

## A robotikai jog útja

Jack M. Balkin\*

### BEVEZETÉS

Melyek azok a fő problémák, amelyeket a robotika a jog számára jelent? E kérdés megválaszolásához Ryan Calo szerint a kiberjoggal kellene analógiát vonnunk. Először is meg kell határoznunk az új technológia "alapvető tulajdonságait", majd fel kell tennünk a kérdést, hogy a jognak hogyan kell válaszolnia az ezen alapvető tulajdonságok által felvetett problémákra.<sup>1</sup>

Calo beszámolója a robotika által a jog számára jelentett problémákról egyszerűen fantasztikus, és úgy vélem, hogy a jövőben sok kutatás kiindulópontja lesz ezen a területen. Ez az esszé az ő gondolataira épít, és egy kissé más irányba viszi őket.

Először is, én más tanulságokat vonok le, mint Calo a kiberjog történetéből, valamint a jog és az új technológia találkozásából. Nem hiszem, hogy hasznos egy új technológia "alapvető tulajdonságairól" beszélni, amelyeket aztán a jogra is alkalmazhatunk. Éppen ellenkezőleg, meg kell próbálnunk nem úgy gondolkodni a technológia jellemzőiről, mintha ezek a jellemzők függetlenek lennének attól, hogy az emberek hogyan használják a technológiát az életükben és a másokkal való társadalmi kapcsolataikban. Mivel a technológia társadalmi életben való használata fejlődik, és mivel az emberek folyamatosan új módokat találnak arra, hogy a technológiát jóra vagy rosszra használják, nem biztos, hogy hasznos lenne a használat bizonyos jellemzőit egy adott pillanatban megragadni és "alapvetőnek" címkézni. Másodsor, a robotika által a jog számára felvetett problémákat kicsit másképp írom le, mint Calo, részben azért, mert nem teszek éles különbséget a robotok és a mesterséges intelligencia (MI) ágensek között. Az innováció előrehaladtával a

---

\*Knight professzor az alkotmányjog és az első alkotmánymódosítás területén, Yale Law School.

1. Ryan Calo, *Robotics and the Lessons of Cyberlaw*, CALIF103. L. REV. 513, 514-15 (2015).





A kétféle technológia közötti különbségtétel a jog szempontjából talán sokkal kevésbé fontos, mint amilyenek jelenleg tűnik; még nem tudjuk, hogy a két technológia közötti határok egyre inkább elmosódnak-e (ahogyan azt várom), vagy tovább differenciálódnak.

Ehelyett két kulcsfontosságú problémát emelek ki, amelyeket a robotika és a mesterséges intelligencia jelent a jog számára. Az első az, hogy hogyan osszuk el a nem emberek cselekedeteiből eredő emberi jogokat és felelősségeket. Az emberi lények a robotokat és a mesterséges intelligenciával rendelkező entitásokat arra fogják használni, hogy új dolgokat hozzanak létre, és hogy megsértsék más emberi lények jogilag védett érdekeit. Az AI-ügynökök regényeket fognak írni, de kémkedni is fognak az emberek után, és rágalmazni fogják őket. A robotok új dolgokat fognak feltalálni, de el is fogják törni az emberek végtagjait. Calo azt sugallja - szerintem helyesen -, hogy még messze vagyunk attól, hogy a robotokat és az AI-ügynököket öntudatos, jogokat vagy felelősséget hordozó entitásokként kezeljük.<sup>2</sup> Ezért a jog számára - legalábbis a történelem e korai szakaszában - az a kulcskérdés, hogy hogyan osszuk el a jogokat és a kötelezettségeket az *emberek* között, amikor a robotok és a mesterséges intelligenciával rendelkező entitások előnyöket teremtenek vagy sérüléseket okoznak.

A második probléma az első általánosítása. Én ezt "helyettesítési hatásnak" nevezem - annak a ténynek, hogy az emberek a robotokat és a mesterséges intelligencia-ügynököket élő dolgokkal, és különösen emberekkel helyettesítik. De ezt csak bizonyos módon és csak bizonyos célokra fogják tenni. Ez a helyettesítés valószínűleg kontextusfüggő, instabil és gyakran opportunista. A helyettesítés problémája a jog számos különböző területét érinti, és az ígéretek szerint még nagyon sokáig zavarba fog hozni bennünket.

## I.

### A KIBERJOG TANULSÁGAI

Oliver Wendell Holmes Jr. 1897-es híres előadásában, *A jog útja* című művében amellest<sup>3</sup> érvelt, hogy mivel a jogot a társadalmi élet kumulatív erői hozzák létre, a jogot társadalmi funkciója és gyakorlati felhasználása szempontjából kell vizsgálnunk. Holmes egy protorealista kiáltványt kínált a formalizmus és a jogi fogalmak lényegi jellemzőibe vetett hit ellen. Minden amerikai jogász, Calo és jómagam is, Holmes pragmatizmusának szellemi leszármazottja. A jog esszencializmusának elutasítása segít nekünk abban, hogy a jog társadalmi kapcsolatokban rejlő alapjára és a jognak a tekintély- és hatalmi viszonyok fenntartásában - és esetenként misztifikálásában vagy mentegetésében - játszott szerepére összpontosítsunk. Hasonló antiformalizmust kellene alkalmaznunk, amikor a technológiáról és a jognak az innovációval való találkozásáról gondolkodunk.

Amikor azt vizsgáljuk, hogy egy új technológia hogyan hat a jogra, nem arra kell összpontosítanunk, hogy mi a technológia *lényege*, hanem arra, hogy a technológia a társadalmi élet mely jellemzőit *emeli* újonnan *előtérbe*. Milyen problémákat helyez előtérbe egy új technológia, amelyeket korábban nem hangsúlyoztak vagy kevésbé tekintettek fontosnak.

2. *Lásd id.* at (517a robotok közeljövöbeli képességeinek "konzervatív" megítélése), (539"Keveset nyerünk, és sokat veszítünk azzal, ha úgy teszünk, mintha a mai robotok bármilyen szándékot mutatnának.").

3. Oliver Wendell Holmes, Jr., *The Path of the Law*, 10 HARV. L. REV. 457 (1897).

fontos? Az emberi tevékenység vagy az emberi lét mely aspektusait helyezi előtérbe, hangsúlyozza vagy problematizálja a technológiai változás? És ha a technológia egy bizonyos problémát vagy életünk egy bizonyos jellemzőjét a háttérből az előtérbe helyezi, milyen következményekkel jár az emberi szabadságra nézve?<sup>4</sup>

A technológiánk legmarkánsabb jellemzői attól függhetnek, hogy az emberek hogyan használják azt. Gyakran előfordul, hogy az emberek úgy használják a technológiát, ahogyan azt a tervezők nem látták előre vagy nem tervezték. Ez különösen igaz a Jonathan Zittrain által "generatív" nevezett technológiákra, <sup>5</sup>amelyek számos helyszínt és lehetőséget kínálnak az innovációra. Ami különösen fontosnak és kiemelkedőnek tűnik a technológiában, az idővel változik, ahogy az emberek az új technológiával és az új technológia segítségével dolgoznak. Ráadásul a technológia, akárcsak a jog, közvetíti az emberek közötti társadalmi kapcsolatokat - beleértve a hatalmi és ellenőrzési viszonyokat is. Mivel ezek a viszonyok mindig változnak, az is változhat, hogy mi a legérdekesebb vagy legaggasztóbb egy technológiában.

