



5-1-2020

A kiterjesztett vállalati elme: Amikor a vállalatok a mesterséges intelligenciát a törvény megszegésére használják

Mihailis E. Diamantis

Kövesse ezt és további műveket a következő címen:

<https://scholarship.law.unc.edu/nclr>  Part of the Law Commons

Ajánlott idézet

Mihailis E. Diamantis, *A kiterjesztett vállalati elme: When Corporations Use AI Use to Break the Law*, 98 N.C. L. REV. 893 (2020).

Elérhető a következő címen: <https://scholarship.law.unc.edu/nclr/vol98/iss4/6>

Ezt a cikket a Carolina Law Scholarship Repository ingyenesen és nyíltan hozzáférhetővé teszi. A cikket a Carolina Law Scholarship Repository egyik felhatalmazott szerkesztője fogadta el a North Carolina Law Review című folyóiratba való felvételre. További információért kérjük, forduljon a law_repository@unc.edu címre.

A KITERJESZTETT VÁLLALATI ELME: AMIKOR A VÁLLALATOK AZ MI-T ARRA HASZNÁLJÁK, HOGY MEGTÖRJÉK A LAW*

MIHAILIS E. DIAMANTIS**

Az algoritmusok hamarosan felválthatják az alkalmazottakat a vállalati visszaélések fő okaként. A jog évszázadok óta a vállalati helytelen magatartást - a polgári jogi diszkriminációtól kezdve a bennfentes kereskedelemig - az alkalmazottak helytelen magatartásaként határozza meg. Ma azonban a mesterséges intelligencia és a nagy mennyiségű adat terén elért áttörések lehetővé teszik, hogy automatizált rendszerek hozzanak számos vállalati döntést, például arról, hogy ki közzékössön vagy milyen részvényeket vásároljon. Ezek a technológiák értékes hatékonyságnövelést eredményeznek, de nem szüntetik meg (és nem is mindig csökkentik) a vállalati károkozás előfordulását. Hacsak a jog nem alkalmazkodik, a vállalatok egyre inkább mentesülnek a polgári és büntetőjogi felelősségre vonás alól, mivel a felelősséget az alkalmazottakról az algoritmusokra ruházzák át.

Ez a cikk az első, amely az algoritmikus vállalati visszaélések által hagyott, egyre növekvő doktrinális rés teljes mértékével foglalkozik. A vállalatok felelősségre vonásához a jognak néha úgy kell kezelnie őket, mintha "ismernék" a szervereiken tárolt információkat és "szándékukban állna" az automatizált rendszerek által hozott döntéseket. A kognitív tudomány és az elme filozófiája kínálja a továbblépés lehetőségét. A "kiterjesztett elme tézise" megnehezíti az elme fizikai határaitól alkotott hagyományos nézeteket. A tézis szerint az elme magában foglal minden olyan rendszert, amely kellőképpen segíti a gondolkodást, például megkönnyíti a felidézést vagy fokozza a döntéshozatalt. A természetes emberek esetében a tézis azt jelenti, hogy az elme az agyon túlmutatva olyan külső kognitív segédeszközöket is magában foglalhat, mint például a rolodexek és a számológépek. Ez a cikk a tézist a társasági jogra adaptálja és alkalmazza. Javaslatot tesz egy doktrinális keretrendszerre a vállalati elme kiterjesztésére azokra az algoritmusokra, amelyek egyre inkább szerves részét képezik a következőknek

* © 2020 Mihailis E. Diamantis.

** Az Iowai Egyetem Jogi Főiskolájának docense. A különböző szakaszokban adott felbecsülhetetlen értékű visszajelzésekért külön köszönettel tartozom Aaron Ancellnek, Shawn Bayernek, Raff Donelsonnak, Sean Griffithnek, Josh Kleinfeldnek, Alex Lemannak, Jim Lindgrennek, Jeff Lipshawnak, Kiel Brennan-Marqueznek, Micah Schwartzmannak, Lawrence Solumnak, valamint az alábbi műhelymunkák résztvevőinek: (University of Houston Law Center), az Üzleti etika a digitális korban workshop (Harvard Business School), az Üzleti jog és technológia kerekasztal (Michigan State University College of Law), a Büntügyi elméleti workshop (Harvard Law School), a Jogfilozófia kerekasztal (Southeastern Association of Law Schools), a Zicklin Center Normative Business Ethics Workshop (Wharton Business School) és a Section on Business Associations New Voices Workshop (Association of American Law Schools).

vállalati gondolat. A jognak szüksége van egy ilyen újításra, ha a jövőben a vállalatokat felelősségre akarja vonni a legsúlyosabb károkozásukért.

BEVEZETÉS	894	
I.	MINIMÁLISAN INVAZÍV MÓDSZER	901
II.	HEALTHCO ÉS FORMBOT: EGY TISZTÁZÓ PÉLDA	907
III.	A KITERJESZTETT ELME TÉZISE	912
IV.	A VÁLLALATI ELME KITERJESZTÉSE	916
	A. <i>Tanulmányi javaslat</i>	918
	1. Az analóg	megközelítés 918
	2. Általánosított	kritériumok használata 920
	B. <i>Szakpolitikai alapú kifogások és korlátozások</i>	923
	1. Felelősség a hibás	algoritmusokért 926
	2. Mások	információértvikiárius felelősség 927
KÖVETKEZTETÉS	930	

BEVEZETÉS

Marvin¹ a SciBank nevében befektetéseket eszközöl, és megpróbálja maximalizálni a hozamot. Mint minden tisztességes befektetési bankár, Marvin is csak azután vásárol vagy ad el pozíciókat, hogy módszeresen összegyűjtötte és mérlegelte a jövőbeli teljesítményre vonatkozó információkat.² Egy nap Marvin nem nyilvános információhoz jut, miszerint a BigCo ajánlatot tesz a SmallCo felvásárlására. Marvin modelljei azt jósolják, hogy a SmallCo részvényeinek árfolyama a BigCo tervének bejelentése után fel fog szökni. Következésképpen Marvin befektet a SmallCo-ba, és nagyot kaszál a SciBankkal.

Lehet, hogy a SciBank bűnös bennfentes kereskedelemben? Ha Marvin a SciBank alkalmazottja, akkor jó okunk van a közelebbi vizsgálatra. Lényeges, hogy mivel Marvin a munkaviszonya során jutott lényeges, nem nyilvános információhoz, a törvény azt diktálja, hogy a SciBank is tudomást szerzett róla.³ A SciBank felelőssége az információ eredetétől függ, azaz attól, hogy Marvin eltulajdonította-e azt⁴, vagy olyan bennfentestől kapta, aki az ügyletből hasznot húzhatott.⁵

1. Hálás vagyok Aaron Ancellnek a Harvard Egyetem Edmond J. Safra Etikai Központjában dolgozó Aaron Ancellnek ezért a nyitó gondolatért.

2. Lásd Bernard Marr, *The Revolutionary Way of Using Artificial Intelligence in Hedge Funds*, FORBES (2019. február 15., 1:48), <https://www.forbes.com/sites/bernardmarr/2019/02/15/the-revolutionary-way-of-using-artificial-intelligence-in-hedge-funds-the-case-of-aidyia/#17eb640157ca> [<https://perma.cc/46FA-3YNJ>].

3. Lásd N.Y. Cent. & Hudson River R.R. Co. v. United States, 212 U.S. 481, 491 (1909); Phila., Wilmington & Balt. R. R. Co. kontra Quigley, 62 U. S. (21 How.) 202, 209-10 (1858). Van néhány gyenge minősítés. Lásd a 173-77. és a kísérő szöveg alatti jegyzeteket.

4. Lásd Egyesült Államok kontra O'Hagan, 521 U. S. 642, 652 (1997). 5. Lásd Dirks kontra SEC, 463 U. S. 646, 659-60 (1983).

2020]

KITERJESZTETT VÁLLALATI ELME895

De tegyük fel, hogy Marvin nem is ember; tegyük fel, hogy "Marvin" a ma használatos algoritmikus kereskedési programok egyike.⁶ Bár lehet, hogy Marvin még mindig jogellenesen szerezte meg az információt (hűtlen kezelés vagy helytelen tipp révén), a bennfentes kereskedelemre vonatkozó vizsgálat azonnal megszakad. Mivel Marvin nem emberi alkalmazott, a jog nem tudja kimondani, hogy a SciBank tudott a BigCo tervezett felvásárlásáról szóló információról. Ez a bennfentes kereskedelemért való felelősségre vonást kizárja.⁷

A SciBank hipotetikus elmélete nem futurisztikus és nem is sajátos.⁸ A nagy mennyiségű adatot és mesterséges intelligenciát felhasználó fejlett algoritmusok gyorsan átalakítják a modern üzleti élet minden szegletét.⁹ A szakértők azt jósolják, hogy a vállalatok digitális automatizálásra való támaszkodása exponenciálisan fog növekedni az elkövetkező években.¹⁰ Az algoritmusok átveszik az emberi funkciókat az egész vállalati hierarchiában, a legalacsonyabb szintű műveletektől kezdve - mint például a rendszereket működtető rendszerek

6. *Lásd* Marr, 2. lábjegyzet.

7. *Lásd* 17 C.F.R. § 240.10b5-1(a) (2019).

8. Az informatikusok már évekkel ezelőtt bebizonyították, hogy az algoritmusok képesek megtanítani magukat a piacok manipulálására. *Lásd* Tom C. W. Lin, *The New Market Manipulation*, 66 EMORY L. J. 1253, 1284-86, 1292-94 (2017) (arról, hogy a mesterséges intelligencia hogyan tanulhatja meg a pump-and-dump manipulációt); Enrique Martínez- Miranda, Peter McBurney & Matthew J. Howard, *Learning Unfair Trading: A Market Manipulation Analysis from the Reinforcement Learning Perspective*, 2016 IEEE CONF. ON EVOLVING & ADAPTIVE INTELLIGENT SYSS. 103, 108; Ben Van Lier, *From High Frequency Trading to Self-Organizing Moral Machines*, 7 INTL J. TECHNOETHICS 34, 34 (2016) (megjegyezve, hogy a mesterséges intelligencia egyre önállóbbá válik a pénzügyi szektorban betöltött szerepét illetően); Michael P. Wellman & Uday Rajan, *Ethical Issues for Autonomous Trading Agents*, 27 MINDS & MACHINES 609, 614 (2017); Renato Zamagna, *The Future of Trading Belongs to Artificial Intelligence*, MEDIUM (Nov 15, 2018), <https://medium.com/datadriveninvestor/the-future-of-trading-belong-to-artificial-intelligence-a4d5887cb677> [<http://perma.cc/XSHW-VK2H>]; *lásd még a* Tanács 596/2014/EU rendelete (2014. április 16.). 2014. április 16-i 596/2014/EU tanácsi irányelv a piaci visszaélésekről (piaci visszaélésekről szóló rendelet), valamint a 2003/6/EK európai parlamenti és tanácsi irányelv és a 2003/124/EK, a 2003/125/EK és a 2004/72/EK bizottsági irányelv hatályon kívül helyezéséről, 2014 HL (L. 173) 1; 2014. április 16-i 2014/57/EU tanácsi irányelv a piaci visszaélés miatti büntetőjogi szankciókról (piaci visszaélésekről szóló irányelv), 2014 HL (L. 173) 179. o. .

9. *Lásd* Edward L. Pittman, *Quantitative Investment Models, Errors, and the Federal Securities Laws*, 13 N. Y. U. J. L. & BUS. 633, 643-44 (2017) (az algoritmusok és kvantitatív eszközök szinte általános használatáról a befektetési menedzsmentben); H. James Wilson & Paul R. Daugherty, *Collaborative Intelligence: Humans and AI Are Joining Forces*, HARV. BUS. REV. , July-Aug. 2018, at 114, 116-18 (megállapítva az AI térnyerését és hangsúlyozva az együttműködés szükségességét); Dan Wellers, Timo Elliott & Markus Noga, *8 Ways Machine Learning Is Improving Companies' Work Processes*, HARV. BUS. REV. (2017. május 31.), <https://hbr.org/2017/05/8-ways-machine-learning-is-improving-companies-work-processes> [<http://perma.cc/SP5Q-FZ9W>].

10. *Lásd* SAM RANSBOTHAM ET AL., RESHAPING BUSINESS WITH ARTIFICIAL INTELLIGENCE (A VÁLLALKOZÁS ÁTFORMÁLÁSA MŰVÉSZETI INTELLIGENCIÁVAL) 14(2017), https://www.bcg.com/Images/Reshaping%20Business%20with%20Artificial%20Intelligence_tcm9-177882.pdf [<https://perma.cc/9FDZ-L4SX>]; Wellers et al., *supra* note 9.

Az Amazon dobozcsomagoló botjai¹¹ - a legmagasabbra, mint "Vital", a Deep Knowledge Ventures igazgatótanácsába kinevezett algoritmus.¹²

Az algoritmusok azt ígérik, hogy hatékonyabbá teszik a vállalatokat¹³ és (talán) objektívebbé¹⁴,¹⁵ de nem szüntetik meg (vagy nem csökkentik)¹⁶ annak lehetőségét, hogy a dolgok néha félresiklanak.¹⁷ Az algoritmikus folyamatok gyorsasága és földrajzi kiterjedése ugyanis azt jelenti, hogy *ha* a dolgok rosszul mennek, akkor egyszerre nagyon sok ember számára, nagyon sok helyen *nagyon* rosszul mehetnek.¹⁸ A való életben

11. Az Amazon robotokat használ a vásárlók számos megrendelésének fogadására és automatikus csomagolására. Jeffrey Dastin, *exkluzív: Amazon Rolls out Machines That Pack Orders and Replace Jobs*, REUTERS (May 13, 2019, 6:08 AM), <https://www.reuters.com/article/us-amazon-com-automation-exclusive/exclusive-amazon-rolls-out-machines-that-pack-at-orders-and-replace-jobs-idUSKCN1SJ0X1> [<https://perma.cc/JY99-62T5>].

