a jövő bizonytalanságával kapcsolatosak". Az ML magáévá teszi azt a hamis feltételezést, hogy a képzési adatok eloszlása a jövő eloszlásának közeli közelítése lehet. A valóságban az adatok jövőbeli eloszlását az ML csak megjósolni tudja, nem pedig megtanulni.

<sup>133</sup> M Hildebrandt, "A jog mint számítás a mesterséges jogi intelligencia korában".

<sup>134</sup> vö. K Gödel, "Die Vollständigkeit der Axiome des logischen Funktionenkalküls" (1930) 37 Monatshefte für Mathematik und Physik 349; JM Rogers és RE Molzon, "Lessons about the Law from Self-Referential Problems in Mathematics" (1992) 90 Michigan Law Review. 5.

Hildebrandt azt állítja, hogy az előrejelzések befolyásolják azokat a viselkedéseket, amelyeket állítólagosan előrejeleznek, és a Goodhart/Campbell/Lucas-hatás újrafogalmazásával arra a következtetésre jut, hogy "amikor egy intézkedés célponttá válik, megszűnik jó intézkedésnek lenni". Bár sokféle jelenbeli jövőt tervezhetnénk, csak egyetlen valódi jelen (a valóság) létezik, és "a jövő megjóslásának legjobb módja, ha megteremtjük azt". Azzal zárja, hogy Arendt "emberi állapotát" a Parsons/Luhman- niai "kettős kontingencia" fogalmával lehet a legjobban magyarázni, és hogy a természetes nyelv az anticipáció folyamatát annak folyamatává teszi, hogy előre jelezzük, hogyan anticipálnak minket mások. A természetes nyelv értelmezhetőségén alapuló anticipáció koevolúciós jellege tehát radikális bizonytalanságot generál, ami viszont a társadalmi rendszerek komplexitásának és bizonytalanságának megszilárdítása, stabilizálása és csökkentése érdekében sajátos viselkedésminták kialakítását követeli meg. A jobbiztonság tehát döntő fontosságú a konszolidáció és stabilizáció e folyamatában, és a szövegvezérelt jellegéből adódóan a jövőnek a múlt skálázásán alapuló befagyasztása nélkül valósul meg.

Hildebrandt tanulmánya a kódvezérelt jog természetének vizsgálatával zárul, rámutatva a "jogi tervezés általi jogvédelem" koncepciójának problémáira és saját megközelítésének - "jogi védelem tervezés általi jogvédelem" - megfogalmazására. Ez utóbbi nem a jogi normák betartására vagy érvényesítésére összpontosít, hanem a jogi védelemnek a kódvezérelt és adatvezérelt környezetek technológiai infrastruktúrájába való beágyazására irányuló stratégiákra. Legalábbis azt állítja, hogy a kódvezérelt döntéshozatali rendszereknek hatékony jogorvoslati joggal kell rendelkezniük az automatizált döntésekkel szemben, és jogilag indokolni kell azokat. Végül következtetése azonban az, hogy a jog nem kiszámítható, mivel szövegvezérelt és többféleképpen értelmezhető, ami azt jelenti, hogy különböző módon lehet kiszámítani. A "jogi számítás" megtervezésének módját tehát inkább az "emberek" és a bíróságok, mint a nagy technológiai vagy jogi cégek szoftverfejlesztői választhatják meg.

### **8.3 John Morison - A demokratikus szingularitás felé?**

Az "Egy demokratikus szingularitás felé?" című fejezetben. John Morison az automatizált rendszerek bevezetésével kapcsolatos alapkérdéseket és a demokráciára és a jogrendszerek legitimitására gyakorolt hatásukat vizsgálja. A tanulmány először is felvázolja az online konzultációk ígéretének hiányosságait, és azt, hogy ezek még nem váltották ki demokratikus potenciáljukat. Morison ezután részletezi, hogy az új megfigyelési technológiák fejlődése, amelyek az emberi létezés gyakorlatilag minden aspektusát számszerűsíthető adatokká alakítják át, amelyeket a viselkedés megfigyelésére használnak, és nagyobb előrejelző képességeket tesznek lehetővé. Ezután megvizsgálja, hogyan használták ezeket a technológiákat a "biológiai felügyelet" lehetővé tételére a Covid-19 világjárvány összefüggésében.