Ma az emberek a pénztárcájukban és a zsebükben mikroszámítógépeket hordanak, amelyeket "mobiltelefonoknak" neveznek. Harminc évvel ezelőtt az emberek talán azt állították volna, hogy a mobiltelefon lényeges tulajdonsága, hogy az ember otthonán kívül is tud telefonálni. De a mobiltelefonoknak korántsem ez a tulajdonsága az elsődleges módja annak, hogy az emberek ma már használják őket. Ehelyett az emberek szokásai a mobiltechnológia fejlődésével párhuzamosan alakultak ki. Amire ma az emberek a telefonjukat használják - a különböző alkalmazások kaleidoszkópján keresztül -, az a "telefon" szó is egy elmúlt kor reliktumának tűnik.

1991-ben, a World Wide Web széles körű elterjedése előtt az internetet főleg e-mailre és fájlcsere-re használták olyan protokollokon keresztül, mint a Gopher. (A keresési funkciók kezdetlegeseek voltak, és a Google-t még nem találták fel.) Akkoriban úgy képzelhettük, hogy az internet "alapvető" jellemzői a joghatósági határok tetszőleges átlépésére, a digitális információk gyors és olcsó továbbítására, valamint az anonim kommunikáció megkönnyítésére való képességek voltak. De ez azért volt így, mert a történelemnek azon a pontján az emberek így képelték el, hogy az emberek így fogják használni a technológiát, és hogy ezek a használatok milyen társadalmi hatalmi, cselekvési és kiszolgáltatottsági viszonyokat fognak létrehozni. Röviden, egy új technológia jellemzői részben a jelenlegi használat termékei, részben pedig az emberi képzelet munkája a potenciális lehetőségekről és lehetőségekről, veszélyekről és fenyegetésekről.

1999-re világossá vált, hogy az államok ellenőrizhetik az internetes adatforgalom jellemzőit, és hogy az internet által biztosított valóban anonim kommunikáció mértékét túlbecsülték. A területi kormányok hatókörétől mentes, többszörös kibervilágról szóló álmok átadták helyüket a szűrővel kapcsolatos növekvő aggodalmaknak.

---

4. Jack M. Balkin, *Digitális beszéd és demokratikus kultúra: A véleménynyilvánítás szabadságának elmélete az információs társadalomban*, N79.Y.U. L. REV. 1,2 (2004).

5. JONATHAN ZITTRAIN, AZ INTERNET JÖVŐJE: (2008)67.

és felügyelet. Lawrence Lessig *Code* című könyvének első kiadása<sup>6</sup> hangsúlyozta, hogy az internet mennyit változott az 1995 és 1999 között. Az egyik fő téma az volt, hogy ezek a változások megmutatták, hogy az internetet nem kell semmilyen módon megtervezni, és hogy az internet vagy jellegzetesen szabadabbá, vagy ellenőrzötté válhat.<sup>7</sup> Hasonlóképpen, ahogy Yochai Benkler rámutatott, az internet nem egyetlen kommunikációs médium - akár egy a sokhoz, akár sokan az egyhez, akár sokan a sokhoz -, hanem inkább egy olyan platform, amely többféle médiumot hoz létre, sok különböző tulajdonsággal, lehetőséggel és társadalmi cselekvési lehetőséggel.<sup>8</sup> Mire Benkler 2006-ban megírta *A hálózatok gazdagsága* című könyvét, Clay Shirky pedig 2008-ban megjelentette a *Here Comes Everybody* című könyvét,<sup>9</sup> az internet ismét átalakulni látszott. Most az volt a legfeltűnőbb jellemzője, hogy képes volt csökkenteni a szervezés költségeit, megkönnyíteni a crowdsourcingot és a nyílt forráskódú projekteket, valamint aláásni az információtermelés professzionális normáit a zenétől az újságíráson át a tudományig terjedő területeken. Calo saját beszámolója az internet<sup>10</sup> "alapvető" jellemzőiről inkább egy visszatekintő nézet, amely a társadalmi és technológiai innováció több mint két évtizedére épül, mintsem az eredendő tulajdonságok időtlen beszámolója.

Az 1991-es internet annyira különbözik a 2015-ös internettől, hogy szinte elképzelni is nehéz. Bár vannak családi hasonlóságok a kettő között, a különbségek óriásiak, ami a számtalan ember elképesztő leleményességének köszönhető, akik a társadalmi és gazdasági kapcsolatokban éppúgy újítottak, mint a hardver és a szoftver terén.

Azt mondhatnánk tehát, hogy a robotika számára a kiberjog legfontosabb tanulsága az, hogy figyelni kell az affordancia és a képzelet, az eszközök és a hatalmi viszonyok, a technológiai szubsztrátum és a társadalmi használat közötti kapcsolatokra. A robotika azon jellemzői, amelyek jelenleg felhívják a figyelmünket, amelyek megzavarják a jogi önelégültségünket, amelyek rejtélyeket okoznak a jog számára és lehetőségeket a társadalmi innováció számára, nem az utolsó rejtélyek és lehetőségek lesznek, amelyekkel szembe kell néznünk.

Valójában az, amit a jogászok "technológiának" neveznek, általában valami sokkal összetettebb dolog rövidítése. Amikor "technológiáról" beszélünk, akkor valójában arról beszélünk, hogy (1) az emberek hogyan lépnek kapcsolatba az új találmányokkal, és (2) az emberek hogyan

---

6. LAWRENCE LESSIG, KÓD: (2005, 1. 216. kiadás). (1999).

7. *Id.* at (25) "A hálózatnak nem egyetlen módon kell működnie; nincs egyetlen architektúra, amely meghatározza a hálózat jellegét."

8. *Lásd* YOCHAI BENKLER, THE WEALTH OF NETWORKS: HOW SOCIAL PRODUCTION TRANSFORMS MARKETS AND FREEDOM 370 (2006) ("[T]he Internet allows a radically more diverse suite of communications models than any of the twentieth-century systems allowed.")

9. CLAY SHIRKY, ITT JÖN MINDENKI: (2008).

10. Calo, *Supra* note, 1,518-25. o. Calo szerint az idő múlásával az internet kulcsfontosságú jogi problémái egyértelműbbé és stabilabbá váltak. *Id.* 525-27. o. Ha ez igaz - és



én valamivel szkeptikusabb vagyok, mint Calo -, az azért van, mert visszatekintve tisztábban látjuk az utat, amelyet a technológiai és társadalmi innováció több mint húszéves folyamatos kölcsönhatása során bejártunk.

kölcsönhatásba lépni más emberekkel, akik ezeket az új találmányokat használják, vagy feltételezik ezeket az új találmányokat. Amit mi a technológia hatásainak nevezünk, az nem annyira a *dolgok* jellemzői, mint inkább az ezeket a dolgokat alkalmazó *társadalmi kapcsolatok* jellemzői. Ezek a társadalmi viszonyok magukban foglalják a hatalmi és tekintélyi viszonyokat, az emberi szervezet és termelés formáit, valamint az emberi szexualitás, társulás és családi élet jellemzőit.