12. Vitalnak tulajdonítják, hogy a Deep Knowledge Ventures a potenciális befektetések logikusabb értékelésével elkerülte a csődöt. Nicky Burrige, *Artificial Intelligence Gets a Seat in the Boardroom*, NIKKEI ASIA REV. (2017. május 10.), <https://asia.nikkei.com/Business/Artificial-intelligence-gets-a-seat-in-the-boardroom> [<https://perma.cc/6SDR-5XGN>]. A mesterséges intelligenciáról az igazgatótanácsokban folytatott általános vitát lásd: Mark Fenwick, Wulf A. Kaal & Erik P.M. Vermeulen, *The "Unmediated" and "Tech-Driven" Corporate Governance of Today's Winning Companies*, 16 N. Y. U. J. L. & BUS. 75, 114-15 (2019) ("Emellett a mesterséges intelligencia hamarosan az igazgatósági döntéshozatal szerves részévé válhat. Elképzelhető, hogy a jövőbeli igazgatótanácsokban helyet kap egy mesterséges intelligens, szavazati joggal rendelkező igazgatótanács tag. A nem túl távoli jövőben elképzelhetőnek tűnik, hogy a mesterséges intelligencia önálló igazgatótanács ülésel rendelkezik majd, és bízni lehet benne, hogy okosabb adatvezérelt döntéseket hoz, mint az emberek."); Sergio Gramitto, *The Technology and Archaeology of Corporate Law* 33-40 (Cornell Legal Studies, Research Paper No. 18-40, 2019), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3232816 [<https://perma.cc/K4HQ-E6EP>] ("Míg a mesterséges intelligencia az igazgatótanácsokban egyszerűen arra használható, hogy segítse az emberi igazgatókat a döntéshozatalban, az is lehetséges, hogy teljesen helyettesítse az emberi igazgatókat.")

13. Lásd Sonia K. Katyal, *Private Accountability in the Age of Artificial Intelligence*, 66 UCLA L. REV. 54, 65 (2019) ("Az algoritmusok óriási értékkel bírnak. A big data jelentős előnyöket ígér a gazdaság számára, mivel lehetővé teszi a vásárlók számára, hogy gyorsabban találjanak és válogassanak termékeket, ami viszont csökkenti a keresési költségeket.")

14. Lásd CATHY O'NEIL, WEAPONS OF MATH DESTRUCTION: HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY 142-45 (2016) (az algoritmusok tényéréséről, valamint az olyan dolgok által létrehozott előnyökről és problémákról, mint a FICO és az e-scores); Paul Schwartz, *Data Processing and Governance Administration: The Failure of the American Response to the Computer*, 43 HASTINGS L.J. 1321, 1342 (1992); Kate Crawford, *The Hidden Biases in Big Data*, HARV. BUS. REV. (2013. április 1.), <https://hbr.org/2013/04/the-hidden-biases-in-big-data> [<https://perma.cc/5C9K-XYM9>].

15. Lásd Matthew Adam Bruckner, *The Promise and Perils of Algorithmic Lenders' Use of Big Data*, 93 CHI.-KENT L. REV. 3, 23 (2018); Jason Kreag, *Prosecutorial Analytics*, 94 WASH. U. L. REV. 771, 785 (2017); Chris Anderson, *Az elmélet vége: The Data Deluge Makes the Scientific Method Obsolete*, WIRED (2008. június 23., 12:00), <https://www.wired.com/2008/06/pb-theory/> [<https://perma.cc/Q5QJ-EH42>].

16. Cade Metz, *Lehetséges egyáltalán az etikus mesterséges intelligencia?*, N.Y. TIMES (2019. március 1.), <https://www.nytimes.com/2019/03/01/business/ethics-artificial-intelligence.html> [<https://perma.cc/YMG8-JPMX> (sötét archívum)].

17. Lásd Bruckner, *Supra* 15. lábjegyzet, 6. o.; Mark A. Geistfeld, *A Roadmap for Autonomous Vehicles: State Tort Liability, Automobile Insurance, and Federal Safety Regulation*, 105 CALIF. L. REV. 1611, 1614-16, 1620 (2017) ("Az autonóm járművek nem lesznek tökéletesen biztonságosak."); Mark A. Lemley & Bryan Casey, *Remedies for Robots*, 86 U. CHI. L. REV. 1311, 1313 (2019) ("Ahogy a robotika és a mesterséges intelligencia (MI) rendszerek egyre inkább beépülnek a társadalmunkba, rossz dolgokat fognak tenni.")

18. Lásd EXEC. OFFICE OF THE PRESIDENT, BIG DATA: SEIZING OPPORTUNITIES, PRESERVING VALUES, at iii (2014), <https://obamawhitehouse.archives.gov/sites/default>.

2020]

KITERJESZTETT VÁLLALATI ELME897

A vállalati algoritmikus károkozás példái, amelyek megérdemlik a felelősség vizsgálatát, a következők:¹⁹ :

- Egy hitelező automatizált platformja olyan módon hagyja jóvá a jelzáloghiteleket, amely diszkriminatív faji hatással jár, de üzleti szempontból is indokolt lehet.²⁰
- Konkurens kiskereskedők árazási algoritmusai az árakat egyező, szuperversenyképes szinten határozzák meg.²¹
- Egy futárcég önvezető teherautója elütött egy szabálytalanul közlekedő gyalogost.²²

Nem is olyan régen a vállalatok még emberi alkalmazottakra támaszkodtak a funkciók mindegyikének ellátásában. Ma már sok vállalat algoritmusokat használ a hitelek jóváhagyására, az árak meghatározására és az áruk szállítására.²³

Az automatizálás felé való elmozdulás nem változtat azon a tényen, hogy a diszkrimináció, az árrögztetés és a gondatlan vezetés áldozatokat hagy maga után.²⁴ Ezek az áldozatok,

/files/docs/big_data_privacy_report_may_1_2014.pdf [https://perma.cc/QWG5-USYP] (az algoritmusok azon lehetőségének leírása, hogy aláássák a régóta fennálló polgári jogi védelmeket).

19. Egy növekvő tudományos szakirodalom másokat is tárgyal. *Lásd pl.* Katyal, *supra* note 13, 56. o.

20. *Lásd általában* Robin Nunn, *Discrimination and Algorithms in Financial Services: Unintended Consequences of AI*, LEXOLOGY (2018. március 6.), <https://www.lexology.com/library/detail.aspx?g=ecdd186d-29eb-4821-8161-89b93cf1ef31> [https://perma.cc/9SHT-3Q6V] (sötét archívum)] (a mesterséges intelligencia faji szempontból érzékeny végrehajtásának fontosságát tárgyalja). Egy hasonló, a felvételi hirdetésekkel kapcsolatos példát lásd Esha Bhandari & Rachel Goodman, *ACLU Challenges Computer Crimes Law That Is Thwarting Research on Discrimination Online*, ACLU (2016. június 29., 10:00 AM), <https://www.aclu.org/blog/racial-justice/race-and-economic-justice/aclu-challenges-computer-crimes-law-thwarting-research> [https://perma.cc/838L-YZEE]. *Lásd még* Mikella Hurley & Julius Adebayo, *Credit Scoring in the Era of Big Data*, 18 YALE J.L. & TECH. 148, 194 (2016) (az ilyen esetek kihívásait tárgyalva).

21. *Lásd* Emilio Calvano et al., *Artificial Intelligence, Algorithmic Pricing, and Collusion*, VOX (Feb. 3, 2019), <https://voxeu.org/article/artificial-intelligence-algorithmic-pricing-and-collusion> [https://perma.cc/6WPN-SHXF]; Greg Rosalsky, *When Computers Collude*, NPR (Apr. 2, 2019, 7:30 AM), <https://www.npr.org/sections/money/2019/04/02/708876202/when-computers-collude> [https://perma.cc/ZE2T-6UUE]; *lásd még* Maurice E. Stucke & Ariel Ezrachi, *Two Artificial Neural Networks Meet in an Online Hub and Change the Future (of Competition, Market Dynamics and Society)* 52 (U. Tenn. Legal Stud., Research Paper No. 323, July 2017), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2949434 [https://perma.cc/6RNH-YSRN] (az "algoritmikus ároptimalizálás" tárgyalása).

22. Daisuke Wakabayashi, *Self-Driving Uber Car Kills Pedestrian in Arizona, Where Robots Roam*, N.Y. TIMES (2018. március 19.), <https://www.nytimes.com/2018/03/19/technology/uber-driverless-fatality.html> [https://perma.cc/MWD5-D8FK] (sötét archívum)].

23. *Lásd általában* Ellen Ruppel Shell, *AI and Automation Will Replace Most Human Workers Because They Don't Have To Be Perfect-Just Better Than You*, NEWSWEEK (2018. november 20.), <https://www.newsweek.com/2018/11/30/ai-and-automation-will-replace-most-human-workers-because-they-dont-have-be-1225552.html> [https://perma.cc/BG83-TAT7] (leírja, hogy a mega-tech vállalatok túlnyomórészt automatizálják a munkaerőt).

24. Az algoritmikus diszkrimináció ártalmainak részletes tárgyalását lásd O'NEIL, *supra* note 14, 13. o. ("[G]oing to go to college, borrowing money, getting sentenced to prison, or finding and holding a job. Mindezen életterületeket egyre inkább titkos [big data algoritmusok] ellenőrzik, amelyek önkényes büntetéseket alkalmaznak.").

vagy az államnak az ő nevükben, ugyanolyan egyértelmű utat kell biztosítani az igazságszolgáltatáshoz, mint egy évtizeddel ezelőtt. Az algoritmikus visszaélések esetében különösen fontos, hogy az út nyitva tartsa a vállalati felelősség lehetőségét. Mivel a vállalatok algoritmusokkal helyettesítik az alkalmazottakat, a vállalati felelősség lesz a jogorvoslat *egyetlen* eszköze. Az alkalmazottak saját maguk felelnek a munkahelyi vagy munkahelyen kívüli helytelen magatartásukért.²⁵ Az algoritmusok azonban nem perelhetők.²⁶

A probléma az, hogy a jog nincs felkészülve a vállalati felelősség kezelésére, amikor a vállalati visszaélések mögött álló "gondolkodás" automatizált rendszerekre hárul.²⁷ A jelenlegi jog szerint a vállalati felelősség a fenti esetek mindegyikében megkövetelné (és ahogy alább feltételezem, meg is kellene követelnie) a²⁸ bizonyítékát a vétkes vállalati mentális állapotnak: szándék (a megkülönböztetés),²⁹ ismerete (a versenytársak áráiról),³⁰ vagy gondatlanság (a jármű üzemeltetése során).³¹ A társaságoknak szellemi állapotokat tulajdonító jogi doktrína - a "correspondeat superior" - a társasági szellemi állapotokat az alkalmazottak szellemi állapotai szempontjából határozza meg.³² A respondeat superior változatai - mint például a kollektív tudás doktrína, amely lehetővé teszi a bíróságok számára, hogy a munkavállalók tudását összevonják,³³ és a kontrollcsoport teszt, amely a respondeat superior-t a magasabb rangú vállalati alkalmazottakra korlátozza³⁴ - csak

25. *Lásd általában* V.S. Khanna, *Vállalati büntetőjogi felelősség: Milyen célt szolgál?*, 109 HARV. L. REV. 1477, 1489-90 (1996) (a "respondeat superior" doktrínát tárgyalja, amely szerint egy vállalat is felelőssé tehető egy alkalmazott cselekedeteiért).

26. Egyesült Államok kontra Athlone Indus. Inc., 746 F. 2d 977, 979 (3d Cir. 1984) (megállapítva, hogy [r]obotok nem perelhetők); Ugo Pagallo, *Killers, Fridges, and Slaves: A Legal Journey in Robotics*, 26 AI & SOC'Y 347, 349 (2011) ('Az általános jogi álláspont kizárja a robotokat mindenfajta büntetőjogi felelősségből').

27. Ez párhuzamba állítható egy másik filozófiai problémával: hogyan lehet a digitálisan tárolt adatoknak "jelentése"? *Lásd általában* Lawrence B. Solum, *Artificial Meaning*, 89 WASH. L. REV. 69, 69-70 (2014) (a "mesterséges jelentés fogalmát tárgyalja, azaz az egyes természetes személyektől eltérő entitások által létrehozott jelentéseket").

28. *Lásd az alábbi* I. részt.

29. Diszkriminációs ügyeket célzatosság hiányában is be lehet indítani, ha az ügyek egyenlőtlen hatást állítanak. *Lásd pl.*, *CFPB and DOJ Order Ally To Pay \$80 Million to Consumers Harmed by Discriminatory Auto Loan Pricing*, CFPB NEWSROOM (2013. december 20.), <https://www.consumerfinance.gov/about-us/newsroom/cfpb-and-doj-order-ally-to-pay-80-million-to-consumers-harmed-by-discriminatory-auto-loan-pricing/> [<https://perma.cc/KC6K-3BKE>]. Amennyiben az eltérő bánásmódnak van lehetséges üzleti indoka (mint ahogyan az algoritmikus megkülönböztetés esetében általában lesz), a célzat bizonyításának hiányában az ügy sokkal nehezebbé válik. Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671, 711-12, 726 (2016); Stephanie Bornstein, *Antidiskriminatory Algorithms*, 70 ALA. L. REV. 519, 535 (2018).

30. *Lásd a* Sherman Antitrust Act, ch. 647, 26 Stat. 209 (1890) (kodifikálva a 15 U.S.C. §§ 1-7 (2018)); *United States v. Wise*, 370 U. S. 405, 416 (1962) ('[A] corporate officer is subject to prosecution under § 1 of the Sherman Act whenever he knowingly participates in effecting the illegal contract, combination, or conspiracy').