Morison érvelése szerint a digitális technológiák konvergenciája a felügyelet és a társadalmi ellenőrzés egy új, mindent átható és átfogó formáját eredményezi, amelyet ő "algoritmikus kormányzásnak" nevez. A demokratikus konzultációk hiányosságaira épülő felismerésekre építve azt állítja, hogy ezek a fejlemények azzal a veszéllyel járnak, hogy a konzultációról és a deliberatív demokráciáról szóló elképzelések szinte teljesen feleslegessé válnak, "mivel a tényleges preferenciák közvetlenül mérhetők anélkül, hogy a preferenciák képviselőéhez szükség lenne egy közvetítő politikai folyamatra". E technológiák totalizációja - véli - lehetővé teszi az egyéni "preferenciák" megállapítását a statisztikai következtetés, a profilalkotás és a "radikális adatosítás" szélesebb körű folyamatának különböző eszközeivel, amely "hamis emancipációt kínál azáltal, hogy természeténél fogva mindenre kiterjedőnek és pontosnak tűnik".





Morison arra a következtetésre jut, hogy ezek a fejlemények a kormányzás újszerű, poszt-politikai formáját jelentik, és képesek először "aláásni, majd meghaladni az állampolgárság számos olyan alapvető tulajdonságát, amelyek jelenleg a kormány és a kormányzott közötti kapcsolaton belüli alku részeként jelennek meg". Következtetése az, hogy ellen kell állnunk a politika depolitizálására irányuló törekvéseknek azáltal, hogy kivonjuk azt a részletes döntéshozatal folyamatából, és algoritmikus kormányzással helyettesítjük, aminek viszont úgy kell ellenállnunk, hogy magukat a folyamatokat vitatjuk meg és kérdőjelezzük meg.

#### **8.4 Jennifer Cobbe - Jogi szingularitás és a jogi reflexivitása**

A "Jogi szingularitás és a jogi reflexivitása" című könyvében Jennifer Cobbe amellet érvel, hogy a jogi mesterséges intelligenciával foglalkozók - különösen az úgynevezett "jogi szingularitás" iránt érdeklődők - félreértik mind a jog, mind a technológia természetét. Azt állítja, hogy nemhog nem oldanák meg a jog nagyon is valós problémáit, de potenciálisan még rosszabbá is tehetik azokat, és új és nagyobb problémákat okozhatnak.

A különböző tudományágakból származó felismerésekre és fogalmakra támaszkodva Cobbe érvelése abból az elképzelésből indul ki, hogy a jog reflexív társadalmi intézményként működik. Szerinte a jog nemcsak tükrözi a társadalmat, hanem jelentősen befolyásolja a társadalmat és annak intézményeit. Mint ilyen, érvelése szerint a jog nem lehet semleges, mivel természeténél fogva kontextuális és a kor körülményeitől függő, valamint az adott kor normatív feltételezéseivel és prioritásaival átitatott. A jog végső soron alkotóinak (jogalkotók), alkalmazóinak (ügyvédek) és bírúinak (bírák) érdekeit és céljait tükrözi. Ezt a reflexív működést úgy kell érteni, hogy az különbözik attól a szereptől, amelyet a jog a társadalomban reflexivitása következtében játszik. Cobbe szerint a jog szerepe a társadalomban - történelmileg és napjainkban egyaránt - a tőke hatalmának megszilárdítása, a gazdagok pozíciójának megerősítése, az egyenlőtlenségek megerősítése és a fennálló érdekek védelme a külső kihívásokkal szemben. Cobbe szerint a jognak mint reflexív társadalmi intézménynek nemcsak az volt a szerepe, hogy tükrözze a társadalomban lévő egyenlőtlenségeket és igazságtalanságokat, hanem olyan társadalmi feltételeket teremt, amelyek megismétlik, megerősítik és újrakódolják azokat a társadalomban. Cobbe továbbá azt állítja, hogy az algoritmusok is reflexívek: "Ahogy a jog a maga reflexivitásában a társadalmat azok szubjektív feltételezéseit, felfogása és céljai szerint formálja, akik írják és gyakorolják" - mondja - "úgy az algoritmusok esetében is". Mint ilyen, ahogyan az is számít, hogy a jogban dolgozók milyen célokat és prioritásokat követnek, és kinek az érdekeit szolgálják, úgy az is számít, hogy milyen célokat és prioritásokat követnek az algoritmikus rendszerek tervezése, alkalmazása és használata során.