Hasonlóképpen, a technológiai innováció nem csak az eszközök és technikák innovációját jelenti, hanem a gazdasági, társadalmi és jogi viszonyok innovációját is magában foglalhatja. Ahogyan társadalmilag és gazdaságilag újítunk, úgy változhat az is, hogy mi tűnik a legjelentősebbnek és legfontosabbnak az eszközeinkkel kapcsolatban. Ha az 1991-es internet gyökeresen másképp néz ki, mint a 2015-ös, az nem csak azért van, mert az eszközeink fejlődtek, hanem azért is, mert az emberek életmódja, munkája és egymással való kapcsolata is fejlődött az eszközökkel való kölcsönhatásban. Ahogy a világunk megtelt robot- és mesterséges intelligencia technológiákkal, életünk és társadalmi, politikai és gazdasági erőviszonyaink is megváltoznak, ami új és váratlan kihívások elé állítja a jogot.

## II.

### MIT EMEL KI A ROBOTIKA ÉS A MESTERSÉGES INTELLIGENCIA?

#### A. *Legyünk fizikálisak*

Calo kifejti, hogy a robotok új problémákat vetnek fel a jog számára, mivel a robotok nem csak gazdasági vagy érzelmi károkat, hanem fizikai károkat is okozhatnak.<sup>11</sup> Calo ezt a tulajdonságot "megtestesülésnek" nevezi,<sup>12</sup> de ez a kifejezés kissé félrevezető lehet. A számítógépek mindig is fizikai formában testesültek meg - gondoljunk csak a múlt hatalmas IBM nagyszámítógépeire. Sőt, a számítógépek fizikai megtestesülésük miatt mindig is képesek voltak fizikai kárt okozni - például egy laptop fizikai sérülést okozhat, ha eldobják, vagy ha ráesik egy ember lábára, és eltörik a lábujja.

Calo inkább azt állítja, hogy a robotok a *programozásuk*, pontosabban a hardver, az operációs rendszer és a szoftver együttes hatása *miatt* okozhatnak fizikai károkat.<sup>13</sup> Ezek teszik lehetővé, hogy a robot kölcsönhatásba lépjen a környezetével, és fizikai hatást gyakoroljon a világra. Tehát bár úgy tűnhet, hogy Calo aggodalma elsősorban a fizikai károkozás képességével kapcsolatos, a mélyebb kérdés az, hogy a kárt - legyen az fizikai, érzelmi vagy pénzületi - a programozhatóság és az interaktivitás okozza. Az a tény, hogy a robotoknak testük van, más szerepet játszik abban, amit Calo "társadalmi valenciának" nevez - többek között abban, hogy az emberek hajlamosak úgy kezelni az interaktív mozgó tárgyakat, mintha azok élnének.<sup>14</sup> De itt sem az a fontos, hogy a robotoknak fizikai testük van.

---

11. *Id.* 515,542.

12. *Id.* at13.532.

*Id.* 533-35. o. 14.



formája - így a kenyérpíritóké is -, de interaktivitásuk sajátos társadalmi jeleket vált ki az emberekből.

Az, hogy egy robot képes fizikai sérülést okozni, nem igazán a robottechnológia "alapvető" jellemzője. A jogászok számára ez a robotika különösen *kiemelkedő* jellemzője, de azért kiemelkedő, mert - amint Calo kifejti - a bírák már az internet korai napjaitól kezdve úgy döntöttek, hogy a meglévő kártérítési jogi doktrínákat úgy alakítják át, hogy korlátozzák az új digitális technológiák által okozott tisztán gazdasági károkért való felelősséget.<sup>15</sup> Hasonlóképpen, a Kongresszus egy sor olyan jogalkotási mentességet hozott létre az online szolgáltatók számára, amelyek korlátozták a szerzői jogok megsértéséért vagy a digitális hálózatokat használó káros kommunikációért való felelősséget.<sup>16</sup> A fizikai károkozás problémája nem egyszerűen a technológia alapvető jellemzőinek sajátja. Inkább abból adódik, ahogyan egy új technológia kölcsönhatásba lép a már meglévő társadalmi és jogi világgal.

Egy új technológia egy már mozgásban lévő társadalmi világba lép be, ahol már létező feltételezések és elvárások vannak arról, hogy mi a valószínű és mi a valószínűtlen, mi a lehetséges és mi a lehetetlen, a társadalmi életre vonatkozó normák, valamint paradigmaticus példák vannak arra vonatkozóan, hogy hogyan működnek a dolgok, és mit kell tennünk a problémákra adott válaszként. A jog szempontjából ezeket a feltételezéseket és elvárásokat arról, hogy hogyan működik a világ, a *szabályozás színterének* nevezhetnénk.<sup>17</sup> Az új technológia megzavarja a szabályozás meglévő színterét, és a különböző szereplőket arra készíti, hogy összevesszenek azon, hogyan fogják és hogyan kell használni a technológiát. Miközben az emberek a technológiával kapcsolatban egymásnak feszülnek és egymással vitatkoznak, nem csak technológiai, hanem társadalmi, gazdasági és jogi szempontból is újításokat hajtanak végre, ami a jog számára új problémákhoz vezet. Ahelyett, hogy azt mondanánk, hogy a jog az új technológia alapvető jellemzőire reagál, talán jobb lenne azt mondani, hogy az új technológia felhasználásával kapcsolatos társadalmi küzdelmek a jog meglévő jellemzőibe illeszkednek, és megzavarják a helyzetek kategorizálására vonatkozó elvárásokat.

A programozható egységek fizikai károkozásra való képességének a robotika "alapvető jellemzőjeként" való minősítése egy másik hátrányt is rejt magában. Elhanyagolhatjuk a mesterséges intelligenciát és öntanulást alkalmazó rendszerek sokféleségét. Valójában, ahogy Calo rámutat, a "robotok" és a "mesterséges intelligencia" között kontinuitás van.<sup>18</sup> Ez azért van így, mert - akárcsak maga az internet - a robotokat és más interaktív entitásokat sem kell egy bizonyos módon megtervezni. És mivel a lehetséges kialakítások és a különböző felhasználási lehetőségek kontinuumma létezik, az új technológiák lehetséges hatásainak is kontinuumma van. Az öntanuló algoritmusok növelhetik vagy csökkenthetik a hőmérsékletet.

---

15. *Id.* 536-37.

16. *Id.*; Online Copyright Infringement Liability Limitation Act, U17.S.C. § 512(a)(1) (2012); Communications Decency Act of U.1996,47S.C. § (2302012).

17. *Vö.* Jack M. Balkin & Reva B. Siegel, *Principles, Practices, and Social Movements*, 154 U.

PA. L. REV. 927, 931 (2006) (a "szabályozási színteret" úgy definiálja, mint "a problémák paradigmaticus halmazát", amelyet egy jogi "elvnek meg kell oldania, és a paradigmaticus gyakorlatok sajátos halmazát, amelyet az elv legitimál vagy delegitimál").

18. Calo, *fentebb*, 1,530-32. o.

egy házban, bekapcsolja a készülékeket, bezárja vagy kinyitja a kapukat, és értesíti a biztonsági szolgálatot. Az algoritmusok képesek értékpapírokat vásárolni és eladni; képesek holografikus vetületeket létrehozni, amelyek úgy néznek ki és úgy viselkednek, mint az emberek; képesek fenyegetni, szórakoztatni, másolni, rágalmazni, becsapni, figyelmeztetni, vigasztalni vagy csábítani. Ezek a különböző hatások a fizikai, a gazdasági, a társadalmi és az érzelmi határokat feszegetik.