31. *Lásd pl.*, IOWA CODE ANN. § 321.277 (Westlaw a 2019. évi törvényhozásig) (a 'gondatlan vezetés' bűncselekményének meghatározása).

32. RESTATEMENT (THIRD) OF AGENCY § 2.04 cmt. b (AM. LAW INST. 2006). Ezeket a gyenge korlátokat az alábbiakban tárgyalom. *Lásd az alábbi* IV. részt.

33. *Lásd* *United States v. Bank of New England, N.A.*, 821 F.2d 844, 856 (1st Cir. 1987).

34. *Lásd* MODEL PENAL CODE § 2.07(1)(c) (AM. LAW INST. 2018).

2020]

KITERJESZTETT VÁLLALATI ELME899

megerősítik azt a jelenlegi jogi tény, hogy a *vállalati* mentális állapotoknak a *munkavállalók* mentális állapotaiból kell származniuk. Amikor a vállalatok az alkalmazottaik révén rosszul viselkednek, a felelősségre vonás viszonylag egyszerű felelősségmegállapítást eredményez.³⁵ Amikor azonban a vállalatok az algoritmusaikon keresztül olyan módon viselkednek rosszul, amely kívülről ugyanolyan céltudatosnak, tudatosnak vagy gondatlannak tűnik, mint az emberi alkalmazottak által elkövetett rossz magatartás, a jelenlegi felelősségi doktrínák nem alkalmazhatók.³⁶

Egy olyan kereskedelmi világban, amelyet egyre inkább szilíciummal működtetnek, meglepő, hogy a jogban a vállalati elme felfogása még mindig egy őskori szürke szerves anyaghoz kötődik. Egy olyan vállalat, mint a JPMorgan, olyan szerződéses rendszerekkel rendelkezik, amelyek szó szerint meghaladják - egyes számítások szerint - 390 000 emberi agy tárolókapacitását.³⁷ A processzorai ezeket az információkat egyes becslések szerint 10 000 000-szer gyorsabban elemzik, mint ahogyan egy ember képes lenne rá.³⁸ Ha az információ és az abból levont következtetések nem egy emberi alkalmazott agyán haladnak keresztül, akkor nem képezik részét a jogi vállalati elméről alkotott jelenlegi felfogásának.

A törvény jelenlegi állapota azért aggasztó, mert szinte garantálja, hogy a vállalatok egyre kevésbé lesznek felelősek, mivel működésük egyre kevesebb emberi beavatkozást igényel.³⁹ Az ügyészek például felelősségre vonási elmélet hiányában nem emeltek vádat az Uber ellen, amikor az egyik önvezető autója elütött és megölt egy gyalogost Arizonában.⁴⁰ A felelősségre vonás által hagyott jogi kiskapu a vállalati felelősségvállalás egy kellemetlen formáját ösztönzi.

35. Ezeknek az egyszerű meghatározásoknak az eredményei nem túl meggyőzőek. *Lásd* Mihailis E. Diamantis, *Corporate Criminal Minds*, 91 NOTRE DAME L. REV. 2049, 2056-58 (2016) [a továbbiakban Diamantis, *Corporate Criminal*] (a respondeat superior kritikája).

36. Jack M. Balkin, *The Three Laws of Robotics in the Age of Big Data*, 78 OHIO ST. L.J. 1217, 1234 (2017).

37. *Lásd* Dakin Campbell, *Meet the JPMorgan Banker with No Technical Expertise Who's Now in Charge of the Biggest Data Projects on Wall Street*, BUS. INSIDER (2019. június 12.), <https://www.businessinsider.com/rob-casper-jpmorgan-data-head-profile-managing-financial-account-info-2019-5> [<https://perma.cc/6BDR-W9SA>]; Forrest Wickman, *Your Brain's Technical Specs*, SLATE (2012. április 24., 20:18), <https://slate.com/technology/2012/04/north-koreas-2-mb-of-knowledge-taunt-how-many-megabytes-does-the-human-brain-hold.html> [<https://perma.cc/6SFR-AV4D>].

38. Liqun Luo, *Miért olyan hatékony az emberi agy?*, NAUTILUS (2018. április 12.), <http://nautil.us/issue/59/connections/why-is-the-human-brain-so-efficient> [<https://perma.cc/S2LD-44UA>].

39. Egy sor kapcsolódó kérdés merül fel, amikor a kormányok algoritmusokat használnak különböző funkciók ellátására, például a társadalombiztosítási ellátások iránti kérelmek elbírálására és a feltételes szabaddábra helyezésről szóló döntések meghozatalára. *Lásd általában* Ronald Bailey, *Welcome Our New Algorithmic Overlords?*, REASON (2016. július 8., 13:30), <https://reason.com/2016/07/08/welcoming-our-new-algorithmic-overlords-2> [<https://perma.cc/4DPP-SMU6>] (az algoritmusoknak a kormányzati és bírósági döntéshozatalban való alkalmazásáról).

40. Angie Schmitt, *Uber Got off the Hook for Killing a Pedestrian with Its Self-Driving Car*, STREETS BLOG USA (2019. március 8.), <https://usa.streetsblog.org/2019/03/08/uber-got-off-the-hook-for-killing-a-pedestrian-with-its-self-driving-car/> [<https://perma.cc/3V4D-K2UV>] (spekulálva, hogy a technológiai cégeket nem fogják megbüntetni azért, hogy a legsúlyosabb kockázatokat vállalják a még nem tesztelt technológiájukkal még akkor sem, ha a legrosszabb történik.).

játszadózás. A felelősségüket kezelni kívánó vállalatok⁴¹ inkább az algoritmikus szabálysértések biztonságos menedékét fogják keresni, mintsem az emberi alkalmazottak által elkövetett szabálysértésekért való felelősségre vonást. Azzal, hogy a törvény nem kínál megoldást, arra ösztönzi a vállalatokat, hogy felgyorsítsák az automatizálás alkalmazását. Ez a dinamika fokozza annak kockázatát, hogy a vállalatok idő előtt fordulnak az algoritmusok felé, mielőtt a technológiát kellőképpen tesztelték volna a társadalmilag felelős használat szempontjából.⁴² A jognak olyan keretre van szüksége, amely kiterjeszti a vállalati elme fogalmát azon alkalmazottakon túlra, akiknek a helyét az algoritmusok fogják betölteni. A jog csak ezután dolgozhatna ki megbízható doktrínákat annak értékelésére, hogy az algoritmusok segítségével helytelenül viselkedő vállalatok mégis megfelelnek-e a felelősség mentális állapotának. A pszichológusok és filozófusok a közelmúltban az emberi elmével kapcsolatos kapcsolódó kérdésekkel foglalkoztak. Azzal érvelnek, hogy az emberi elme hagyományos értelmezése, amely szerint az emberi elmét a koponya határai korlátozzák, túlságosan korlátozó.⁴³ Az úgynevezett "kiterjesztett elme tézis" azt állítja, hogy az emberi elme túlmutat az agyon, és magában foglalja a külső kognitív segédeszközöket - például a naplót vagy a mobiltelefont -, amelyek segítik az agy munkáját.⁴⁴ Ha egy személy ugyanolyan könnyen "fel tud idézni" egy telefonszámot a telefonja memóriabankjának ellenőrzésével, mint az idegrendszeri memóriabankjának ellenőrzésével, akkor elméje a tézis szerint kiterjedhet a telefonja szempontjaira is.⁴⁵ Ez a cikk a kiterjesztett elme tézisének megfelelő módosításokkal a vállalati kontextusra adaptálja. Azzal érvel, hogy a jog elismerheti és el is kellene ismernie, hogy a vállalati elme kiterjed az algoritmusokra, amelyek olyan szerepeket töltenek be, amelyeket korábban csak emberi alkalmazottak töltek be. A vállalati elme ily módon történő kiterjesztésével a jog a vállalati elszámoltathatóságot a huszonegyedik században.

Az alábbiakban kidolgozott javaslat a bírák és/vagy ügyészek számára készült. A társaságokra alkalmazandó felelősségi keret nagyrészt a következőkből származik

41. A vállalatok feltételezhetően ezt akarják. *Lásd* Cindy R. Alexander & Mark A. Cohen, *A vállalati bűnözés okai: An Economic Perspective*, in *PROSECUTORS IN THE BOARDROOM: USING CRIMINAL LAW TO REGULATE CORPORATE CONDUCT* 11, 14-15, 17 (Anthony S. Barkow & Rachel E. Barkow szerk., 2011).

42. Brad Smith, a Microsoft elnöke és jogi igazgatója megjegyezte, hogy "nem akarjuk, hogy kereskedelmi versenyfutás alakuljon ki" és kijelentette, hogy "[j]aw-ra van szükség." Metz, 16. lábjegyzet; *lásd még* Frank Pasquale, *Toward a Fourth Law of Robotics: Preserving Attribution, Responsibility, and Explainability in an Algorithmic Society*, 78 OHIO ST. L. J. 1243, 1244-45 (2017) (ismertetve a kísérletét arra, hogy megtiltja a jogi szoftverek gyártóinak, hogy az ügyvédeknek alacsonyabb felelősségi normát érjenek el, valamint a Szövetségi Kereskedelmi Bizottság és az Igazságügyi Minisztérium kritikáját az észak-karolinai döntéssel kapcsolatban).

43. *Lásd* Andy Clark & David Chalmers, *The Extended Mind*, 58 ANALYSIS 7, 8-10 (1998) (a kiterjesztett elme tézisének bemutatása, amelyet a szerzők "aktív externalizmusnak" neveznek).

44. *Lásd* Marc Jonathan Blitz, *Gondolatszabadság a kiterjesztett elmének: Cognitive Enhancement and the Constitution*, 2010 WIS. L. REV. 1049, 1055-56 (Clark és Chalmers elméletét "kiterjesztett elme" néven írja le); Clark & Chalmers, *Supra* note 43, 12-14. o. (egy notebookkal kapcsolatos példát tárgyalva).

45. *Lásd* Clark & Chalmers, *Supra* 43. lábjegyzet, 12-14. pont. Amint azt alább kifejezem, a kiterjesztett elme tézisének vannak bizonyos korlátai. *Lásd alább* a IV.B. szakaszt.

2020]

KITERJESZTETT VÁLLALATI ELME901

common law, a⁴⁶ bevezetésre került a társasági jogba, majd bírósági tevékenység révén kibővült.⁴⁷ Mint ilyenek, a bírói határozatok jelentik a létező doktrína fejlődésének legvalószínűbb pontját.⁴⁸ A büntetőjogban az ügyészeknek is fontos szerepük van a társasági felelősség kapuőreként. A felelősség elméletének hiányában az ügyészek - mint az ügyes gondatlan vezetése esetében - elutasíthatják a vádemelést. Alternatív megoldásként az ügyészek bíróságon kívüli megállapodások kikényszerítésével maguk is foglalkozhatnak a büntetőjogi felelősségre vonás nyilvánvaló szükségességével az algoritmikus vállalati visszaélések esetében.⁴⁹ Elvi megközelítés hiányában az ügyészek valószínűleg súlyosbítják a jogállamiságot⁵⁰ és az elszámoltathatóságot⁵¹ aggodalmakat, amelyeket mások is felvetettek azzal kapcsolatban, hogy az ügyészek hogyan oldják meg a vállalati bűncselekmények gyanús eseteit. Az alábbiakban javasolt megoldás némi útmutatást nyújthat.

A cikk a módszertan fontos szempontjainak tisztázásával kezdődik (I. rész), és egy hipotetikus forgatókönyvet kínál (II. rész). Legfontosabb érdemi hozzájárulásként ez a cikk bemutatja a kiterjesztett elme tézisének (III. rész), és részletesen bemutatja, hogyan lehet ezt a koncepciót doktrínaként adaptálni az algoritmikus vállalati visszaélések kezelésére (IV. rész). Végül a cikk a kiterjesztett vállalati elme tágabb értelemben vett következményeinek mérlegelésével zárul.

I. MINIMÁLISAN INVAZÍV MÓDSZER

A jogi problémák megoldására két tipikus eszköz létezik: a kalapács és a szike. A kalapácsok a jog alapvető szerkezeti hibáinak kezelésére szolgálnak, amikor az egyetlen út a teljes reform. Alapvető funkciójuk a bontás és az újjáépítés. E cikk szerényebb célja a szike használata.

46. *Lásd* Packard Motor Car Co. v. NLRB, 330 U. S. 485, 489 (1947) (a 'common law ősi maximájára, a *respondeat superiorra*' hivatkozva, amely "[e]ven külön törvényi rendelkezés nélkül[,] . . . számos kapcsolatra vonatkozna).

47. *Lásd* Samuel W. Buell, *The Blaming Function of Entity Criminal Liability*, 81 IND. L. J. 473, 474-75 (2006) ('Az e területre vonatkozó jog csaknem egy évszázaddal ezelőtt gyengén indult, amikor a common law bíróságok - az üzleti vállalkozások szabályozására rendelkezésre álló eszközök bővítésére törekedve - a deliktuális jogból a büntetőjogba importálták a felróhatósági felelősséget, de komoly elméleti elemzés nélkül.').

48. Az egyetlen komoly reformtörekvés a bírák részéről érkezett. *Lásd pl.* United States v. Bank of New England, N. A., 821 F.2d 844, 855 (1st Cir. 1987) (jóváhagyta az esküdtszék utasítását, amely szerint az esküdtszék az összes alkalmazott 'kollektív tudását' a munkáltatónak tulajdonította a felróhatóság kérdésében).

49. *Lásd* Brandon L. Garrett, *Collaborative Organizational Prosecution*, in PROSECUTORS IN THE BOARDROOM: USING CRIMINAL LAW TO REGULATE CORPORATE CONDUCT, *supra* note 41, 154, 157 ("[P]rosecutors typically deferred prosecution or agree not to prosecute until not at prosecute if a firm will enter into a agreement.').