Miután a jog és a mesterséges intelligencia reflexivitásának leírásával lerakta az alapokat, Cobbe két részből álló érvelését fejleszti tovább. Az elsőben azzal érvel, hogy az algoritmikus rendszerek nem képesek - és talán soha nem is lesznek képesek - az emberi ügyvédek és bírák helyettesítésére. Rávilágít néhány hibás feltételezésre, amelyre a jogi mesterséges intelligencia hívei támaszkodnak, például arra az elképzelésre, hogy a mesterséges intelligencia rendszerek képesek jogi érvelésre, vagy semleges és objektív döntőbírók lehetnek. Hangsúlyozza azonban, hogy a jogi mesterséges intelligenciával kapcsolatos kritikák, amelyek az algoritmikus rendszerek technikai korlátaira összpontosítanak, bár fontosak, nem érintenek olyan strukturális kérdéseket, amelyek őt elsősorban érdeklik.

Érvelésének második részében Cobbe ezért éles kritikát fogalmaz meg a jogi mesterséges intelligencia hatalmi viszonyairól és strukturális hatásairól, ahogyan azt általában elképzelik. A kormányzatiasság foucault-i elméletét alkalmazva azzal kezdi, hogy a jogi mesterséges

intelligencia ideológiai alapjait és motivációit a neoliberális racionalizálási folyamat részeként határozza meg; a jog minőségi, ~~norma~~értékeit felváltják a statisztikai és gazdasági gondolkodáson alapuló, állítólag racionális, objektív, mennyiségi mérőszámok és logikák. E racionalizálási folyamat részeként a jog gyakran valószínűsíti a joggal kapcsolatos

a jogi mesterséges intelligencia hívei lassúnak, költségesnek, hatástalannak, összetettnek, kiszámíthatatlannak, optimalizálásra szorulónak, és így a techno-megoldáson alapuló beavatkozásra alkalmasnak tartják. Azzal, hogy a jogi mesterséges intelligencia mellett érvelők olyan kifejezésekkel érvelnek, amelyek alapvetően a jog működésének minőségével foglalkoznak, a jogi mesterséges intelligencia - és általánosabban a "jogi szingularitás" - hívei nem csupán a jog társadalomban betöltött szerepének természetét nem veszik figyelembe, hanem azt a fajta piacorientált és kereskedelemorientált gondolkodásmódot helyezik előtérbe, amely hozzájárul ahhoz, hogy e szereppel kapcsolatos problémák egyáltalán kialakuljanak. Cobbe szerint a jog társadalomban betöltött szerepének kritikai vizsgálata nélkül a jogi mesterséges intelligencia hívei ezért azt kockáztatják, hogy olyan rendszereket fejlesztenek ki, amelyek elsősorban a jog "jobbá" tételét teszik a hierarchiák kiterjesztésében és megerősítésében, a jog kirekesztő hatásainak fenntartásában, valamint a tőke dominanciájának és hatalmának újbóli megerősítésében.

### **8.5 Roger Brownsword - Mesterséges intelligencia és a jog funkcionális teljessége**

Roger Brownsword "A mesterséges intelligencia és a jog funkcionális teljessége" című fejezete azzal a megfigyeléssel kezdődik, hogy "számos technológiai éket vernek a jognak mint szabályalapú vállalkozásnak az elképzelése alá; hogy a jog csatornázó és újracatornázó funkciójával kapcsolatos ékek sokkal jelentősebbek, mint azok, amelyeket a bíraskodásba vernek", és hogy e technológiai ékek közül legalábbis néhány "először a vastag végébe megy". Brownsword azt állítja, hogy ezek a kihívások szükségessé tehetik "jogi gondolkodásunk radikális újraindítását", és potenciálisan átformálhatják "a jogállamiságot és a jog koherenciájáról alkotott elképzelésünket".

Dolgozatának érvelése három tengelyre bontható. Az első az, hogy az algoritmikus vagy "mesterséges intelligencia" eszközök állami és magánszabályozók általi használatát a szabályozási felelősség felülvizsgált fogalmának kell vezérelnie. Másodszor, azt javasolja, hogy ezt a felülvizsgált értelmezést, amelyet "a jogszerűség új mércéjeként" jellemez, be kell ágyazni a jogállamiságba, és ki kell fejezni azokban a szabályokban, amelyek kifejezetten a szabályozási hatáskör gyakorlását szabályozzák, amennyiben az technológiai intézkedésektől függ.". Harmadszor, Brownsword azt javasolja, hogy e szabályok felülvizsgálatához szükség van az általa "új koherenciának" nevezett módszerre, amely a szabályozási intézkedések és a jogszerűség újonnan megállapított mércéinek összeegyeztethetőségére és kölcsönös összefüggéseire összpontosít. A "hagyományos koherencialista gondolkodásmód" felülvizsgálatával - amely arra vonatkozik, hogy az általános jogi elvek hogyan alkalmazandók az egyes tényállási mintákra - Brownsword megfogalmazza, hogy "hogyan kell egy új koherencialista gondolkodásmódot kialakítani, hogy a technológiai intézkedések folyamatos ellenőrzése során ellenőrizni lehessen, hogy azok összeegyeztethetőek-e a szabályozási felelősségi viszonyítási pontokkal".