Félrevezethet bennünket, ha ragaszkodunk a robotika és a mesterséges intelligencia rendszerek közötti túl éles megkülönböztetéshez, mert még nem ismerjük a technológia fejlesztésének és alkalmazásának minden módját. Hasonlóképpen félrevezethet bennünket az is, ha a "lényeges" jellemzők elméletének megőrzése érdekében megpróbáljuk ezeket a hatásokat és károkat "fizikai" hatásokba és károkba zsúfolni. Egy robotmacska, amelyet egy idős ember számára társnak terveztek, hogy simogassa és ápolja, fizikai és érzelmi hatásokkal jár. Egy intelligens otthon, amely szabályozza a hőmérsékletet, megrendeli a bevásárlást, és tájékoztat minket a napunkról, egyesíti a fizikai, a gazdasági és a társadalmi vonatkozásokat. Figyelemmel kell lennünk a technológiák felhasználásának sokféleségére, és ezért az e technológiák által okozott károk és veszélyek sokféleségére is.

Egy technológia számos különböző jellemzőjét azonosíthatjuk kulcsfontosságú vagy alapvető tulajdonságként, de az igazi kérdés mindig az, *hogyan miért* érdekelnek minket ezek a tulajdonságok. Az, hogy miként határozzuk meg egy technológia központi jellemzőit, attól függ, hogy mire *szolgál* a meghatározásunk, és milyen célt szolgál az adott kutatási területünkön. Ha mérnökök vagyunk, akik egy tervezési problémát próbálnak megoldani, akkor a technológia más aspektusaira fogunk összpontosítani, mint ha jogászok vagyunk, akik a jogi jogok és kötelezettségek problémáját próbálják megoldani. Nem meglepő módon a jogászok hajlamosak a technológiát a meglévő jogi doktrínák és a jelenlegi politikai aggályok szemszögéből vizsgálni. A jogászokat érdekelheti - és érdekelnie is kellene -, hogy a mérnökök és az informatikusok hogyan határozzák meg a fogalmaikat; de a jogászok érdeklődése abból fakad, hogy képesek ezeket az elképzeléseket jogi és politikai problémákhoz kapcsolni.

### B. A kód törvénytelen

Mind a robotok, mind a mesterséges intelligencia-ügynökök problémákat okoznak a jog számára, mivel nem mindig lehet megjósolni, hogy mit fognak tenni, amikor kölcsönhatásba lépnek a környezetükkel. Calo ezt a tulajdonságot "emergenciának" <sup>19</sup>vagy "emergens viselkedésnek" nevezi.<sup>20</sup> Egyetértek vele abban, hogy - legalábbis a történelem ezen szakaszában - a "megjelenés" hasznosabb kifejezés, mint az "autonómia", amely számos nehéz kérdést vet fel a mesterséges intelligencia státuszával kapcsolatban. Az "emergencia" problémája azonban keresztezi a fizikai ártalom problémáját. Calo számos példája - a villámkereskedés, az algoritmikus beszéd, valamint az automatizált szerzői vagy szabadalmi jogsértés - a megjelenést foglalja magában, de nem fenyeget fizikai sérüléssel.<sup>21</sup> Mindazonáltal e károk némelyike - például az értékpapírpiacok destabilizálódása - lehet

---

19. *Id.* 532.

20. *Id.* at21 .515.

*Id.* 541-43.

katasztrofálisak, még akkor is, ha nem járnak fizikai sérüléssel. Bizonyára az ilyen jellegű sérülések ugyanolyan fontosak az új technológiával való jogi találkozás szempontjából, de ezek még inkább aláássák azt a feltételezést, hogy a fizikai sérülés okozásának képessége a minket érdeklő technológiák "alapvető" jellemzője.

Ezzel szemben az emergens viselkedés gondolata sokkal "lényegesebbnek" tűnhet a robotika számára, különösen, ha a robotokat általában a mesterséges intelligencia rendszerekkel együtt tekintjük. De ennek a tulajdonságnak a fontosságát is az új technológiák és a meglévő törvényszerűségek kölcsönhatása hozza létre. Ha a jognak egyszerű vagy nyilvánvaló megoldásai lennének a robotok és az AI-rendszerek által okozott sérülésekre, akkor a jog szemében nem tartanánk különösebben megkülönböztető vagy kiemelkedő jelentőségűnek azt a tényt, hogy a robotok és az AI-rendszerek saját maguk hozzák meg döntéseiket, még akkor sem, ha ez a tény technológiai, erkölcsi vagy filozófiai okokból fontos lehet. Amint azonban Calo utal rá, a robotok és a mesterséges intelligenciával rendelkező rendszerek kialakuló viselkedése új nehézségeket okoz a meglévő jogi doktrína számára.<sup>22</sup> Ezért érdekli ez a jogászokat.

Az emergencia problémája az a probléma, hogy kit fogunk felelőssé tenni azért, amit a kód tesz. Lawrence Lessig híres diktátuma, miszerint "A kód a törvény", azt állította, hogy a számítógépes hardver és szoftver kombinációi - a szabályozás más módozataihoz hasonlóan - korlátozhatják és irányíthatják az emberi viselkedést.<sup>23</sup> Az emergencia a fordított problémát is felveti: az öntanuló rendszerek sem kiszámíthatóak, sem pedig nem korlátozhatók a megfelelő viselkedésre vonatkozó emberi elvárások által. A kód törvénytelen.

Ezt az aggodalmat két problémára oszthatjuk. Először is, ki kell osztanunk a felelősséget az interaktív és/vagy öntanuló rendszerek által okozott fizikai, gazdasági vagy érzelmi sérülésekért. Számos különböző lehetséges szereplőt vonhatunk felelősségre, beleértve a tulajdonost, az üzemeltetőt, a kiskereskedőt, a hardver tervezőjét, az operációs rendszer tervezőjét vagy a programozó(ka)t, hogy csak néhány lehetőséget említsünk.

Milyen mértékű vétkességet kell megkövetelnünk ezektől a potenciális vádlottaktól? A legegyszerűbb eset, amelyre a jog már jól felkészült, egy olyan vádlottat érint, aki egy robotot úgy programoz, hogy az kárt okozzon egy másiknak. A legtöbb esetben azonban nehéz lesz bizonyítani akár a szándékos károkozási szándékot, akár a károkozás bekövetkeztének tudatát. Ilyenkor a gondatlansági elméletekhez lehet fordulni. De bár a jövőben *valamikor bekövetkező* sérülés kockázata előre látható, ha valaki új technológiát vezet be, az, hogy hogyan és mikor következik be a sérülés, nem biztos, hogy a fenti listán szereplő potenciális alperesek mindegyike számára előre látható. Ha a jog azt reméli, hogy a felelősséget az emberekre és a vállalatokra hárítja, akkor a robot- és mesterséges intelligencia rendszerek által okozott sérülések az előreláthatóság hagyományos fogalmait feszegethetik. A vétkesség nélküli felelősség hagyományos megoldás, de ez elfojthatja az innovációt egy fejlődő területen, és nem biztos, hogy megfelelő megoldás a büntetőjog kontextusában.



22. *Id.* 543.