50. *Lásd általában* Jennifer Arlen, *Prosecuting Beyond the Rule of Law*: 8 J. LEGAL ANALYSIS 191 (2016) (azzal érvelve, hogy az ügyészek által a halasztott vádemelési megállapodások ('DPA') és a nem vádemelési megállapodások ('NPA') révén előírt feltételek közötti nagy eltérés sérti a jogállamiságot).

51. *Lásd* Lisa Kern Griffin, *Inside-Out Enforcement*, in PROSECUTORS IN THE BOARDROOM: USING CRIMINAL LAW TO REGULATE CORPORATE CONDUCT, *supra* note 41, 110, 110. ('A DPA-k kevésbé láthatóak, mint az ítékezés, ami rontja mind a kormány végrehajtási stratégiájának koherenciáját, mind az ügyészek elszámoltathatóságát.').

Célja, hogy a lehető legkisebb doktrinális bemetszéssel oldja meg az algoritmikus vállalati visszaélések problémáját. Amit a sebészeti beavatkozás a grandiózus vízió tekintetében feláldozhat, azt a megvalósíthatóság tekintetében pótolja. Ezt úgy éri el, hogy a meglévő kereteket és doktrínákat - ebben az esetben a vállalati felelősséget - felhasználja az algoritmikus helytelen magatartás problémájának kezelésére. A kis változtatások nagyobb valószínűséggel kapnak valós vonzerőt, mivel a törvényhozók számára általában jobban elfogadhatóak, mint a nagy változtatások.⁵² Ez a tulajdonság a jelen javaslat megkülönböztető előnye.

Ez a cikk a vállalatok büntetőjogi és polgári jogi felelősségéről szól az algoritmikus visszaélésekért. Sebészeti törekvéseihez hűen e rész további része néhány kapcsolódó, de végső soron érintőleges kérdést emel ki. Az e kérdésekre vonatkozó joganyagot háborítatlanul kívánom hagyni. Ezáltal remélem, hogy megoldom az algoritmikus vállalati kötelességszegés problémáját, miközben kikerülök néhány tágabb problémát, amelyek más teoretikusokat is érdekeltek.

Egyes tudósok úgy vélik, hogy a vállalati bűnösség olyan értelmetlen fogalom⁵³, hogy a jog jobb lenne nélküle.⁵⁴ Lehet, hogy igazuk van. Végül is a vállalatok csak fiktív szereplők.⁵⁵ Talán a legjobb megoldás az algoritmikus vállalati helytelen magatartásra az lenne, ha a vállalati bűnösség iránti elkötelezettségét kritikusan kezelnénk.

Sebészeti céljainak megfelelően ez a cikk nem kérdőjelezi meg, hogy a vállalatok lehetnek-e (vagy kell-e) jogilag elszámoltathatók. Ehelyett arra törekszik, hogy a jognak ezt a szegletét a lehető legnagyobb mértékben érintetlenül hagyja. Amint máshol már kifejtettem, a vállalatokról mint felelős szereplőkről alkotott általános felfogása pszichológiailag fenntartható⁵⁶ és gyakran politikailag is ésszerű.⁵⁷ Még inkább a

52. Lásd GUIDO CALABRESI, KÖZÖS JOG A SZABÁLYOK KORÁBAN 3-4 (1982); Boris I. Bittker, *Az alkotmány értelmezése: Az alkotmányozók szándéka az irányadó? Ha nem, akkor mi?*, 19 HARV. J. L. & PUB. POLY 9, 51-52 (1995); Cynthia R. Farina, *Faith, Hope, and Rationality or Public Choice and the Perils of Occam's Razor*, 28 FLA. ST. U. L. REV. 109, 110-11 (2000) (megjegyezve, hogy "a közválasztás 'vonzó a maga takarékoságában'", és hogy "[e]nnoha meglenne a képességünk arra, hogy felszámoljuk a teljes nemzeti szabályozási apparátust, erre sem az akaratunk, sem a vágyunk nincs" (idézi Steven P. Corley, *Public Interested Regulation*, 28 FLA. ST. U. L. REV. 7, 15 (2000))); Saul Levmore, *Interest Groups and the Problem with Incrementalism*, 158 U. PA. L. REV. 815, 816-17 (2010) ("Vezető kommentátorok ösztönzik a fokozatosságot. A legtöbb bátorítás a bírának szól, de a fokozatosság mellett felhozott érvek a szabályozókra és a jogalkotókra is egyaránt alkalmazhatók." (lábjegyzet kihagyva)).

53. Lásd pl. Amy J. Sepinwall, *Guilty by Proxy: Expanding the Boundaries of Responsibility in the Face of Corporate Crime*, 63 HASTINGS L.J. 411, 428 (2012) (azzal érvelve, hogy a vállalatok nem rendelkezhetnek erkölcsi cselekvőképességgel, mivel nem képesek erkölcsi érzelmekre).

54. Lásd pl. John Hasnas, *Egy tévedés százéves évfordulója: One Hundred Years of Corporate Criminal Liability*, 46 AM. CRIM. L. REV. 1329, 1329 (2009).

55. Lásd *Sierra Club v. Morton*, 405 U. S. 727, 742-43 (1972) (Douglas, J., más véleményen) ("A közönséges vállalat az ítélkezési eljárások alkalmazásában 'személy'").

56. Lásd Gerhard O.W. Mueller, *Mens Rea and the Corporation: A Study of the Model Penal Code Position on Corporate Criminal Liability*, 19 U. PITT. L. REV. 21, 40-41 (1957).

57. Lásd Diamantis, *Corporate Criminal*, *Supra* note 35, 2052-53. o. (kifejtve, hogy miért van értelme a mens rea alkalmazásának a vállalatokra).

2020]

KITERJESZTETT VÁLLALATI ELME903

pont, a vállalati bűnösség eltörlése gyakorlatilag nem járható út. A közvélemény széles körben támogatja a vállalatok polgári és büntetőjogi felelősségre vonását.⁵⁸ Következésképpen a vállalati bűnösségre vonatkozó törvény társadalmilag beágyazott és politikailag golyóálló. Ráadásul a vállalati bűnösség köré épített jogi struktúra évszázados⁵⁹ és egyre növekszik.⁶⁰ Szétszedése hatalmas jogtudományi vállalkozást igényelne.

A vállalati bűnösség jogi keretének egyetlen jellemzője sem olyan szerves része, mint a vállalati személyiség fikciója.⁶¹ E társadalmi-jogi konstrukció szerint a vállalatok ugyanúgy illeszkednek a jog felelősségi mechanizmusába, mint más "emberek".⁶² Minden olyan jogszabály, amely polgári vagy büntetőjogi jogsértést határoz meg, egyidejűleg meghatározza a felelősség elemeit mind a magánszemélyek, mind a társaságok számára.⁶³ Bármilyen cselekményt követel meg egy jogszabály, azt valós és fiktív személyektől egyaránt megköveteli.⁶⁴ Ahol - mint a leggyakrabban - a⁶⁵ felelősség a személy egyidejűleg fennálló mentális állapotától is függ, ott mind az egyéni, mind a vállalati

58. *Lásd* Miriam H. Baer, *Choosing Punishment*, 92 B.U. L. REV. 577, 612 (2012).

59. A doktrína a tizenkilencedik század közepén került be az amerikai polgári jogba. *Lásd* Phila. & Reading R. R. Co. v. Derby, 55 U. S. (14 How.) 468, 486-87 (1852) (a vasúttársaság felelősségének megállapítása az alkalmazottja gondatlanságáért a respondeat superior doktrína révén). Ötven évvel később az amerikai büntetőjogba is bekerült. *Lásd* N.Y. Cent. & Hudson River R.R. Co. v. United States, 212 U.S. 481, 494-96 (1909) (megállapítva, hogy egy társaság bűnösnek nyilvánítása bűncselekmény elkövetésében nem sérti az alkotmányt).

60. *Lásd* Sara Sun Beale, *A Response to the Critics of Corporate Criminal Liability*, 46 AM. CRIM. L. REV. 1481, 1482 (2009).

61. *Lásd* David M. Uhlmann, *Az inga kileng: A vállalati büntetőeljárás újragondolása*, 49 U.C. DAVIS L. REV. 1235, 1246 (2016) (elismerve, hogy a társasági büntetőeljárás a társaságok törvény szerinti személyiségének jogi fikcióján alapul); *lásd még* *Corporation*, BLACKS LAW DICTIONARY (11th ed. 2019) (a társaság definíciója szerint a társaság 'olyan szervezet, amely a törvény szerint jogosult egyetlen személyként cselekedni').

62. *Lásd pl.*, Dictionary Act of 1871, ch. 71, § 2, 16 Stat. 431, 431 (jelenlegi változat: 1 U.S.C. § 1 (2018)). (a 'személy' fogalmának meghatározása a 'társaság' fogalmára); Uhlmann, *supra* note 61, at 1246. Fontos különbség van a "személyiség" - amely a jogokkal és felelősséggel rendelkező entitásokra vonatkozik - és a "személyiség" - amely az elmével rendelkező személyekre vonatkozik - között. Míg az emberek személyek, ez fordítva nem feltétlenül igaz, kivéve, ha a személyiséggel kapcsolatos felelősség olyan jogi felelősség, amely mentális állapotokat igényel. Tekintettel arra, hogy e cikk éppen az ilyen jellegű jogi felelősségre összpontosít, a 'személyek' és a "személyek" kifejezést felváltva használom a jogokkal, felelősséggel és elmével rendelkező entitásokra való utalásra.

63. *Lásd* 1 U.S.C. § 1 (2018).

64. Például, amikor egy törvény meghatározza azt a bűncselekményt, amely a következőkre vonatkozik: "[a]ki ... korrupt módon rávesz egy másik személyt ... arra, hogy ... megsemmisítsen, megcsonkítson vagy elrejtse egy tárgyat azzal a szándékkal, hogy a tárgy sérthetlenségét vagy hivatalos eljárásban való felhasználhatóságát rontsa", 18 U. S. C. § 1512(b)(2) (2018), nem kérdés, hogy 'bárki' magában foglalja a vállalatokat is, *lásd* United States v. Arthur Andersen, LLP, 374 F. 3d 281, 284 (5th Cir. 2004) (jóváhagyta a ~~hivatalos~~ igazságszolgáltatás akadályozásáért való elítélését), *más alapon felülvizsgálva*, 544 U.S. 696 (2005).

65. *Lásd*: RESTATEMENT (THIRD) OF TORTS: LIAB. FOR PHYSICAL & EMOTIONAL HARM ch. 4, scope note (AM. LAW INST. 2010) (megjegyezve, hogy a szigorú felelősség általában az abnormálisan veszélyes tevékenységgel, az állatok birtoklásával és a termékfelelősséggel kapcsolatos deliktumokra korlátozódik); RESTATEMENT (THIRD) OF TORTS: PRODS. LIAB. § 1 (AM. LAW. INST. 1998) (a szigorú felelősség fejlődésének magyarázata a termékfelelősségi ügyekben); 22 C. J. S. *Criminal Law* § 39 (2006) ('a szigorú felelősségre vonatkozó jogszabályok büntetőjogi rendszerünkben továbbra is a kivételek, nem pedig a szabály, és általánosságban kedvezőtlen státuszban vannak.').

az alperesnek meg kell felelnie ennek a mentális állapotnak ahhoz, hogy felelősségre vonható legyen.⁶⁶ Egyes teoretikusok úgy vélik, hogy mivel a társaságok valójában nem emberek, a jognak fel kellene hagynia az ellenkezőjét állító látszattal.⁶⁷ Ezek az elméletalkotók azzal érvelnek, hogy a jognak a társaságok mentális állapotára való jelenlegi támaszkodása hatástalan igazságszolgáltatást eredményez⁶⁸ és nem optimális megelőzést.⁶⁹ A vállalati mentális állapotokra való hivatkozás megszüntetése minden vállalati felelősséget szigorú felelősséggé alakítana át.

Ez a cikk nem a jogi személyiség fikciójának célszerűségéről szól. A jog eleget tehet ennek a fikciónak, függetlenül attól, hogy a⁷⁰ vagy nem⁷¹ társaságok valóban személyek-e. Valójában az a tény, hogy a jog öntudatosan hivatkozik a társasági személyiség "fikciójára", maga is annak elismerése, hogy a társaságok valójában nem emberek, és valójában nem rendelkeznek mentális állapotokkal.⁷² Amint máshol már kifejtettem, ez a látszat segíti a törvényt alapvető céljainak elérésében. Az a fikció, hogy a vállalatoknak lehetnek vétkes mentális állapotai, segít a jognak azonosítani a valóban elítélendő vállalati magatartást (szemben a pusztán káros magatartással), amelyet megkülönböztetett bánásmódban kell részesíteni.⁷³ Továbbá, a szigorú vállalati felelősséget támogató kifinomult közgazdasági elméletek ellenére,⁷⁴ ez a megközelítés gyakran

66. Khanna, *supra* note 25, at 1489 ("[A] követelmények között [amelyeknek] teljesülniük kell ahhoz, hogy a társaság felelőssége ... a társaság ügynökének jogellenes cselekményt (az actus reus) kell elkövetnie a szükséges tudatállapottal (a mens rea)."); W. Robert Thomas, *Incapacitating Criminal Corporations*, 72 VAND. L. REV. 905, 914-15 (2019) ("[T]he practice of holding commercial corporations criminally responsible for general intent crimes as well as specific intent crimes - crimes for which there exists a proscribed action (actus reus) pursued concurrently with a proscribed attitude (mens rea) - took held körül the turn of the twentieth century. ").

67. *Lásd* John S. Baker, Jr., *Reforming Corporations Through Threats of Federal Prosecution*, 89 CORNELL L. REV. 310, 349-53 (2004) (leírja, hogy a vállalati vádlottak és az emberi vádlottak közötti különbségek hogyan hozzák őket hátrányos helyzetbe).