Brownsword azzal zárja tanulmányát, hogy olyan új intézményi konfigurációkra van szükség, amelyek jobban alkalmasak a koherens jogi érvelés táplálására és fenntartására, valamint "a szabályozók globális közösségeikért viselt felelősségének" támogatására.

### **8.6 Sylvie Delacroix - Automatizált rendszerek és a változás szükségessége**

Sylvie Delacroix "Automatizált rendszerek és a változás szükségessége" című fejezete a mesterséges intelligencia rendszerek alkalmazását vizsgálja az erkölcsi döntéshozatalban. Az egyik probléma ezzel az elképzeléssel az, hogy a mesterséges intelligencia nem jutnának el az erkölcsi tudatosság olyan formájához, amelyet mi a saját tapasztalatainkra jellemzőnek

ismerhetnék el. Különösen az erkölcsileg terhelt kontextusokban alkalmazható automatizált rendszerek kifejlesztésére irányuló jelenlegi erőfeszítések kevés figyelmet fordítanak azokra a nehézségekre, amelyek a morális változás elkerülhetetlen szükségességéből fakadnak. Ha azt a nézetet valljuk, hogy az etika egy "folyamatban lévő munka", akkor alapvető fontosságú lenne, hogy az értékképzés folyamatában megmaradjon az emberi szereplők szerepe.

Az olyan rendszerek, mint a jog, amelyek célja, hogy egyszerűsített útmutatást nyújtsanak az együttéléshez, konformitást idézhetnek elő, és eloszlatják a változásra irányuló nyomást. Az automatizált rendszerek hasonlóan "erkölcsileg kockázatosak". A veszélyek közé tartozik a gépi tanulás visszatekintő jellege. Delacroix az inverz megerősítő tanulás (IRL) használatát vizsgálja lehetséges megoldásként. Az IRL olyan rendszerek tervezését foglalja magában, amelyek a környező (emberi) ágensek viselkedéséből egy erkölcsileg megterhelt "hasznossági függvényt" következtetnek. Érdeklődés mutatkozik a mesterséges intelligencia alkalmazása iránt, hogy az adatok lehetővé tegyék a rendszer betanítását, elkerülve ezzel az adatok "kézzel történő" kiválasztásának szükségességét. A hasznossági függvény nem rögzített, hanem az ágensek erkölcsi ítéleteire vonatkozó új adatok fényében fejlődik.

Az erkölcsi döntések automatizálásának lehetősége vonzónak tűnhet annak fényében, hogy az ember nehezen tudja értékelné azokat a gyakorlatokat, amelyekhez hozzászokott. Az erkölcsi képesség azonban nem őrizhető meg pusztán a kognitív éberség révén, mivel ez utóbbit a gondolkodásnak az adott társadalmi környezetben való elmerülés során elsajátított szokásai kondicionálják. Egyesek számára, akik "realista" álláspontot képviselnek az erkölcsi kérdésekben, vonzó egy olyan szuperintelligencia lehetősége, amely az igazságosság útjára terel bennünket. Delacroix azonban arra figyelmeztet, hogy "rendszeres gyakorlás nélkül erkölcsi izmaink egyszerűen elsorvadnak, és képtelenek leszünk az együttélés alternatív, jobb módjainak mérlegelésére".

Szerinte pragmatikusabb cél lenne, ha "olyan ember-számítógép interakciókat hoznánk létre, amelyek alkalmasak arra, hogy elűzzék az erkölcsi álmodást". Ennek egyik módja az lenne, ha az embereket "végfelhasználóként" vonnánk be a döntési körbe, mint az "interaktív gépi tanulás" esetében. De "ha nem történik radikális változás azokban a tervezési döntésekben, amelyek azt a módot irányítják, ahogyan ezek a rendszerek interakcióra szólítanak fel bennünket, lusta, normatív állapotokat, akkor ez a hatás drámai lesz, olyannyira, hogy valószínűleg aláássa az ember által kiváltott változás lehetőségét".