23. LESSIG, *Supra* note, 6,88-896,20,. o.

Az elsőből következik egy második probléma - az ok-okozati felelősség. A robotok és az AI-egységek által használt programozás és algoritmusok valószínűleg sok kéz munkája lesz. A digitális hálózatok példája tanulságos. Jonathan Zittrain generativitás koncepciója az internet kimeríthetetlennek tűnő innovációs kapacitását a lehetséges variációk több rétegével magyarázza.<sup>24</sup> Van egy hardverréteg, amelyen keresztül a digitális tartalom áramlik, egy protokollréteg, amely közös nyelvet hoz létre a cseréhez, és egymást követő, egymásra épülő szoftverrétegek.<sup>25</sup> És van egy "társadalmi réteg" is - a társadalmi tevékenységek, szokások és szervezetek, amelyek a generatív technológiák körül alakulnak ki, és érdekes és újszerű módokat találnak azok használatára és megváltoztatására.<sup>26</sup> Az innováció minden egyes rétegben előfordulhat, sok különböző ember munkáját ötvözve.

A szoftverek szintjén például a nyílt forráskódú projektek kihasználják az internet azon képességét, hogy csökkenti a szervezési költségeket. Ezek a projektek - és sok kereskedelmi projekt is - gyakran folyamatosan változnak az egymást követő frissítések és javítások révén. A szoftverprogramok platformként is szolgálhatnak a harmadik felek által készített végtelenül sokféle alkalmazás, bővítmény és módosítás számára.

A robotok és az AI-algoritmusok az innováció ezen gyakorlataira építenek; a hardver, a protokoll és a szoftver több rétegéből érthetjük meg őket, amelyek többféle lehetőséget kínálnak a több szereplő általi innovációra és változtatásra. Minél több lehetőség van az innovációra és a változtatásra, annál nehezebb lesz azonban megtalálni és kimutatni a felelősséget a másnak kárt okozó, kialakulóban lévő viselkedésért.

A szoftverek - különösen az érett és összetett szoftverek - valószínűleg hibásak vagy kiszámíthatatlan eredményeket produkálnak. A hibákat nehéz lehet észrevenni, és többszörös módosítások és kiegészítések kombinációja révén alakulhatnak ki. A több kéz által végzett szoftverfejlesztés rétegeiből eredő hibákért való felelősség megállapítása ördögien nehéz lehet. És amennyiben a robotok és a mesterséges intelligenciaprogramok megtanulják, hogyan módosítsák saját kódjukat, a felelősség kérdései még diffúzabbá válnak.

A biztonság problémája az ok-okozati felelősség problémájának a másik oldala: minél több lehetőség van az innovációra, annál több a lehetséges célpont a hackerek számára. Itt Lessig híres gondolata a kód általi szabályozásról az ellenkezőjére fordul. A kód mint az embereket szabályozó törvény helyett a kód egyfajta törvénytelenséggel rendelkezik, amely elkerüli az emberi szabályozást. Ahogy James Grimmelmann rámutatott, a szoftverek által történő szabályozás nagy részére jellemző, hogy a szoftverek feltörhetők és nem robusztusak.<sup>27</sup> Ha egy szoftverrendszert egyszer feltörnek, az meghibásodhat.

---

24. ZITTRAIN, *Supra* note at 5,67.

25. *Id.* Zittrain ezeket "fizikai" rétegeknek, "alkalmazási" rétegeknek és "tartalmi" rétegeknek nevezi. Én kissé eltérő terminológiát használok.

26. *Id.*

27. James Grimmelmann, *Regulation by Software*, YALE 114 L.J. 1741-431719, (2005).

azonnal, olyan módon, amely nem azonnal nyilvánvaló; és a rendszer lehet, hogy emberi beavatkozás nélkül nem képes magától helyreállni.<sup>28</sup>

A helyzetet tovább rontja, hogy sok robot és mesterséges intelligencia rendszer valószínűleg folyamatosan kapcsolatban lesz az internettel, és folyamatosan új információkat és új programokat vesz fel több forrásból. Az önvezető autókat például interaktív robotok óriási hálózatának részeként tervezhetnénk, amelyek folyamatosan információkat küldenének egymásnak a helyi vezetési körülményekről. Az operációs rendszer rendszeres frissítéseit a végfelhasználó tudta nélkül tölthetnék le az egyes autókra.<sup>29</sup> Valójában arra kell számítanunk, hogy a leghasznosabb és legszélesebb körben alkalmazott robot- és mesterséges intelligencia-rendszerek némelyike az internetes felhőhöz lesz csatlakoztatva. Ez azt jelenti, hogy ezek a rendszerek nem lesznek önálló egységek, hanem folyamatosan frissülnek majd más robotokkal és mesterséges intelligenciaegységekkel, valamint központi és decentralizált információforrásokkal való kommunikáció révén. A felhőalapú robotika biztonsági problémáitól eltekintve a balesetekért viselt felelősséget is megnehezíti és megosztja.

Calo elmagyarázza, hogy ezek a problémák az adatok "promiszkuitásából" következnek, ami általában a digitális hálózatokra jellemző.<sup>30</sup> A "promiszkuitás" metaforája arra utal, hogy a hálózat csomópontjai folyamatosan információt osztanak meg sok más csomóponttal (mint egy személy, aki sok különböző emberrel közösül). A "promiszkuitás" metaforája a fertőzéssel (a nemi betegségek analógiájára), a hackeléssel és a károkozással szembeni sebezhetőségre is utal. Calo az adatok promiszkuitását a megtestesülés problémájához kapcsolja - vagyis a fizikai károkozás képességéhez.<sup>31</sup> Ez azért van így, mert az adatok promiszkuitása és az átprogramozhatóság könnyűsége nem volt akkora probléma, amikor az egyetlen kár, ami ebből fakadt, nem fizikai volt.

Én másképp látnám a kérdést. Az internet összekapcsolt jellegét, a digitális generativitást és az innováció többretegűségét az emergensia problémájával hoznám összefüggésbe. Ez azért van így, mert a jog szempontjából - szemben a mérnöki szemlélettel - az emergensia által felvetett probléma a robotok és a mesterséges intelligencia rendszerek kiszámíthatatlan viselkedéséért való felelősség megállapításának problémája.

Calo hangsúlyozza, hogy ez a kiszámíthatatlanság az algoritmusok összetettségéből adódik, beleértve azokat is, amelyek tapasztalatból tanulnak.<sup>32</sup> Ez minden bizonnyal fontos oka a jogi problémának, de nem az egyetlen. Az innováció többretegűsége, a digitális rendszerek generativitása és az adatok könnyű áramlása szintén okai annak, hogy nem biztos, hogy meg tudjuk jósolni, hogy az öntanuló vagy

---

28. *Id.* 1742.

29. Vö. Jack Boeglin, *Az önvezető autók költségei: Reconciling Freedom and Privacy with Tort Liability in Autonomous Vehicle Regulation*, 17 YALE J.L. & TECH. 171, 181 (2015) (olyan "kommunikatív" automatizált járművek leírása, amelyek dinamikusan töltik le más járművek adatait).

30. Calo, *Supra* note, 1,532-33515,519., o.

31. *Id.* 533-35.