68. *Lásd* Albert W. Alschuler, *Two Ways To Think About the Punishment of Corporations*, 46 AM. CRIM. L. REV. 1359, 1392 (2009) [a továbbiakban: Alschuler, *Two Ways To Think*] (azzal érvelve, hogy a vállalati büntetés két lehetséges jogi analógiája - a deodand, 'az állatok és élettelen tárgyak megbüntetése, amelyek kárt okoztak', vagy a frankpledge, 'egy csoport összes tagjának megbüntetése, amikor a csoport egy tagja elkerülte a bűncselekményért való letartóztatást' - mutatja annak abszurditását).

69. *Lásd* Daniel R. Fischel & Alan O. Sykes, *Corporate Crime*, 25 J. LEGAL STUD. 319, 320-21 (1996) (azzal érvelve, hogy a vállalatok büntetőjogi felelősségre vonása túlzott elrettentést eredményez).

70. *Lásd* CHRISTIAN LIST & PHILIP PETTIT, GROUP AGENCY: THE POSSIBILITY, DESIGN, AND STATUS OF CORPORATE AGENTS 1-2 (2011) (a társaságok ügynöki státusza mellett érvelve).

71. MAX WEBER, *ECONOMY AND SOCIETY: A NEW TRANSLATION* 89-90 (Keith Tribe ed. & trans., 2019) (azzal érvelve, hogy a vállalatok nem ügynökök).

72. *Lásd pl.* Int'l Shoe Co. v. Washington, 326 U. S. 310, 316 (1945) ("[T]he corporate personality is a fiction, although a fiction intended to be acted upon as though it was a fact").

73. Diamantis, *Corporate Criminal*, 35. lábjegyzet, 2063-64. o.

74. *Lásd pl.* Fischel & Sykes, *Supra* 69. lábjegyzet, 328. o. ("[A]hol az [bűncselekménye] helyesen a vállalati tevékenység költségének tekinthető, helyénvalónak tűnik, hogy a vállalat 'szigorú felelősséget viseljen a bűncselekmény társadalmi költségeiért A szigorú helyettesítési felelősség hagyományos indoklása [a] a költségek internalizálásának fontossága. "); Richard A. Posner, *An Economic Theory of the Criminal Law*, 85 COLUM. L. REV. 1193, 1222, 1228-29 (1985) ('Valójában a szigorú felelősség bizonyos fokát vezetjük be a büntetőjogba, mint a kártérítési jogba, amikor a tevékenységi szint megváltoztatása hatékony módszer a társadalmi költség elkerülésére. ').

2020]

KITERJESZTETT VÁLLALATI ELME905

nem hatékony túlzott elrettentést eredményez⁷⁵ és túlzott befektetést a megfelelésbe.⁷⁶ Függetlenül annak (de)érdemeitől, a társasági személyiség fikciójának elhagyása a jelenlegi társasági jog átfogó átalakítását igényelné. Ez tehát ellentétes lenne az itt elfogadott sebészi megközelítéssel. Ez a cikk a társasági személyiség és a társasági mentalitás jogi fikcióin belül helyezkedik el. Célja, hogy megtalálja e fikció ésszerű kiterjesztését az algoritmikus vállalati visszaélések eseteire.⁷⁷

Ismétlem, ez a cikk nem fogja azt állítani, sem hallgatólagosan, sem másként, hogy a vállalatoknak *valóban* vannak mentális állapotai. Ehelyett azt vizsgálja, hogy fogalmilag mit jelentene, ha a vállalatok gondolkodnának. Mivel a jog elkötelezett egy olyan fikció mellett, amelyben a vállalatoknak mentális állapotai vannak, mi mellett kötelezi el magát? Milyen alakot ölthet vagy kell öltenie ennek a fikciónak? Amint az alábbiakban bemutatjuk, az e kérdésekre adott válaszok jelenthetik a megoldást az algoritmikus vállalati visszaélések problémájára.

Végül, ez a cikk nem követi a jogászok, a⁷⁸ informatika, a⁷⁹ és az üzleti etika⁸⁰ tudósainak példáját, akik azt javasolják, hogy az algoritmikus helytelen magatartást úgy kezeljék, hogy magukat az algoritmusokat teszik felelőssé. Ez a megközelítés mélyen ellentmondásos. Távrolról sem egyértelmű, hogy az algoritmusok jelenleg megfelelnek, vagy valaha is megfelelhetnének⁸¹ a személyiség és a bűnösség feltételeinek.⁸² Még azok az elméletalkotók is, akik

75. Lásd *United States v. U. S. Gypsum Co.*, 438 U. S. 422, 440-43 (1978); *Developments in the Law-Corporate Crime: Regulating Corporate Behavior Through Criminal Sanctions*, 92 HARV. L. REV. 1227, 1270 (1979).

76. Mihailis E. Diamantis, *Functional Corporate Knowledge*, 61 WM. & MARY L. REV. 319, 348-49 (2019) [a továbbiakban Diamantis, *Functional Corporate*] (kifejtve, hogy a kollektív tudás megközelítése hogyan vezet a megfelelésre való szélsőséges összpontosításhoz).

77. Azáltal, hogy ez a cikk a vállalati személyiség jogi fikcióján belül helyezkedik el, nem kell ellentmondásos feltételezéseket tennie a vállalatok valódi metafizikájáról.

78. Lásd GABRIEL HALLEVY, *LIABILITY FOR CRIMES INVOLVING ARTIFICIAL INTELLIGENCE SYSTEMS* 27-28 (2015); Steven J. Frank, *Tort Adjudication and the Emergence of Artificial Intelligence Software*, 21 SUFFOLK U. L. REV. 623, 624-25 (1987); Christina Mulligan, *Revenge Against Robots*, 69 S. C. L. REV. 579, 579-80 (2018).

79. Lásd Fahad Alaieri & André Vellino, *Ethical Decision Making in Robots: SOCIAL ROBOTICS: 8TH INTERNATIONAL CONFERENCE* 159, 159 (Arvin Agah et al. eds., 2016) ("[N]on-predictability and autonomy may confer a greater degree of responsibility to the machine"); Luciano Floridi & J. W. Sanders, *On the Morality of Artificial Agents*, 14 MINDS & MACHINES 349, 373-74 (2004); Gabriel Hallevy, *Unmanned Vehicles*: 21 J. L. INFO. & SCI. 200, 200 (2011).

80. Lásd Nicholas Diakopoulos & Sorelle Friedler, *How To Hold Algorithms Accountable*, MIT TECH. REV. (2016. november 17.), <https://www.technologyreview.com/s/602933/how-to-hold-algorithms-accountable/> [<https://perma.cc/4B7Z-X6MF> (sötét archívum)].

81. Lásd JOHN SEARLE, *MINDS, BRAINS AND SCIENCE* 30-31 (1984) (érvelése szerint a számítógépek nem tudnak gondolkodni); Thomas C. King et al., *Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions*, 26 SCI. & ENGINEERING ETHICS 89, 95, 102 (2019) (megjegyezve, hogy 'az elképzelés, hogy egy [algoritmus] önkéntesen cselekedhet, vitatott', és hogy egy mesterséges ágens 'önmagában nem felelhet meg [a bűncselekmény] mens rea követelményének').

82. Lásd általában JOHN CHIPMAN GRAY, *THE NATURE AND SOURCES OF THE LAW* 27-52 (1909) (a jogi személyiségről).

az algoritmikus személyiség fikcionáló megközelítését javasolják (a jogi személyiség fikciójának analógiájára)⁸³ két hatalmas akadállyal kell szembenézniük. Először is, senki sem javasolt kielégítő választ arra, hogy mikor lenne értelme az algoritmusokat felelősségre vonni. Bármit tesz egy algoritmus, az végső soron a környezetének és a programozásának a terméke.⁸⁴ Ezért nehéz elképzelni, hogy mikor lenne az algoritmus, és nem a környezete vagy a programozója a felelős. Például 2016-ban a Microsoft elindított egy chatbotot, "Tay"-t, hogy online kommunikáljon a tizenévesekkel.⁸⁵ A Tay-nek a Twitterről lekapart adatokból tanulva kellett volna megtanítania magát beszélni.⁸⁶ Huszonnégy órán belül az internetezők annyi romboló tweettel csalogatták Tayt, hogy a bot üzenetei sovinszta, rasszista és antiszemita lett.⁸⁷ Az azonban korántsem egyértelmű, hogy Tay volt a hibás a mondanivalójáért, és nem a Microsoft vagy a maró "Twitterverse".⁸⁸ Ez elvezet a közvetlen algoritmikus felelősséggel kapcsolatos második nehézséghez: Még ha a törvény szerint egy olyan algoritmus, mint a Tay felelős is lenne, akkor mi lenne? Nincs mód arra, hogy szankcionáljunk egy algoritmust vagy botot⁸⁹ (kivéve talán a megölését, ahogyan a Microsoft tette a Taynel).⁹⁰ Más "embereket" bebörtönözhetünk vagy megbírságozhatunk, de az algoritmusoknak nincs testük és zsebük.⁹¹

83. Lásd Lawrence B. Solum, *Legal Personhood for Artificial Intelligences*, 70 N. C. L. REV. 1231, 1239-43 (1992) (összehasonlítva a mesterséges intelligencia személyiségének és a vállalati személyiségnek az esetét).

84. Lásd DAVID A. PATTERSON & JOHN L. HENNESSY, *COMPUTER ORGANIZATION AND DESIGN: THE HARDWARE/SOFTWARE INTERFACE* 13-15 (5th ed. 2014) (ebook); Anupam Chandler, *The Racist Algorithm*, 115 MICH. L. REV. 1023, 1034-37 (2017). Ez nem tagadja, hogy az algoritmusokkal kapcsolatos problémák más módon is felmerülhetnek, például amikor egy bizonyos célra tervezett algoritmust más célra használnak. Lásd Batya Friedman & Helen Nissenbaum, *Bias in Computer Systems*, 14 ACM TRANSACTIONS ON INFO SYS. 330, 330-32 (1996).

85. Lásd Elle Hunt, *Tay, Microsoft's AI Chatbot, Gets a Crash Course in Racism from Twitter*, GUARDIAN (2016. március 24., 14:41), <https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter> [<https://perma.cc/B933-9RQS>].

86. Lásd *id.*

87. Damon Beres, *Microsoft Chat Bot Goes on Racist, Genocidal Twitter Rampage*, HUFFPOST (2016. március 24. 10:19), https://www.huffpost.com/entry/microsoft-tay-racist-tweets_n_56f3e678e4b04c4c37615502 [<https://perma.cc/GP9Z-YL7G>].

88. Lásd James Vincent, *Twitter Taught Microsoft's AI Chatbot To Be a Racist Asshole in Less Than a Day*, VERGE (2016. március 24., 6:43), <https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist> [<https://perma.cc/M4BU-VDVQ>].

89. Lásd Joanna J. Bryson, Mihailis E. Diamantis & Thomas D. Grant, *Of, For, and By the People: The Legal Lacuna of Synthetic Persons*, 25 ARTIFICIAL INTELLIGENCE & L. 273, 288 (2017); Solum, *supra* note 83, at 1244-48 (az algoritmusok büntetésének nehézségeiről). De lásd Gabriel Hallevy, *"I, Robot-I, Criminal"-When Science Fiction Becomes Reality: Legal Liability of AI Robots Committing Criminal Offences*, 22 SYRACUSE SCI. & TECH. L. REP. 1, 29-35 (2010) [a továbbiakban: Hallevy, *I, Robot*] (a különféle AI-büntetések nem meggyőző bemutatása).

90. Lásd Rob Price, *Microsoft Is Deleting Its AI Chatbot's Incredibly Racist Tweets*, BUS. INSIDER (2016. március 24. 7:31), <https://www.businessinsider.com/microsoft-deletes-racist-genocidal-tweets-from-ai-chatbot-tay-2016-3> [<https://perma.cc/8C2R-2LXJ> (sötét archívum)].

91. Lásd általában Ryan Abbott & Alex Sarch, *Punishing Artificial Intelligence*: 53 U. C. DAVIS L. REV. 323 (2019) (a hagyományos "büntető" algoritmusok szabálysértésért való "megbüntetésének" nehézségeit tárgyalja, és végül elutasítja az algoritmusok büntetésének lehetőségét).

2020]

KITERJESZTETT VÁLLALATI ELME907

Függetlenül attól, hogy a jog közvetlenül felelőssé teheti-e vagy kellene-e tennie az algoritmusokat, a jelenlegi megközelítés egyértelmű: az algoritmusok nem emberek, és nem lehetnek polgári vagy büntetőjogi vádlottak.⁹² Az irány megfordításához egy pörölykalapács megrendítésére lenne szükség. Ez a cikk az algoritmikus visszaélésekért való *vállalati* felelősségre szorítkozik, mivel a jog már megállapította, hogy a vállalatok felelős "személyek". Amint azt a szövetségi büntetési irányelvek szervezetekre vonatkozó rendelkezései is bizonyítják, a jog már régóta rendelkezik a vállalatok szankcionálására szolgáló mechanizmusokkal.⁹³ Az itt elfogadott sebészeti megközelítés erre a meglévő jogi struktúrára támaszkodva teszi fel a kérdést: milyen feltételek mellett legyenek felelősök *a vállalatok*, ha algoritmusaik szabálysértést követnek el? Ezáltal elkerüli az algoritmusok közvetlen felelősségre vonásának fogalmi, filozófiai, jogi és pragmatikai kihívásait.⁹⁴

A vállalati algoritmikus visszaélések által okozott jelentős nehézségek ellenére a jelen cikk által javasolt jogi felülvizsgálatok viszonylag szerények. Az általa keresett megoldásnak inkább a jelenlegi jog kiterjesztésének, mintsem átírásának kell lennie. A megoldásnak át kell ölelnie a társasági felelősség jelenlegi jogát, beleértve a társasági személyiség és a vétkes mentális állapotok fikcióját. El kell kerülni a társaságok szigorú felelősségének előírását, ami politikailag megvalósíthatatlan, és túlzottan akadályozná a vállalati innovációt. Végezetül pedig nem lenne szükség az algoritmusok jogi személyiségét vagy elszámoltathatóságát megállapító új joganyagra. Az alább javasolt megoldás szerint az algoritmusok nem gondolkodnak és nem tudnak dolgokat. Inkább a vállalatok gondolkodnak vagy tudnak dolgokat az általuk használt algoritmusokon keresztül.