## 8.7 Ryan Abbott + Alex Sarch - Mesterséges intelligencia büntetése

Ryan Abbott és Alex Sarch "A mesterséges intelligencia büntetése: azt a kérdést vizsgálja, hogy lehetséges-e büntetőjogi szankciók alkalmazását fontolóra venni a mesterséges intelligenciával rendelkező entitásokra. Rámutatnak, hogy a büntetőjog nem emberi entitásokra való alkalmazásának gondolata nem újdonság, mivel a vállalatok büntetőjogi felelőssége elvileg jól megalapozott. Szerintük a vállalati személyek büntetőjogi szankciókkal sújthatók, ha a szervezetek olyan rendszerszintű károkat okoznak, amelyek nem vezethetők vissza az egyes emberek cselekedeteire. Ha egy mesterséges intelligencia-rendszer képes hasonlóan "visszavezethetetlen" károkat - "kemény mesterséges intelligencia-bűncselekményeket" - okozni, akkor legalábbis alaposan meg kell fontolni a mesterséges intelligencia büntetésének lehetőségét.

Bár nincs közvetlen kilátás egy olyan "erős mesterséges intelligencia" létrehozására, amely képes pontosan lemásolni az emberi ~~kognitív~~ elképzelhető olyan eset, amikor a mesterséges intelligencia rendszerek nem csak kiszámíthatatlanul és megmagyarázhatatlanul, hanem minden közvetlen emberi irányítástól függetlenül is képesek viselkedni. A gépek ekkor olyan károkat okozhatnak, amelyek bűncselekménynek minősülnének, de ahol nincs azonosítható emberi személy, aki bűncselekménnyel vétkesen cselekedett volna, vagy ahol gyakorlatilag nem védhető a büntetőjogi vétkesség egyetlen emberi szereplőhöz kötése, mint például a sok egyén által egy bizonyos időn keresztül vagy egy összetett szervezeti környezetben végrehajtott cselekmények esetében. Ilyen körülmények között a mesterséges intelligenciával rendelkező szervezet megbüntetésének közvetve

visszatarthatja az azt kifejlesztő és forgalmazó személyeket, és a jog kifejező funkciója teljesülhet a mesterséges intelligenciával kapcsolatos egyes eredmények jogi elítélésével.

Abbot és Sarch ezután megvitatta, hogy lehetséges-e mentális állapotokat tulajdonítani a mesterséges intelligencia rendszereknek. Ezt már megtették - mutatnak rá - a vállalati entitások esetében, az ügynöki elméleteken keresztül, mint például a respondeat superior elv esetében. Ennek során a mesterséges intelligenciáért felelős felhasználó vagy fejlesztő mentális állapotát kell vizsgálni. Ez azonban azzal a problémával ütközik, hogy nem biztos, hogy egyértelmű, ki a felelős a mesterséges intelligenciáért. A súlyos károkért végső soron felelős személyek felelősségre vonására szolgáló büntetőjogi technikák - például a vélelmezett felelősség doktrínája - alkalmazása azzal a veszéllyel jár, hogy túlzott felelősséget rónak a fejlesztőkre, és elfojtják az értékes innovációt.

Azzal a kényes kérdéssel kapcsolatban, hogy a mesterséges intelligencia rendszereknek jogi személyiséget kell-e tulajdonítani, Abbot és Sarch figyelmeztetnek a "jogok kúszásának" veszélyére, ahogyan az már megtörtént, amikor számos joghatóságban az alkotmányosan védett emberi jogokat a vállalatoknak tulajdonították. Az egyik megoldás szerintük az lenne, ha egy "felelős személyt - a gyakorlatban valószínűleg egy vállalatot - jelölnének ki a mesterséges intelligencia végső garanciavállalójaként. A szerzők azt is megvizsgálják, hogy milyen költségekkel és előnyökkel járna, ha bizonyos mesterséges intelligencia-rendszerek esetében kötelező lenne a nyilvántartás és a jelentéstétel.

Arra a következtetésre jutottak, hogy "a kemény mesterséges intelligenciával kapcsolatos bűncselekmények növekvő lehetősége" ellenére "a mesterséges intelligencia büntetésének radikális eszköze" túlreagálás lenne. Jobb lenne "a büntetőjog szerény kiterjesztése, beleértve - ami a legfontosabb - az AI-alkalmazások helytelen tervezése, működtetése és tesztelése köré összpontosító új gondatlansági bűncselekményeket, valamint a törvényes kötelezettségeiket elmulasztó kijelölt felek lehetséges büntetőjogi szankcióit".