32. *Id.*

az interaktív rendszerek, ha egyszer rászabadítjuk őket a világra. Ez egy újabb ok, amiért az "emergencia" jobb kifejezés, mint az "autonómia". Egy önállóan tanuló rendszer nem attól lesz *autonómabb*, hogy rétegesen strukturált, generatív, moduláris, hibás, és sok kéz munkája. Ezek a tulajdonságok inkább azokkal az okokkal kapcsolatosak, amelyek miatt a törvény az emergenciával törődik: fokozzák annak lehetőségét, hogy az öntanuló rendszerek olyan módon oldanak meg problémákat és végeznek el feladatokat, amit mi nem látunk előre, vagy nem várunk el.

Az emergenciamodell problémája az ellenkező irányba is működik - az *ártalmak* helyett az emergens viselkedés többszörös társadalmi *hasznot* hozhat létre. Például a robotok és az AI-rendszerek új találmányokat és irodalmi műveket hoznak létre. A kérdés az, hogy ki fogja élvezni a szellemi tulajdonjogokat, vagy ehelyett az alkotások a köztulajdonba kerülnek.<sup>33</sup> A keresőalgoritmusok rangsorok formájában beszédet produkálnak; automatikusan kitölthetik vagy kiegészíthetik a keresési lekérdezéseket is. Ezt a kifejezést a keresőmotor-vállalat beszédének részeként kell-e kezelnünk, és ezért az Első Alkotmánykiegészítés alapján védenünk kell-e a szabályozással szemben?<sup>34</sup>

Valószínűleg jelentős opportunitizmust fogunk látni a jogi érvekben, amelyeket az emberek a robotokkal és a mesterséges intelligencia rendszerekkel kapcsolatban hoznak fel. Az emberek elismerést (és jogi védelmet) fognak követelni minden olyan előnyért, amelyet robotjaik létrehozhatnak, miközben tagadni fogják a felelősséget, ha robotjaik tulajdon pusztítanak el, szerzői jogokat sértenek, fenyegetőznek vagy rágalmaznak. Az emberek lekicsinylik az emergenciát, amikor a rendszer olyan előnyöket termel, amelyeket saját maguknak szerezhetnek meg, és hangsúlyozzák az emergenciát, amikor a rendszer kárt okoz. Ez egy változata a régi mondásnak, miszerint "a sikernek ezer apja van, a kudarc árva". A jog célja, hogy megfeleljen ennek a kiszámítható opportunitizmusnak, és hogy a robotok által létrehozott alkotások és a robotok által okozott károk esetében is ésszerű haszon- és felelősségmegosztást találjon ki.

### III.

#### A HELYETTESÍTÉSI HATÁS

A robotika harmadik jellemzője, amelyet Calo lényegesnek nevez, az a tendencia, hogy az emberek úgy reagálnak a robotokra, mintha emberekkel lépnének kapcsolatba. Ezt a tendenciát "társadalmi valenciának" nevezi,<sup>35</sup> és azt állítja, hogy "a történelem talán minden más technológiájánál nagyobb mértékben a robotok társadalmi jelentéssel bírnak az emberek számára".<sup>36</sup>

---

33. Annemarie Bridy, *A kreativitás kódolása*: 2012 STAN. TECH. L. REV. 5; Ralph D. Clifford, *Intellectual Property in the Era of the Creative Computer Program: Will the True Creator Please Stand Up?*, TUL71. L. REV. 1675 (1997); Pamela Samuelson, *Allocating Ownership Rights in Computer-Generated Works*, U1185. PITT. L. REV. 47 (1985).

34. Tim Wu, *Machine Speech*, 161 U. PA. L. REV. 1495 (2013); Stuart Minor Benjamin, *Algorithms and Speech*, 161 U. PA. L. REV. 1445 (2013); Eugene Volokh & Donald M. Falk, *Google: First Amendment Protection for Search Engine Search Results*, 8 J.L. ECON. & POL'Y 883 (2012); Frank Pasquale, *Search as Speech: Two Scenarios*, CONCURRING OPINIONS (2012. május 29., 9:53), <http://www.concurringopinions.com/archives/2012/05/search-as-speech-two-scenarios.html>.

35. Calo, *fenti* 1,545-46532., o.

36. *Id.* 546.

Természetesen az emberek nem feltétlenül úgy reagálnak a robotokra, mintha azok emberek lennének. Lehet, hogy úgy reagálnak rájuk, mintha állatok lennének. Gondoljunk csak egy robotkutyára, amelyet az ember betanít trükkökre, vagy egy robotfókára, amelyet idős betegek számára terveztek, hogy megسيمogassák. Amit Calo leír, az lehet antropomorfizmus (emberi tulajdonságok tulajdonítása) vagy zoomorfizmus (állati tulajdonságok tulajdonítása). Az általánosabb pont azonban az, hogy az emberi lények cselekvőképességet vetítenek ki nem élő dolgokra. Az emberek emóciókat, öröm- és fájdalomérzetet, a másokkal való kapcsolatteremtés képességét, valamint azt a képességet is kivetíthetik, hogy gondoskodjanak másokról, és hogy ők maguk is gondoskodjanak róluk. Az emberi vagy állati érzelmek kivetítése élettelen tárgyakra egyidős a történelemmel. Az emberek hallják a szél üvöltését és az óceán morajlását; ügynöki tevékenységet és hűséget vetítenek hajóikra és autóikra. Az emberség kivetítése arra, ami nem emberi, az én visszatükröződése a külvilágra.

Érdemes még egyszer megjegyezni, hogy ez a tendencia nem csak a robotikára jellemző, hanem a mesterséges intelligencia rendszerekre is. Spike Jonze 2013-as *Her*<sup>37</sup> című filmje egy férfiről szól, aki nem egy robotba, hanem egy operációs rendszerbe szeret bele. A robotokat az emberek talán azért tekintik élőnek, mert mozognak; a mesterséges intelligencia rendszerek viszont azért tekintik őket élőnek, mert beszélnek. És ha az antropomorfizmus az aggodalom tárgya, akkor az emberek a beszéd erejét sokkal inkább az emberiséghez társítják, mint a mozgás erejét.<sup>38</sup>

A robotika első két jellemzője, amelyet Calo azonosított - a fizikai károkozásra való képesség és a kialakuló viselkedés - nyilvánvaló problémákat okoz a felelősség megállapításában a kártérítési és büntetőjogban. Ezért különösen érdekesek a jogászok számára. Amit Calo "társadalmi valenciának" nevez, az azonban sokkal összetettebb jelenség. Nem korlátozódik a jogi felelősség kérdésére, hanem minden olyan módot érint, ahogyan a robotok és a mesterséges intelligencia-ügynökök beavatkozhatnak a társadalmi kapcsolatokba. Nem meglepő módon Calo a lehetséges következmények változatos sorát kínálja fel. Rámutat például arra, hogy (1) minél antropomorfabb egy robot, annál inkább a robotot fogják az emberek hibáztatni, mint a robotot használó embert;<sup>39</sup> (2) a robotok jelenléte egy megfigyelési rendszerben fokozza azt a szubjektív érzést, hogy az embert megfigyelik;<sup>40</sup> (3) az emberek nagyobb kockázatot vállalnak az antropomorf robotok integritásának megőrzése érdekében, mint az eszközöknek szánt dolgok esetében;<sup>41</sup> és (4) az emberek megkülönböztetett érzelmi sérüléseket szenvedhetnek a robottársak elvesztése miatt.<sup>42</sup> Végül, Kate Darling munkásságára építve, Calo azt javasolja, hogy az antropomorf robotok erkölcsi kárt okozhatnak: az emberek, akik

---

37. HER (Annapurna Pictures 2013).

38. *Lásd pl.* JOHN HEATH, THE TALKING GREEKS: SPEECH, ANIMALS AND THE OTHER IN HOMER, AESCHYLUS, AND PLATO 1 (2005) (azzal érvelve, hogy az ókori görögök a beszéd erejét tekintették az emberi és az állati természet közötti legfontosabb különbségnek).