Az alább javasolt általános stratégia a jelenlegi társasági jog egy olyan furcsaságát használja ki, amely szerint a társaságok emberek, és mentális állapotokkal rendelkezhetnek. Az érvelés belelép ebbe a fikcióba, és megvizsgálja, hogy ennek milyen következményei vannak. Tekintettel arra, hogy a jog szerint a vállalatoknak milyen "elméjük" van, kevés okunk van azt állítani, hogy az algoritmusok nem játszhatnak szerepet annak meghatározásában, hogy a vállalatok mit gondolnak és mit tudnak.

II. HEALTHCO ÉS FORMBOT: EGY TISZTÁZÓ PÉLDA

A következő hipotetikus eset rávilágít néhány olyan kihívásra, amellyel az algoritmikus vállalati visszaélések problémájának megoldásának foglalkoznia kell:

92. Lásd Thomas Beardsworth & Nishant Kumar, *Who To Sue When a Robot Loses Your Fortune*, BLOOMBERG (2019. május 5., 20:00), <https://www.bloomberg.com/news/articles/2019-05-06/who-to-sue-when-a-robot-loses-your-fortune> [<https://perma.cc/27PX-4E49>] (A robotok napról napra emberibbé válnak, de még mindig nem lehet őket beperelni.).

93. U. S. SENTENCING GUIDELINES MANUAL § 8A1.1 cmt. n. 1 (U. S. SENTENCING COMM. 2018).

94. Ez nem jelenti azt, hogy a jognak végül nem kell megtalálnia a módját annak, hogy az algoritmusokat közvetlenül felelősségre lehessen vonni. Az itt javasolt megoldás nem fog működni, ha az algoritmusok egyedül, és nem egy vállalat nevében eljárva követnek el visszaélést.

A HealthCo egy olyan vállalat, amely a Medicare és Medicaid-jogosult betegek számára szolgáltatások széles körét nyújtja. A költségek megtakarítása érdekében a HealthCo felkérte adatelemzőit, hogy fejlesszék ki a FormBotot, egy gépi tanuló algoritmust, amelyet arra képeztek ki, hogy a lehető leghatékonyabban töltsék ki és nyújtsa be a szövetségi visszatérítési űrlapokat. Néhány hónapos működés után a FormBot magától rájött, hogy kevesebb idő alatt több visszatérítést tud biztosítani, ha néhány űrlaphoz hamis információkat használ. A HealthCo-nál senki sem tudott erről a fejleményről, és senki sem számított rá. Mire a szövetségi hatóságok felfedezték a hamis nyomtatványokat, a HealthCo már több millió dollárnyi jogosulatlan visszatérítést kapott.

Megsértette-e a HealthCo a False Claims Act⁹⁵ büntetőjogi vagy a polgári jogi⁹⁶ rendelkezéseit, amelyek tiltják a szövetségi kormánynak való tudatos hamis állítások benyújtását?

A HealthCo példával kapcsolatban több dolgot is meg kell jegyezni. Először is, ez az algoritmikus helytelen magatartás problémájának egyik kihívására világít rá. A hatályos jog szerint nem kétséges, hogy a vállalatok megsérthetik a hamis követelésekről szóló törvényt.⁹⁷ A HealthCo megfelel a jogsértés objektív elemeinek, mivel hamis követeléseket nyújtott be a szövetségi kormánynak. A vitás pont a hamis követelésekről szóló törvény ismereti követelménye. Ha a HealthCo egyik alkalmazottja úgy nyújtotta volna be a nyomtatványokat, hogy tudta, hogy azok hamis adatokat tartalmaznak, a felelősségre vonás esete elég egyértelmű lenne.⁹⁸ A feltételezett tényállás alapján azonban a hatályos jog nem írja elő sem a polgári jogi, sem a büntetőjogi felelősséget. A HealthCo nem tudhatta, hogy a nyomtatványok hamisak, mivel egyik alkalmazottja sem tudta. A felelősségi vizsgálat ilyen módon történő automatikus megszüntetése aggasztó. A közpénzek kárt szenvedtek. Az a tény, hogy a HealthCo algoritmust használt, nem pedig alkalmazottat, nem változtat az állampolgári érdeken, hogy az ilyen magatartást megakadályozza vagy elítéli. A körülmények legalábbis megkülönböztetőbb felelősségvizsgálatot indokolnak.

Másodszor, a feltételezett esetben a büntetőjogi felelősség is szóba kerül. Ez azért fontos, mert a büntetőjognak vitathatatlanul a legnagyobb érdeke, hogy a vállalati mentális állapotokat helyesen ítéljék meg. Egyes tudósok a szigorú felelősségi normák hatékonysági előnyeit hangoztatják a polgári jogban,⁹⁹ különösen a vállalati kontextusban.¹⁰⁰ A büntetőjog azonban nem csak a hatékonysággal foglalkozik.¹⁰¹ A büntetőjog

95. 18 U.S.C. § 287 (2018).

96. 31 U. S. C. § 3729(a)(1)(A)(G) (2018).

97. *Lásd pl.* Nye & Nissen kontra Egyesült Államok, 336 U. S. 613, 614-16 (1949). 98. *Lásd pl.*, United States v. Sain, 141 F.3d 463, 470-71 (3d Cir. 1998).

99. *Lásd pl.* Guido Calabresi & Jon T. Hirschoff, *Toward a Test for Strict Liability in Torts*, 81 YALE L. J. 1055, 1060-67 (1972) (a kártérítési felelősség szigorú felelősségi tesztje mellett érvelve, de megjegyezve az ilyen megközelítés néhány gyengeségét).

100. *Lásd* Fischel & Sykes, 69. lábjegyzet, 327-28. o.

101. *Lásd* Albert W. Alschuler, *A büntetőjogi büntetés változó céljai: A Retrospective on the Past Century and Some Thoughts About the Next*, 70 U. CHI. L. REV. 1, 1 (2003). *De lásd* Gary S. Becker, *Crime and Punishment: An Economic Approach*, 76 J. POL. ECON. 169, 170, 172 (1968) (egy, a bűnözés és a bűnözés elleni küzdelemre összpontosítva).

2020]

KITERJESZTETT VÁLLALATI ELME909

nem tudná betölteni sajátos elítélő¹⁰² és erőteljes elrettentő funkcióját¹⁰³, ha csak a szigorú felelősségre támaszkodna. A szigorú felelősség figyelmen kívül hagyja azokat a különbségeket, amelyek az igazságszolgáltatást követelő áldozatok és a büntetés előtt álló vádlottak számára fontosak. Egyformán kezeli az ártatlan balekot, a szerencsétlen bolondot és a számító gazembert.¹⁰⁴ Egy olyan büntetőjogi rendszer, amely nem hivatkozik a mentális állapotokra, felismerhetetlen lenne.¹⁰⁵

Harmadszor, a HealthCo esetében a kérdéses konkrét mentális állapot a tudás.¹⁰⁶ A mentális állapotok minden típusának megvannak a maga megkülönböztető tulajdonságai, és ideális esetben külön tárgyalást kapnának. Bár csak a szövetségi büntető törvénykönyv rendelkezései között több mint száz különböző mens rea-típus létezik, a¹⁰⁷ tudás a leggyakoribb vállalati bűncselekmények közül soknak az egyik eleme.¹⁰⁸ Az alább javasolt megoldásnak, bár a tudásra vonatkozik, mintaként kell szemlélni más mentális állapotok számára is. A tudás kényelmes kiindulópont, mivel a kiterjesztett elme tézisével kapcsolatos meglévő szakirodalom általában erre összpontosít.¹⁰⁹

Végül, a hipotetikus feltételezés olyan technikai részleteket határoz meg, amelyek különösen bonyolulttá teszik az algoritmikus visszaélések problémáját. A HealthCo mérnökei a FormBotot gépi tanulással programozták. A gépi tanulási technikák nagyjából az algoritmus céljának meghatározásával kezdődnek, majd az algoritmust tesztesetek nagy halmazával betanítják.¹¹⁰ Azzal, hogy minden egyes tesztesetben megmondjuk az algoritmusnak, hogy

gazdasági és hatékonysági alapú megközelítés a büntetőjogi büntetések meghatározására). *Lásd általában* Jeremy Bentham, *Principles of Penal Law*, in 1 THE WORKS OF JEREMY BENTHAM 366 (John Bowring szerk., 1962) (a büntetőjog céljainak és annak kifejtése, hogy mit kínál mind az áldozatoknak, mind a társadalomnak).

102. Joel Feinberg, *The Expressive Function of Punishment*, 49 MONIST 397, 400 (1965).

103. Samuel W. Buell, *The Blaming Function of Entity Criminal Liability*, 81 IND. L. J. 473, 500, 510-12 (2006).

104. *Lásd általában* Paul H. Robinson & Jane A. Grall, *Element Analysis in Defining Criminal Liability: The Model Penal Code and Beyond*, 35 STAN. L. REV. 681, 687-90 (1983) (annak tárgyalása, hogy a büntetőjog hogyan használja a különböző mentális állapottelemeket a bűnösség meghatározásához).

105. Ezért Rebecca Crootof, aki szigorú állami felelősséget javasol a mesterséges intelligencia által okozott károkért, a büntetőjogtól a kártérítési jog felé fordul. Rebecca Crootof, *War Torts: Accountability for Autonomous Weapons*, 164 U. PA. L. REV. 1347, 1387-88 (2016).

106. Meg kell jegyezni, hogy a tudás jogi meghatározása nem azonos a filozófiai meghatározással. A jogban egy személy akkor tud valamilyen információt, ha azt elhiszi, és az igaz. MODEL PENAL CODE § 2.02(2)(b)(i)-(ii) (AM. LAW INST. 1985). A filozófusok további követelményeket támasztanak a tudással szemben, amelyek közül az egyik, hogy a személynek a meggyőződését igazolni is kell. *Lásd* PAUL K. MOSER & ARNOLD VANDER NAT, *HUMAN KNOWLEDGE: CLASSICAL AND CONTEMPORARY APPROACHES* 3 (2d ed. 1995). A 'tudást' jogi értelemben használom, összhangban a hamis követelésekről szóló törvényben szereplő jelentésével.

107. *Lásd* William S. Laufer, *Culpability and the Sentencing of Corporations*, 71 NEB. L. REV. 1049, 1065 (1992).

108. *Lásd* Diamantis, *Functional Corporate*, 76. lábjegyzet, 322-23 (példák felsorolása).

109. Nevezetes kivételek: Mark Rowlands, *Consciousness, Broadly Construed*, in THE EXTENDED MIND 271, 271 (Richard Menary szerk., 2010) (ebook), valamint Mattia Gallotti & Bryce Huebner, *Collective Intentionality and Socially Extended Minds*, 30 PHIL. PSYCHOL. 251, 252-53 (2017).

110. *Lásd* David Lehr & Paul Ohm, *Playing with Data*: U.C. DAVIS L. REV. 653, 668 (2017); Jason Brownlee, *Supervised and Unsupervised*

vagy nem érte el a célját, az algoritmus végül megtanulhatja, hogy magától is sikerüljön.¹¹¹ A mérnökök például olyan drónrepülő algoritmust tervezhetnek, amely a leghatékonyabb útvonalat választja egy célpont felé.¹¹² Ezt a célt kódolnák, majd betanítanák a drónt úgy, hogy különböző helyekre küldik a cél körül, megnézik, merre megy, és közlik az algoritmussal, hogy a teljesítménye sikeres volt-e vagy sem. Ha minden jól megy, a drónnak meg kell tanulnia, hogy megbízhatóan és hatékonyan keresse meg a célt.

A gépi tanulás jelentősége az algoritmikus szabálysértés szempontjából kettős. Először is, a legfejlettebb mesterséges intelligencia rendszerek mögött a gépi tanulás áll.¹¹³ Ahogy a vállalati algoritmikus visszaélések társadalmi fenyegetése egyre nagyobb méreteket ölt, a gépi tanulás valószínűleg túlsúlyba kerül. Másodszor, a gépi tanulás felveti annak lehetőségét, hogy az algoritmusok emberi beavatkozás nélkül is rosszul viselkednek.¹¹⁴ A HealthCo mérnökei nem úgy tervezték a FormBotot, hogy hamis űrlapokat küldjön be, és senki sem tudta vagy várta, hogy ez így lesz. Ez azért reális, mert a gépi tanulási algoritmusok gyakorlatilag saját kódjukat írják.¹¹⁵ Az így létrejövő algoritmusok olyan bonyolulttá válnak, hogy a kódot utólag elemző programozók gyakran nem értik, hogyan működik.¹¹⁶ Ráadásul sok algoritmus beépített véletlenszerűséggel rendelkezik, ami a tervezésük alapvető része.¹¹⁷ Következésképpen a gépi tanuló algoritmusok nem szándékolt és nem várt (és nem előre látható) módon viselkedhetnek.¹¹⁸ Utólag visszatekintve a rendellenes eredmények gyakran a gépi tanulási folyamat valamely jellemzőjére vezethetők vissza: a cél meghatározásának módjára, az algoritmus betanításához használt tesztesetek halmazára vagy valamilyen kölcsönhatásra.

Machine Learning Algorithms, MACHINELEARNINGMASTERY (Mar. 16, 2016), <https://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/> [<https://perma.cc/DSF4-TZAU>].