## **8.8 Lyria Bennett Moses - Not a Single Singularity (Egyetlen szingularitás sincs)**

Lyria Bennett Moses "Not a Single Singularity" című fejezete szerint a mesterséges intelligencia fejlődése a jogban három dimenzió mentén zajlik: egy "x tengely", amely azt írja le, hogy jelenleg mi lehetséges egy adott időpontban, egy "y tengely", amely arra utal, hogy a technológia mivé válhat, és egy "z tengely", amely mentén a mesterséges intelligencia legitimitásáról szóló döntések születnek majd. Úgy véli, hogy a jogi szingularitás felé vezető "egyenes út" helyett a három tengely mindegyikén valószínűleg váltakozó fejlődési és stagnálási időszakok lesznek, és hogy a végeredmény valószínűleg nem egy teljesen kiszámítható jogrendszer lesz.

A jogi mesterséges intelligenciával összefüggésben az emberi és a gépi képességek inkább kiegészítik, mint helyettesítik egymást. A gépek jelenleg képesek szimulálni bizonyos emberi feladatokat, és ha a hangsúly a kiszolgált funkciókon van, akkor úgy is felfoghatók, hogy a racionális viselkedés bizonyos formáit szimulálják. A mesterséges intelligencia-rendszerek jelenlegi állapota azonban olyan, hogy nagyon messze vannak attól, hogy az általános intelligenciát megközelítő eredményeket érjenek el. Ha az egyes technikákat közelebről megvizsgáljuk, élesen kirajzolódnak a joggal kapcsolatos korlátaik. Így a jogszabályok gépi olvashatóvá tételére tett kísérletek a szakértői rendszerek technikáinak alkalmazásával azt mutatták, hogy a legegyszerűbben azokat a jogszabályi szövegeket lehet kódolni, amelyek számítási elemeket tartalmaznak, vagy amelyek viszonylag egyszerű döntési szabályokra támaszkodnak. A jogszabályok kódolásra való törekvésének veszélye az, hogy ez megváltoztatja magának a jogszabályszerkesztésnek a módját. A szakértői rendszerek nem kódolhatják az olyan nyílt végű fogalmakat, mint az "ésszerűség", de az ilyen jellegű általános záradékokat, amelyek - hiányosságuknál fogva - lehetővé teszik a jogi alkalmazkodást.

Elektronikusan elérhető a következő címen:

<https://ssrn.com/abstract=3589184>

Hasonló problémák merülnek fel, amikor a természetes nyelvi feldolgozást és a gépi tanulást használják az esetek előrejelzésére. E technikák jelenlegi állapota nem megfelelő, mivel hajlamosak a faji és egyéb torzítások megismétlésére, amint arra a COMPAS szoftver esete rávilágított. De még ezek a ka-



a jövőben fokozható, a felhasználásukkal kapcsolatos legitimitási problémák merülnek fel. A jogi döntéshozatallal összefüggésben az ítélkezés nem pusztán előrejelző folyamat, hanem mérlegelés és viták eredménye.

Arra a következtetésre jut, hogy a jogban a mesterséges intelligencia egyenlőtlen fejlődése valószínűsíthető, nemcsak a technikailag megvalósítható korlátok miatt, hanem az automatizált döntéshozatal legitimitásával kapcsolatos elkerülhetetlen aggályok miatt is. A technológiai fejlődés "fejlődéshez és hibákhoz egyaránt vezet, és néha evolúciós és forradalmi módon egyaránt bővíti a rendelkezésre álló, a lehetséges, a megfelelő és a legitim lehetőségeket. De sok küszöböt kell átlépni, és nehéz elképzelni egy olyan rendszert, amely a jogot teljesen kiszámíthatóvá tenné anélkül, hogy a jog természete megváltozna.

### **8.9 Dilan Thampapilai - A vitatott fogalmak törvénye**

Dilan Thampapilai fejezetében, a "*A vitatott fogalmak joga? Reflections on Copyright Law and the Legal and Technological Singularities (Gondolatok a szerzői jogról és a jogi és technológiai szingularitásokról)*" című könyvében azzal érvel, hogy míg a jogi szingularitás gondolata magas absztrakciós szinten elég meggyőzőnek tűnik, egy olyan alkalmazott jogterület, mint a szellemi tulajdon (IP) esetében komoly ellenvetésekbe ütközik. A szellemitulajdon-jog a tulajdon, az emberi jogok, a szabad véleménynyilvánítás, a technológiai fejlődés és az oktatás fogalma köré épülő összetett korlátok között mozog, többek között. A szerzői jog csak olyan tökéletlen tulajdonjogi rendszerként működhet, amelyben a bíróságok mérlegelési jogkörükben esetről esetre mérlegelik az egymással versengő érdekeket és értékeket.