39. Calo, *fenti* 1,547-48. o.

40. *Id.* 547.

41. *Id.* 548.

42. *Id.*



akik rosszul bánnak az antropomorf robotokkal, ugyanúgy elfajulhatnak vagy lealacsonyodhatnak, mint azok, akik embereket és állatokat bántalmaznak, még akkor is, ha harmadik félnek ténylegesen nem esik bántódása.<sup>43</sup>

Calo rövid értekezése csak a felszínét karcolja annak, hogy miről van szó, és a "társadalmi érték" talán nem a legjobb módja a leírásnak. Minden példája egy alapvető probléma körül forog, amelyet a robotika és a mesterséges intelligencia társadalomba való bevezetése okoz. A probléma nem az, hogy az emberek *összekeverik* a robotokat az élőlényekkel, mert általában nem így van. A probléma inkább az, hogy a robotokkal és a mesterséges intelligencia rendszerekkel való interakcióik révén az emberek hajlandóak bizonyos kontextusokban és bizonyos célokra az állatokat vagy az embereket *helyettesíteni velük*. Ezt nevezzük *helyettesítési hatásnak*. Az emberek egy entitást ember vagy állat helyettesítésére készítetnek, és úgy is kezelik az entitást - de csak bizonyos módokon. Más szóval, az emberek a robotokat és a mesterséges intelligencia-ügynököket *különleges rendeltetésű állatokként* vagy *különleges rendeltetésű emberi lényekként* kezelik. Calo jól fogalmaz, amikor "a jogi alanyok új kategóriájáról beszél, amely félúton van a személy és a tárgy között".<sup>44</sup> Az elhelyezés azért "félúton" van, mert a státusz hozzárendelése hiányos, kontextusfüggő, instabil és mindenekelőtt opportunistá lehet. Az emberek a robotot bizonyos célokból személyként (vagy állatként), más célokból pedig tárgyként kezelhetik.

Azért nevezem ezt helyettesítési hatásnak, mert sok esetben nem tekintjük a helyettesítő dolgot teljesen azonosnak azzal, amit helyettesít. Inkább csak ideiglenesen, bizonyos kontextusokban vagy bizonyos célokra egyenértékű, és az emberek gyakran fenntartják maguknak a jogot, hogy elutasítsák az állítólagos azonosságot, amikor az nekik megfelel. A helyettesítő tanár egy negyedik osztályos osztályban talán nem részesül a diákok teljes tiszteletében, és valószínűleg nem élvezi a rendes tanár nyugdíjbiztosításának előnyeit. Annak ellenére, amit a vaját, tejszínt és cukrot helyettesítő termékek gyártói mondanak nekünk, nem kezeljük ezeket a táplálékpótlókat valódi terméként. Inkább azért használjuk őket, hogy elkerüljük a valódi dologgal járó kalóriákat vagy zsírt. Vagyis az emberek a helyettesítővel próbálják mindkét világból a legjobbat kihozni - a diákokat alacsonyabb költséggel tanítják meg a törtekre; mi pedig a vaj elégedettségét kapjuk meg feleannyi kalóriával.

Az emberi érzelmek, érzések és célok kivetítése élettelen tárgyakra a helyettesítés egy speciális esete. Képzeletünk a robotot valamilyen szempontból egy emberi lény vagy állat helyére állítja. Ez a "helyettesítő" eredeti jelentése, amely a latin *substituere* szóból származik, amely a *sub+statuere* szóból származik, ami azt jelenti, hogy más alá állítani vagy más helyett állni.<sup>45</sup> A helyettesítő valami, amit felállítanak vagy alá helyeznek valaminek, hogy helyettesítse vagy kiszorítsa azt.

---

43. *Id.* 548-49. o. (idézi Kate Darling, *Extending Legal Rights to Social Robots*, in ROBOT LAW (Ryan Calo et al. szerk.) (megjelenés: 2015), *elérhető a* <http://ssrn.com/abstract=2044797> oldalon).

44. *Id.* 549.

45. *Lásd* MERRIAM WEBSTER ONLINE DICTIONARY, <http://www.merriam-webster.com/szótár/helyettesítő> (a *helyettesítő* etimológiája).

A robotnak az élőlényvel való helyettesítése lehet ártatlan, érzelmi, szinte ösztönös. Az a beteg, aki egy sebészeti robotot hibáztat egy elrontott beavatkozásért, egy részleges emberséget - és ezáltal felelősséget - vetít a technológiára. Az a katona, aki gyászolja a bomba hatástalanító robotja elvesztését, a bajtársiasság, a bátorság és a katonatársak iránti elkötelezettség emberi tulajdonságait vetíti a robotra. Amikor az otthonunkban működő robot társunk személyes adatokat küld rólunk egy vállalatnak, elárulva érezzük magunkat, holott soha nem gondolnánk, hogy egy kamera és egy mikrofon elárulhat minket.

A Calo által leírt helyettesítési formák többsége nem szándékos - egyszerűen azt tükrözhetik, ahogyan az emberek a mozgó tárgyakra és a beszélő programokra reagálnak. De ugyanilyen érdekes formái a helyettesítésnek a szándékos és az instrumentális formák. Egy kormányzat lecserélheti az emberi katonákat robotokra, mert ez utóbbiaknak nincs családjuk, és nem hullásukban térnek haza; az önkényes pusztításuk kevésbé valószínű, hogy aláássa a háború politikai támogatását. A vállalatok robotokkal helyettesíthetik a munkásokat, mert a robotok nem szerveződnek szakszervezetbe, nem igényelnek kávészünetet, és nem szenvednek alkoholproblémáktól, érzelmi problémáktól vagy a teljesítmény és a hiányzások egyéb okaitól. A helyettesítésnek ez a gyakorlata, ellentétben a fent leírt érzelmi vagy ösztönös helyettesítéssel, tudatos és céltudatos. A robotok használatára vonatkozó döntésünk olyan, mint a margarin vagy a tejmentes tejföl használata - úgy tűnik, hogy az emberek minden előnyét kínálja, de semmi költséggel.

A jog gyakorlati okokból vagy közrendi okokból szándékosan úgy is kezelheti a robotokat, mintha élő ügynökök lennének. Például megvédhetjük a robotokat a visszaélések bizonyos formáitól, mert attól tartunk, hogy a robotokkal visszaélő emberek esetleg emberekkel, gyerekekkel vagy állatokkal is visszaélnék. Lehet, hogy jogi fikcióként mens rea-t tulajdonítunk egy robotnak vagy algoritmusnak, hogy megkönnyítsük a felelős felettes jog, a bűnsegély és felbujtás büntetőjogi fogalmainak vagy a bűntársak felelősségének alkalmazását a tulajdonosokra, üzemeltetőkre vagy tervezőkre. Hasonlóképpen, a bíróságok a mesterséges intelligenciával előállított művészetet "bérmunkaként" kezelhetik, hogy minimalizálják a hatályos szerzői jog módosításait. Amikor ilyen jellegű jogi fikciókat alkalmazunk, az embert robottal helyettesítjük, hogy a jog hatékonyan működhessen a kialakuló viselkedés által felvetett jogi rejtélyekkel szemben. Vagy jogi fikciókat fogadhatunk el, hogy a meglévő jogi doktrínákat ideiglenesen működőképesen tartsuk, amíg alaposabb és koherensebb reformokat nem tudunk létrehozni.