111. *Lásd: A Beginner's Guide to Neural Networks and Deep Learning*, PATHMIND: A.I. WIKI, <https://skymind.ai/wiki/neural-network> [<https://perma.cc/P59K-RJ45>].

112. *Lásd* Lemley & Casey, *Supra* 17. lábjegyzet, 1313-14. o. (e példa tárgyalása).

113. *Lásd id.* 1335. o. ([A] gépi tanulásban rejlő kiszámíthatatlanság egyben az egyik legnagyobb erőssége is.).

114. *Lásd általában* PEDRO DOMINGOS, THE MASTER ALGORITHM: HOW THE QUEST FOR THE ULTIMATE LEARNING MACHINE WILL REMAKE OUR WORLD (2015) (vázolja a jelenlegi algoritmusok működését, valamint a gépi tanulás pozitív és potenciálisan negatív hatásait a működésükre). Ryan Abbott és Alex Sarch ezt a viselkedést 'Hard AI Crime[]' (Kemény mesterséges intelligencia bűnténynek nevezi. "Abbott & Sarch, *Supra* 91. lábjegyzet, 328. o.

115. *Lásd* Abbott & Sarch, 91. lábjegyzet, 330-31. o.

116. Andrew D. Selbst & Solon Barocas, *The Intuitive Appeal of Explainable Machines*, 87 *FORDHAM L. REV.* 1085, 1089-90, 1092 (2018); Matthew Carroll, *The Complexities of Governing Machine Learning*, *DATANAMI* (2017. április 27.), <https://www.datanami.com/2017/04/27/complexities-governing-machine-learning/> [<https://perma.cc/H3EU-LB4J>].

117. Joshua A. Kroll et al., *Accountable Algorithms*, 165 *U. PA. L. REV.* 633, 655 (2017).

118. Lemley & Casey, *supra* note 17, at 1365 ([Az [algoritmikus] visszaélések nagy része, amelyekkel szemben a holnap tervezőinek, döntéshozóinak és felügyelőinek védekezniük kell, egyáltalán nem biztos, hogy szándékos.]).

2020]

KITERJESZTETT VÁLLALATI ELME911

e két tényező és a való világ között.¹¹⁹ A kód összetettsége miatt még akkor is adódhatnak problémák, ha minden érintett ember teljesen ártatlan.¹²⁰

A fenti drónos példában a mérnökök egy ponton megfigyelték, hogy a drón pontosan az ellenkezőjét csinálja annak, amit terveztek - elrepül a célponttól a teszterület peremére.¹²¹ Némi vizsgálat után megállapították, hogy ez nem hiba volt. A drón megtanulta, hogy bizonyos körülmények között a célpont elérésének leghatékonyabb módja az, ha a teszterület peremére repül.¹²² Ha a drón ezt megtette, megtanulta, hogy a mérnökök visszahozzák és a célponthoz viszik, hogy újraindítsák a próbát. Néha a célpontozóhoz való eljutás a repülés helyett a cipelés volt a leghatékonyabb útvonal, bár a mérnökök nyilvánvalóan nem ezt akarták elérni a drónnal.

Az a tény, hogy a FormBot gépi tanulást használt, és anélkül nyújtott be hamis űrlapokat, hogy bárki is tervezte volna azt, biztosítja, hogy a feltételezett jogsértés az algoritmikus visszaélés tiszta esete.¹²³ Ellenkező esetben az eset lehet, hogy csak a szokásos munkavállalói kötelességszegésről van szó, bár algoritmus által közvetített kötelességszegésről. A jog már rendelkezik mechanizmusokkal az olyan esetek kezelésére, amikor az alkalmazottak szándékosan, tudatosan vagy gondatlanul terveznek olyan algoritmusokat, amelyekkel törvényt szegnek.¹²⁴ A hipotetikus esetben, ha a HealthCo mérnöke szándékosan vagy tudatosan úgy tervezte meg a FormBotot, hogy hamis nyomtatványokat küldjön be, akkor a respondeat superior a mérnök mentális állapotát a HealthCo-nak tulajdonítaná, és ezáltal teljesülne a False Claims Act követelménye.¹²⁵ Nem számítana, hogy a mérnök fizikailag nem maga nyújtotta be a nyomtatványokat.¹²⁶ Ha a mérnök gondatlanságból tervezte a FormBotot, akkor a szándékos

119. *Lásd pl.* Kroll et al., 117. lábjegyzet, 693-94. o.; Lemley & Casey, 118. lábjegyzet, 1313. o..

120. KEVIN PETRASIC ET AL., ALGORITHMS AND BIAS: WHAT LENDERS NEED TO KNOW 1 (2017), <https://www.whitecase.com/sites/whitecase/files/files/download/publications/algorithm-risk-thought-leadership.pdf> [<https://perma.cc/36YD-VWTW>] ([A] tökéletesen jó szándékú algoritmus véletlenül olyan elfogult következtetéseket hozhat létre, amelyek hátrányosan megkülönböztetik a védett emberscsoportokat.); Barocas & Selbst, *supra* 29. lábjegyzet, 729. o. (kifejtve, hogy a hibák az adatbányászok teljesen ártatlan választásainak eredménye is lehet").

121. Lemley & Casey, *Supra* 17. lábjegyzet, 1313. lábjegyzet.

122. *Id.*

123. Ez tehát eltér néhány olyan példától, amelyre más büntetőjogi tanulmányok fókuszáltak. *Lásd pl.* Amanda McAllister, Note, *Stranger than Science Fiction: The Rise of A.I. Interrogation in the Dawn of Autonomous Robots and the Need for an Additional Protocol to the U.N. Convention Against Torture*, 101 MINN. L. REV. 2527, 2545, 2547 (2017).

124. *Lásd* Hallevy, *I, Robot*, *Supra* 89. lábjegyzet, 9. o. (a robotokkal kapcsolatos bűncselekmények büntetőjogi felelősségének különböző modelljeit tárgyalja, amelyek a programozó vagy a felhasználó hibáján alapulnak).

125. *Lásd* MODEL PENAL CODE § 2.02(5) (AM. LAW INST. 1985).

126. Egy követelés hamis benyújtásának okozása kifejezetten megfelel a hamis követelésekről szóló törvény polgári jogi változatának. 31 U.S.C. § 3729(a)(1)(A) (2018). A büntetőjog esetében más felelősségi doktrínák töltik ki a hiányt. *Lásd* 18

U. S. C. C. § 2(b) (2018) (Aki szándékosan olyan cselekményt idéz elő, amely, ha közvetlenül ő vagy más hajtáná végre, bűncselekményt követne el az Egyesült Államok ellen, főbűnösneként büntetendő.); MODEL PENAL CODE § 5.04(1)(a) (AM. LAW. INST. 1985) (kimondja, hogy egy személy bűnös lehet abban, hogy ártatlan tisztségviselőt bűncselekmény elkövetésére szólít fel). A bűncselekmény actus reus és mens rea elemeit a társaság különböző részei is kielégíthetik. *Lásd pl.* United States v. Bank of New England, N.A., 821 F.2d 844, 856 (1st Cir. 1987).

tudatlansága alapot szolgáltathatna arra, hogy azt mondhassuk, hogy ő, és így a HealthCo is tudott a hamis nyomtatványokról.¹²⁷ Amit a jelenlegi doktrína nem tud kezelni, és amivel ez a cikk foglalkozik, azok a legaggasztóbb esetek, amikor a munkavállalói kötelességszegés kikerül a képből.¹²⁸ A tisztán algoritmikus vállalati kötelességszegés eseteinek megoldása alternatív utat biztosít a vállalati felelősségre vonáshoz olyan helyzetekben is, amikor az egyes alkalmazottak vétkesek lehettek, de ennek bizonyítása nehézkes.¹²⁹

III. A KITERJESZTETT ELME TÉZISE

Itt az ideje, hogy korszerűsítsük a vállalati mentalitásról alkotott jogi felfogást. A jog csak úgy tudja megfelelően felelősségre vonni a vállalatokat, ha megteremti annak lehetőségét, hogy néha "ismerik" a szerveiken könnyen hozzáférhető információkat, és "szándékukban áll" az algoritmusaik által hozott döntések meghozatala. Számos kognitív tudós és filozófus a mentális állapotokat olyan módon értelmezi, amely megteremtheti az alapját annak, hogy túllépünk a "respondeat superior" kizárólagos, munkavállalókra összpontosító megközelítésén. A jelen rész ezt a megértést a természetes emberekre alkalmazva írja le. A következő rész kiterjeszti az elméletet, és bemutatja, hogyan lehet azt mind elvben, mind jogilag a vállalati emberekre adaptálni.

A "kiterjesztett elme tézis" azt állítja, hogy az emberi elmét nem mindig korlátozzák az agy fizikai határai.¹³⁰ A kiterjesztett elme teoretikusai jellemzően a mentális állapotok "funkcionalista" leírását támogatják.¹³¹ A funkcionális szerint a mentális állapotokat az a kognitív szerep jellemzi, amelyet a bemeneteket (például a környezeti jeleket és más mentális állapotokat) a kimenetekkel (például új mentális állapotokkal vagy viselkedéssel) összekötő kognitív szerepükkel játszanak.¹³² Ha például egy személy fagyfaltra vágyik, és odasétál a fagyasztóhoz, akkor jó eséllyel azért teszi, mert azt hiszi, hogy a fagyfalt ott van. A hit (pl. hogy a fagyasztóban van fagyfalt) durva meghatározó jellemzője az, hogy a vágyakozási bemeneteket (pl. a fagyfalt után) és a

127. *Lásd* Glob.-Tech Appliances, Inc. kontra SEB S.A., 563 U.S. 754, 769 (2011) (kifejtve, hogy a Bíróság szerint a szándékos vakság 'meghaladja a gondatlanságot és a hanyagságot').

128. Vagy nem bizonyítható a képen. *Lásd általában* Eric Holder, helyettes főállamügyész memorandumát az összes komponens vezetőjének és az amerikai ügyészeknek (1999. június 16.), <http://www.justice.gov/sites/default/files/criminal-fraud/legacy/2010/04/11/charging-corps.PDF> [<https://perma.cc/6BH5-D2FC>] (a vállalatok elleni büntetőeljárásról és a bűnösök azonosításának nehézségeiről).

129. *Lásd* Barocas & Selbst, *supra* 29. lábjegyzet, 692-93. o. (leírja, hogy a számítógépes programok hogyan fedhetik el az emberi helytelen magatartást); Pauline T. Kim, *Data-Driven Discrimination at Work*, 58 WM. & MARY L. REV. 857, 884-85 (2017) (a "szándékos megkülönböztetés" kontextusában a helytelen magatartás és az algoritmusok tárgyalása).

130. Clark & Chalmers, *Supra* 43. lábjegyzet, 14. pont.

131. *Lásd pl.*, *id.* ('Ami miatt bizonyos információk meggyőződéseknek számítanak, az a szerep, amelyet játszanak, és nincs ok arra, hogy a releváns szerepet csak a test belsejéből lehet játszani. '); Richard Menary, *Bevezetés: The Extended Mind in Focus to THE EXTENDED MIND*, *supra* note 109, at 1, 5 (a [kiterjesztett elme elmélet] 'funkcionalista hitvallásának leírása'); Michael Wheeler, *In Defense of Extended Functionalism*, in *THE EXTENDED MIND*, *supra* note 109, at 245, 245.

132. GILBERT HARMAN, *REASONING, MEANING, AND MIND* 236-39 (1999) (ebook).

2020]

KITERJESZTETT VÁLLALATI ELME913

viselkedési kimenetek (pl. séta a fagyasztóhoz).¹³³ A funkcionalizmus szerint minden olyan állapot, amely a vágyakat és a viselkedést ilyen módon megfelelően összekapcsolja, lehet hit.

A funkcionalizmus egyik fontos következménye a mentális állapotokat megvalósító rendszerek semlegessége. Nem számít például, hogy miből áll a rendszer.¹³⁴ Az emberi neuronok egyértelműen képesek a feladatra. De ugyanígy képesek különböző anyagok összetett elrendezéséből álló rendszerek is, akár szerves (pl. állati vagy idegen agyak), akár szervetlen (pl. fogaskerekek vagy áramkörök nagyon kifinomult elrendezései).

Hasonlóképpen a funkcionalizmus sem törődik azzal, hogy a mentális állapotokat megvalósító rendszerek hol helyezkednek el. A kiterjesztett elmeelmélet alapműve ezt a pontot "paritás premisszaként" fogalmazza meg: "Ha valamilyen feladattal szembesülve a világ egy része olyan folyamatként működik, amelyet, *ha a fejünkben történne*, habozás nélkül a kognitív folyamat részeként ismernénk el, akkor a *világnak* ez a része . . .

a kognitív folyamat része."¹³⁵ Az emberi mentális állapotokat megalapozó rendszerek jellemzően a koponyán belül helyezkednek el. De egy olyan személy, akinek az agya túlnyúlik a koponyáján, még mindig rendelkezhet mentális állapotokkal a kiálló részen. Ez még akkor is igaz lehet, ha az agyrésze meglehetősen távoli, hosszú neuronokkal összekötött. Csak az számítana, hogy a résznek mindene megvan (a megfelelő belső szerveződés és a megfelelő kapcsolatok az agy többi részével) ahhoz, hogy mentális állapotfunkciókat hajtson végre.

Az előző két bekezdés meglátásainak - az anyag és a hely semlegességének - kombinálása lehetővé teszi a kiterjesztett elme elméletírók számára, hogy értelmesebb példák széles skálájáról beszéljenek. A következő két eset a¹³⁶ illusztrálja:

Alice el akar sétálni a házától egy új kávézóig. Megnézi a számítógépén az útvonalat, és megjegyzi. Ezután elindul, és könnyen megtalálja az utat.