A tulajdonjog egyik formájaként a szellemi tulajdonjog a kizárás jogát fejezi ki. A kizárás azonban mindig is tökéletlen eszköz volt a szellemitulajdon-jog céljainak megvalósítására. A szerzői jog biztosítja a jogot a kizárólagos jogok megsértése miatti jogorvoslat igénybevételére, de nem garantálja a jogosultak érdekeinek teljes körű védelmét. A jog legitimitása attól is függ, hogy elismerik-e az al-ternatív érdekeket, és megengedik-e az olyan kivételeket, mint a tisztességes felhasználás doktrínája. A szerzői jog csak annak elfogadásával őrizheti meg legitimitását, hogy többféle érdeket szolgál.

A jogi szingularitás ezzel szemben a teljes jogrendszer egy olyan változatát írja le, amelyet egy emberfeletti intelligencia felügyel. Egy ilyen rendszer a jogi jogok tökéletes érvényesítésének lehetőségét feltételezi. Thampapillai szerint ez természeténél fogva ellentétes azzal a fajta többtényezős döntéshozatallal, amely a szerzői jogi ügyekben előfordul. A méltányos használat doktrínája úgy alakult ki, hogy lehetővé teszi a bíróságok számára, hogy az eléjük kerülő ügyekben összetett kompromisszumokat kössenek. Az ilyen egyensúlyok megteremtéséhez használt koncepciók éppen azért hasznosak, mert az értelmezés során nyíltak és vitathatók. A fogalmak egy olyan értékdimenziót is kifejeznek, amelyet semmilyen adatmennyiség nem képes egyenesen megragadni.

Mindezek miatt Thampapillai azt javasolja, hogy a jogi szingularitás vizsgálatokor szkepticizmusra van szükség. A szerzői jog példája rávilágít arra, hogy "az emberi rendszerek és az észérvek milyen mértékben bizonyultak rugalmasnak és alkalmazkodóképesnek" a technológiai változások kihívásaira való reagálás során; a jogi szingularitás víziója ezzel szemben olyan, amelyben nemcsak az emberi munka, hanem maga az emberi szubjektum is a perifériára szorul.

### **8.10 Christopher Markou + Lily Hands - *Capacitas Ex Machina***

A kötetet záró "*Capacitas Ex Machina: A szellemi cselekvőképesség számítógépes értékelése a*

mesterséges intelligencia "vörös vonala" vagy mércéje?" című kötetben a mesterséges intelligencia és a fejlett technológia alkalmazásának kilátása.

nológiák, mint például az fMRI és az automatizált mentális állapotfelismerés, Christopher Markou és Lily Hands foglalják. Dolgozatuk a mesterséges intelligencia történetének áttekintésével kezdődik, kezdve a szakértői rendszerek (ES) megjelenésével a 20. század közepén, a konnekcionista mesterséges intelligencia-keresés felemelkedésével annak második felében, az automatizált mentális állapotfelismerés (AMSD) és a kapcsolódó biokognitív interfészek kifejlesztéséig. Megvizsgálja e rendszerek valós világban történő megvalósításának elméleti és gyakorlati problémáit, valamint azt, hogy a technológiai fejlődés a közeljövőben várhatóan milyen hatással lesz a pszichiátriára.

A tanulmány ezt az egyelőre hipotetikus problémát azon történelmi erőfeszítések kontextusában helyezi el, amelyek arra irányulnak, hogy a számítógépek segítségével segítsék az orvosi diagnózist, és potenciálisan helyettesítsék az emberi orvosokat és pszichiátereket. Ezután azt vizsgálja, hogy a számítógépes érvelés működhet-e, és hogyan kellene-e működnie az Angliában és Walesben a cselekvőképességgel kapcsolatos döntések kontextusában. Joseph Weizenbaum informatikus kritikáját követve Markou és Hands nemcsak a technikai kihívásokra hívja fel a figyelmet, amelyekkel az orvosi szakismeretek rögzítése, kódolása és alkalmazása során szembesülni kell, hanem arra a normatív -és mélyen politikai - kérdésre is, hogy egy számítógépet bármilyen körülmények között fel kell-e hatalmazni arra, hogy szubjektív ~~pszich~~ és pszichiátriai jelenségek alapján jogi következményekkel járó döntéseket hozzon. A kötetben Hildebrandt, Morison és mások által korábban azonosított problémák közül sokat kiemelve azzal érvelnek, hogy a mentális képtelenség megállapításának engedélyezése egy gép számára nem csupán - legalábbis egyelőre - technikailag kivitelezhetetlen, hanem hogy "a cselekvőképességgel kapcsolatos döntések alapvető emberi jellege és következményei nemcsak az egyénre, hanem a ~~köz~~ nézve is megkövetelik, hogy a cselekvőképességet ne csak meghatározzák, hanem a közösség tagjai értékeljék és hozzárendeljék". Arra a következtetésre jutottak, hogy "minden más azt eredményezné, hogy a gépet az emberi viselkedés és tapasztalat döntőbírójává emelnék egy olyan társadalmi valóságban, amelyhez nem fér hozzá".