Az imént a helyettesítés két fajtáját írtam le - a tudattalan vagy érzelmi, illetve a céltudatos vagy instrumentális helyettesítést. Valójában ezek egy érem két oldala. A helyettesítő egy másik helyett *szolgál* - mint például egy helyettesítő tanár -, vagy egy másik helyett *használják* - mint amikor egy receptben a vaját margarinnal helyettesítjük. A szolgálat és a használat fogalma egyaránt érvényes az érzelmi helyettesítésre és a gyakorlati helyettesítésre. A robotokat az érzelmeink és a gyakorlati szükségleteink szolgálatába állítjuk; arra használjuk őket, hogy teljesítsék vágyainkat és terveinket. Mint Hegel

híres dialektikájában, az úr és a rabszolga dialektikájában, függővé válhatunk attól, amit használunk, és

ami minket szolgál.<sup>46</sup> A robotok és a mesterséges intelligencia-ügynökök elvégezhetnek olyan feladatokat, amelyeket mi magunk már nem tudunk (vagy nem akarunk) elvégezni; olyan érzelmi támogatást nyújthatnak nekünk, amelyet már nem tudunk nélkülözni.

A robotok helyettesítésének egyik fő jellemzője, hogy *részleges*. A robotok és a mesterséges intelligencia entitások átveszik az emberek bizonyos aspektusait és képességeit - például a munkaképességüket, a harci és gyilkos képességüket, a nevelési képességüket, a teremtési képességüket, a vizsgáztalási, szórakoztatási vagy szexuális kielégítési képességüket. A robotok és mesterséges intelligencia-ügynökök alkalmazásával az emberi tulajdonságok és képességek szegmentálhatók, elkülöníthetők és szétszórhatók. Erre gondolok, amikor azt mondom, hogy a robotok és a mesterséges intelligencia-ügynökök "különleges célú emberi lényekként" működnek; egy bizonyos okból vagy funkcióra szolgáló ügynökök, akik az én és az eszköz, illetve a személy és az eszköz közötti határvonalon mozognak.

Első pillantásra a "helyettesítő" fogalma - még hozzá csak részleges helyettesítő - hiányt vagy alsóbbrendűséget sugallhat. A margarin nem olyan egészséges, mint a vaj; az a szerelőtanonc, aki megjelenik, kevésbé képzett, mint az, aki nem tudott eljönni a megbeszélte időpontra. A helyettesítés azonban bőséget és felsőbbrendűséget is jelenthet. A helyettesítés nemcsak a nagyobb előnyök lehetőségét ígéri alacsonyabb költségek mellett, hanem az embernél jobb készségeket és teljesítményeket is kínálhat. Ilyen esetekben a helyettesítő azt ígéri, hogy emberfeletti. A robottárs azt ígéri, hogy jobban kielégíti érzelmi szükségleteinket, a mesterséges intelligenciával működő kereskedőprogram gyorsabban és hatékonyabban végez arbitrázst, mint a földi kereskedő, a robothadsereg hatékonyabban kaszálja el az ellenséget, és így tovább. Valójában éppen a helyettesítés részrehajlása teszi felsőbbrendűvé. A robot házmester nem fárad bele a szükségleteink kielégítésébe, a robotkatoná pedig nem szenved poszttraumás stresszbetegségben, mert egyikük sem teljesen emberi.

Ezt az esszét azzal kezdtem, hogy megjegyeztem, hogy a technológia nem pusztán a személyek és a dolgok kapcsolatát érinti, hanem inkább az emberek közötti társadalmi kapcsolatokat, amelyeket a dolgok közvetítenek. Ha a szóban forgó "dolgok" az embereket helyettesítő vagy különleges célú emberek, akkor az emberek közötti társadalmi kapcsolatok közvetítése még inkább elhomályosul. A robotokkal és a mesterséges intelligencia rendszerekkel való interakcióink interakciók azokkal az emberekkel, akik ezeket az új technológiákat alkalmazzák, még akkor is, ha ezt nem vesszük észre.

#### KÖVETKEZTETÉS

Ebben az esszében a jog és a technológia találkozásának esszencializmusa ellen érveltem, és ehelyett azt javasoltam, hogy mindig tartsuk szem előtt a technológia társadalmi vonatkozásait. Mivel a technológiával együtt újítjuk meg a társadalmi kapcsolatokat is, nem mindig tudjuk megmondani, hogy mi lesz a legfontosabb a

---

46. GEORG WILLIAM FRIEDRICH HEGEL, THE PHENOMENOLOGY OF SPIRIT 117 (Arnold Vincent Miller trans., 1979) (1807) ("[J]ust ahogy az uraság megmutatta, hogy lényegi természete az ellenkezője annak, ami lenni akar, úgy a szolgaság is a beteljesedésében valóban az ellenkezőjévé válik annak, ami közvetlenül van.").

technológia az elkövetkező években. Egy tekintetben azonban a robotika lehetőségei és veszélyei világosabbnak tűnnek számunkra, mint 2015 az internet ígérete és veszélye az 1990-es évek elején. Már most úgy gondoljuk, hogy felfogjuk, miről fog "szólni" a robotika törvénye.

Az egyik ok, amiért erős elképzeléseink lehetnek arról, hogy mit tehet a robotika, és milyen problémákat okozhat, az, hogy gazdag irodalmi példákból álló irodalom áll rendelkezésünkre, amelyek a helyettesítési hatásra vonatkoznak. Valójában a helyettesítési hatás a művészettörténet egyik legrégebbi trópusa lehet. A legendás mesterember, Daidalosz például állítólag olyan élethű szobrokat alkotott, hogy meg kellett őket láncolni, hogy ne szaladjanak el.<sup>47</sup> Az emberiség történelme során az emberek mindig is meséltek olyan emberi műtárgyakról, amelyek "életre keltek", és különböző emberi funkciókat helyettesítettek. A példák Pygmalion Galatea szobrától kezdve a prágai gólemen és *Rossum univerzális robotjain át* a Frankenstein-történet végtelen variációiig, Isaac Asimov robottörténeteinek és a robotika három törvényén át egészen a televíziós műsorokból és filmekből ismert robotkarakterekig, mint például a *Lost in Space*, *Terminator*, *Star Wars* és *Star Trek: A következő generáció*.

Ezek az irodalmi és művészeti források - és még sok más forrást is megemlíthetnénk - már most sejtetni engedik, hogy mit tehetnek a robotok, és milyen problémákat okozhat a társadalomba való bevezetésük. Ne legyünk azonban túlságosan magabiztosak abban, hogy ezt a bizonyos filmet már láttuk, és ne keverjük össze képzeletünk élénkségét a bizonyossággal, hogy mit tartogat számunkra a technológia és a társadalom.

---

47. *Daedalus*, MYTHSENCYCLOPEDIA, <http://www.mythencyclopedia.com/Cr-Dr/Daedalus.html>.