## 9. Következtetés

A kötetet alkotó fejezetek egy olyan szándékos erőfeszítés részét képezik, amely visszavág a mesterséges intelligenciáról szóló, inkább hagiográfiai jellegű beszámolóknak, és segít meghatározni azokat a területspecifikus kihívásokat, amelyeket a fejlesztés során figyelembe kell venni, mielőtt a valós jogi kontextusban megbízhatóan alkalmazhatók lennének. Reméljük, hogy ez a kötet szókratikus belépési pontként szolgál a jogászok, jogtudósok és hallgatók számára egyaránt, ugyanakkor segít áthidalni a szakadékot a mesterséges intelligencia kutatásának technikai dimenziói és a hatékony politikára és kormányzásra gyakorolt normatív következményei között.

Bár az e kötetben szereplő hozzászólások az igazságszolgáltatás automatizálásával kapcsolatos jogi, etikai és politikai kihívások egész sorát azonosítják, segítenek jobban összekapcsolni a jelenlegi és közeljövőbeli kihívásokat azokkal, amelyekkel hosszabb távon kell szembenéznünk. Talán nem az az igazi dán- ger, hogy a jogi szingularitás bekövetkezik, hanem az, hogy amit a modern mesterséges intelligencia vállalkozás "igazságosnak" akar az igazságnak, az valójában nem több, mint amit "elégségesnek", "éppen elégségesnek" tartanak. Egy AGI-paradigmában az AI-bírák kilátása talán triviális. Egyszerűen nem tudjuk, és nem is tudhatjuk. De van okunk azt gondolni, hogy ez egy olyan forгатókönyv lehet, ahol az emberiség számára minden esélyünk elveszett, és egyes vélemények szerint az esélyeink nem jók.<sup>135</sup>

Egy hipotetikus érvelés hosszú távú és meghatározatlan feltételes módokon való rögzítése azonban csak egy kidolgozott módja annak, hogy "1, 2, 3, 1 millió" számoljunk, és ezt jó számtani érzéknek és megbízható üzleti pozíciónak nevezzük. Míg ez utóbbi egyesek számára mégis igaznak tűnik, addig ez csak a totalizálás keretét adja.

---

<sup>135</sup> E Yudkowsky, "A mesterséges intelligencia mint a globális kockázat pozitív és negatív tényezője" in N Bostrom és MM

Ćirković (szerk.) *Global Catastrophic Risks* (Oxford University Press 2008).

A mesterséges intelligencia a jogban elkerülhetetlen, amint az "intelligencia" egy önkényesen meghatározott küszöbértékét elérjük, vagy különböző funkcionális képességeket elérünk. Ahogy Powles és Nissenbaum érvel:

...a szűk körű számítási rejtvényekkel való foglalatosság elvonja a figyelmünket a sokkal fontosabb kérdésről, a társadalmi költségek és a magánnyereség közötti kolosszális aszimmetriáról az automatizált rendszerek bevezetése során. Azt is megtagadja tőlünk, hogy megkérdezzük: Egyáltalán meg kellene-e építenünk ezeket a rendszereket?<sup>136</sup>

Ez a kérdés inspirálta ezt a kötetet, és őszintén reméljük, hogy másokat is arra ösztönöz, hogy munkájuk során hangosan és gyakran tegyék fel ezt a kérdést.

## **TERVEZET CSAK ÉSZREVÉTELEK MEGTÉTELÉRE**

---

<sup>136</sup> J Powles és H Nissenbaum, "The Seductive Diversion of "Solving" Bias in Artificial Intelligence" (*Medium*, December 72018) <<https://medium.com/s/story/the-seductive-diversion-of-solving-bias-in-artificial->

Elektronikusan elérhető a következő címen:  
<https://ssrn.com/abstract=3589184>