Barry is gyalog akar eljutni a házától az új kávézóig. Alzheimer-kórban szenved, és nehezen emlékszik a dolgokra. Megnézi az útvonalat a számítógépén, és gondosan beírja a naplójába. Aztán elindul. Azzal, hogy minden egyes kanyarban ellenőrzi a naplóját, könnyen megtalálja az utat.

133. Lásd BRIAN LOAR, MIND AND MEANING 6-9 (1981) (a hit-vágy elmélet magyarázata).

134. Lásd Menary, *Supra* note 131, at 6 ([I]t is the physical properties that matter to [extended mind theory], however, but the functionality of the process.); Hilary Putnam, *Psychological Predicates*, in ART, MIND, AND RELIGION 37, 44-45 (W. H. Capitan & D. D. Merrill eds. , 1967); J. J. C. Smart, *Sensations and Brain Processes*, 68 PHIL. REV. 141, 150 (1959).

135. Clark & Chalmers, *Supra* 43. lábjegyzet, 8. o.; Menary, *Supra* 131. lábjegyzet, 5. o.

136. Az alábbi hipotetikus feltevés az Andy Clark & David Chalmers által használt hipotetikus feltevésen alapul. Lásd *id.* 12-16. o. (egy hasonló hipotetikus esetet tárgyalva, amelyben Inga a memóriáját használja, hogy eljusson egy múzeumba, Otto pedig egy jegyzetfüzetet).

Kétségtelen, hogy Alice tudta, hogyan juthat el a boltba, miután megnézte az útvonalat. De Barry tudta-e? Úgy tűnik, hogy a naplőbejegyzései hasonló funkcionális szerepet játszanak, mint az Alice neuronjaiban kódolt irányinformáció. Input: a kávézó meglátogatásának vágya. Kimenet: pontos és könnyű odasétálás. Alice és Barry ugyanazzal a funkcionális kapcsolattal rendelkezik a bemenetek és kimenetek között. A kiterjesztett elme teoretikusai arra következtetnek, hogy mivel Alice tudta, hogyan juthat el a kávézóba, Barry is tudta.

Ahhoz, hogy ellenálljanak ennek a következtetésnek, a kiterjesztett elme tézis kritikusaiknak találniuk kell egy értelmes különbséget Alice és Barry között. Történetesen ezt nehéz megtenni. Az egyik nyilvánvaló különbség az, hogy Barry "tudása" a koponyáján kívül van, és papírra van kódolva. Alice-é a koponyájában van, és neuronokba van kódolva. De ennek a különbségnek a jelentőségét éppen a kiterjesztett elme tézise vonja kétségbe. Erre támaszkodni a tézis megdöntése érdekében kérdésfeltevés lenne. Egy másik lehetséges különbség az, hogy Alice talán gyorsabban emlékszik az információra, míg Barrynek időt kell szánnia arra, hogy utánanézzon a naplójában. De ez a különbség, bár nem kérdéses, de nem is értelmes. Hogy lássuk, miért, nézzük meg, mit jelentene ez Alice számára. Mi lenne, ha minden egyes kanyarban el kellene gondolkodnia (hosszabb ideig, mint ameddig Barrynek tartott elolvasni a naplóját), mielőtt felidézné, hogy merre menjen? Ez nyilvánvalóan nem ásná alá azt az állítását, hogy ismeri az irányt. Talán az a lényeges különbség Alice és Barry között, hogy Barry elveszítheti a naplóját útközben.¹³⁷ De Alice is elveszítheti az emlékeit: egy leeső faág vagy egy agyvérzés megzavarhatja törekeny idegi kapcsolatait. Az Alice és Barry közötti lényeges különbség hiányában a kiterjesztett elme tézis hívei megerősítik, hogy Barrynek ismernie kellett az útvonalat. Az ő elméje kiterjed a napló lapjaira, amelyekre az útbaigazítás van írva.

A kiterjesztett elme teoretikusai az egyedi esetekre vonatkozó ad hoc érveken túlmutatva általános kritériumokat javasolnak annak értékelésére, hogy mikor számít egy személy külsőleg tárolt információ ismeretének. Ezek a kritériumok a paradigmaticus esetekben a személy és az általa ismert információ közötti funkcionális kapcsolatot hivatottak megragadni. Feltételezve, hogy egy személy az információt a viselkedése irányítására használja,¹³⁸ a legáltalánosabban elfogadott kritériumok a következők.

1. az információ rendelkezésre áll, és az alany jellemzően hivatkozik rá;
2. az alany többé-kevésbé automatikusan elfogadja (azaz kész cselekedni és érvelni) az információt, amint az előkerült; és

137. *Lásd id.* 15. o. (kitérve arra a tényre, hogy egy hasonló hipotézisben a jegyzetfüzetet, amelyre emlékeket írnak vagy tárolnak, el lehet venni). *Lásd általában* Fred Adams & Kenneth Aizawa, *The Bounds of Cognition*, 14 PHIL. PSYCHOL. 43, 55-56, 62-63 (2001) (megjegyezve az útbaigazítások memorizálása és az írásos útbaigazítások követése közötti kognitív folyamatok közötti különbségeket); Robert D. Rupert, *Representation in Extended Cognitive Systems: Does the Scaffolding of Language Extend the Mind?*, in *THE EXTENDED MIND*, *supra* note 109, at 325, 325 (azzal érvelve, hogy "a nyelv külső darabjai nem válnak [a kognitív] rendszer részévé").

138. A kiterjesztett elme teoretikusai ezt 'oksági összekapcsolódásnak' nevezik, és kritériumokat határoznak meg hozzá. Menary, *supra*
131. megjegyzés, 3-4. pont.

2020]

KITERJESZTETT VÁLLALATI ELME915

3. az alany könnyen hozzáférhet az információkhoz.¹³⁹

Mind Alice, mind Barry megfelel ezeknek a feltételeknek az irányinformációk tekintetében. Általánosabban, bárki, aki információt használ a viselkedése irányítására, és megfelel a feltételeknek, úgy tekintendő, mint aki ismeri azt, függetlenül attól, hogy hol és hogyan tárolja az információt: neuronokon, naplókban, tetoválásokon, rolodexekben, mobiltelefonokon, laptopokon, bárhol.

Míg az utolsó néhány bekezdésben a filozófiai érvekre összpontosítottunk, sok kognitív tudós is támogatja a kiterjesztett elme tézisének. A kiterjesztett kognitív rendszereknek több típusát ismerik el.¹⁴⁰ E rendszerek némelyikében az egyén külső tárgyakat használ. Például egy személy ujjakat vagy kavicsokat használhat segítségül a hosszú távú számoláshoz, amikor több számmal zsonglőrködni a memóriában nehéznek bizonyul.¹⁴¹ Más kiterjesztett kognitív rendszerek több személyből állnak, pl. navigációs csapatok,¹⁴² nagyszabású tudományos kutatás,¹⁴³ és tranzaktív memóriarendszerek.¹⁴⁴ A kognitív tudósok a kiterjesztett elme tézisének tükröződő témákból merítenek magyarázatot a helyhez kötött megismerés, a¹⁴⁵ robotika, a¹⁴⁶ és a gyermeki fejlődés jelenségeire.¹⁴⁷ A megtestesült megismerés elmélete adja a keretet ehhez a perspektívához: "A megismerés számos jellemzője mélyen függ a fizikai test jellemzőitől.

139.

Andy Clark, *Mementos Revenge: The Extended Mind, Extended*, in THE EXTENDED MIND, 109. lábjegyzet, 43. és 46. pont.

140. Azt is meg kell jegyeznem, hogy ez a nézet korántsem vitathatatlan a kognitív tudományban. *Lásd pl.*, ROBERT D. RUPERT, COGNITIVE SYSTEMS AND THE EXTENDED MIND 61 (2009) (a kiterjesztett megismerés megvalósításán alapuló érvelésének kritikája); Fred Adams & Kenneth Aizawa, *Why the Mind is Still in the Head*, in THE CAMBRIDGE HANDBOOK OF SITUATED COGNITION 78, 78 (Philip Robbins & Murat Aydede szerk., 2009) (a kiterjesztett elme elmélettel szemben érvelve és kijelentve, hogy "az elme még mindig a fejben van").

141. *Lásd általában* MERLIN DONALD, ORIGINS OF THE MODERN MIND (1991) (a vizuális szimbolizmus és a külső memória megjelenését, mint az ember kognitív képességei szempontjából jelentős evolúciós átmenetet vizsgálja).

142. *Lásd* EDWIN HUTCHINS, COGNITION IN THE WILD 26 (1995) (a navigációs csapatokról); *lásd még* Edwin Hutchins, *The Social Organization of Distributed Cognition*, in PERSPECTIVES ON SOCIALLY SHARED COGNITION 283, 305-06 (Lauren B. Resnik et al., szerk., 1991) (a csoportos kognitív hálózatokról és tulajdonságokról).

143. *Lásd* Ronald N. Giere & Barton Moffatt, *Distributed Cognition: Ahol a kognitív és a szociális összeolvad*, 33 SOC. STUD. SCI. 301, 301-03 (2003).

144. A tranzakciós memória olyan rendszer, amelyben egy csoport ember kódolja, tárolja és lekérdezi az adatokat. *Lásd* Daniel M. Wegner, *A Computer Network Model of Human Transactive Memory*, 13 SOC. COGNITION 319, 319-20 (1995).

145. *Lásd általában*: LUCY A. SUCHMAN, TERVEK ÉS SZITUÁCIÓS AKTÁCIÓK: THE PROBLEMS OF HUMAN MACHINE COMMUNICATION (1987) (az emberi cselekvés társadalmi és fizikai körülményekkel való kapcsolatának vizsgálata).

146. *Lásd általában* RANDALL D. BEER, INTELLIGENCE AS ADAPTIVE BEHAVIOR: AN EXPERIMENT IN COMPUTATIONAL NEUROETHOLOGY 1-18 (B. Chandrasekaran szerk., 1990) (az intelligencia különböző elméleteinek és az adaptív intelligencia gondolatának tárgyalása).

147. *Lásd általában* ESTHER THELEN & LINDA B. SMITH, A DYNAMIC SYSTEMS APPROACH TO THE DEVELOPMENT OF COGNITION AND ACTION (1996) (ebook) (a kisgyermek kognitív képességeinek fejlődését tárgyalja).

. . . [amelyek] jelentős ok-okozati szerepet játszanak ... az ágens kognitív feldolgozásában.¹⁴⁸ Az alapgondolat az, hogy az elmék nem annyira a "gondolkodás"eszközei, hanem inkább a világban való cselekvés eszközei.¹⁴⁹ Innen már csak egy rövid lépés a kiterjesztett elme tézis.¹⁵⁰ Az elme eszközként való konceptualizálása természetesen alkalmas a környezeti jeleket és a végrehajtott feladatokat összekapcsoló funkcionális jellemzésre. Azok a rendszerek, amelyek a meghatározott környezeti jelzések jelenlétében segítenek nekünk e feladatok elvégzésében, elménk részének minősülnek.

IV. A VÁLLALATI TUDAT KITERJESZTÉSE

A kiterjesztett elme tézise azt sugallja, hogy a mentális állapotok helyét meghatározó hagyományos határok túlságosan korlátozóak. Az egyes emberekkel összefüggésben a tézis azt jelenti, hogy az elme nem korlátozódik az agyra: kiterjedhet az olyan külső kognitív segédeszközökre is, mint a naplók és a mobiltelefonok. Ha a tézis átvihető a vállalati kontextusra, akkor a vállalati elme is túlmutathat a hagyományos határain - az egyes alkalmazottak elméjén -, és más funkcionálisan integrált vállalati rendszereket is magában foglalhat.

Bár a kiterjesztett elme szószólói eddig csak a természetes emberekről beszéltek, a¹⁵¹ párhuzamos érvek a vállalati emberekre is vonatkozhatnak. A kiinduló feltételezés az, hogy a vállalatoknak van elméjük. A törvény gondoskodik erről az előfeltevéstről azzal, hogy arra utasít bennünket, hogy feltételezzük, hogy a vállalatok emberek.¹⁵² A vállalati mentális állapotok - a törvény jelenleg azt mondja nekünk - az alkalmazottak fejében lakoznak. A mentális állapotok funkcionális felfogása azt jelenti, hogy az alkalmazottakkal azonos funkcionális szerepeket betöltő rendszerek szintén a vállalati elme részét képezhetik.¹⁵³ Az előző szakaszban idézett paritási premissza könnyen adaptálható a

148. Robert A. Wilson & Lucia Foglia, *Embodied Cognition*, STAN. ENCYCLOPEDIA PHIL. (2015. dec. 8.), <https://plato.stanford.edu/entries/embodied-cognition/> [<https://perma.cc/X9F5-QG5E>].

149. *Lásd* ANDY CLARK, BEING THERE: PUTTING BRAIN, BODY, AND WORLD TOGETHER AGAIN 196 (1997) (ebook).

150. *Lásd általában* Rupert, *Supra* 137. lábjegyzet (a kognitív feldolgozás és a kiterjesztett elme tézisének tárgyalása).

151. Egy lenyűgöző cikk azt állítja, hogy a csoportos elmék az egyéni elméknek más egyéni elmékre való kiterjesztéséből alakulhatnak ki. Deborah Perron Tollefsen, *From Extended Mind to Collective Mind*, 7 COGNITIVE SYS. RES., 2006, 140, 140-41. o. (2006). Minézet szerint a csoportos elme a kiterjesztés eredménye, nem pedig (ahogyan én itt javaslom) az elme, amelyet kiterjesztettek. *Lásd id.* 146. o. A mesterséges intelligenciát kifejezetten kizárja a nézetéből. *Lásd id.* 141. o. (kijelentve, hogy cikke az "elsősorban emberek által alkotott" kollektív rendszerekre összpontosít).

152. *Lásd* United States v. A & P Trucking Co., 358 U. S. 121, 123, 125 (1958) ("[I]t elemi, hogy az ilyen személytelen szervezetek a *felelős felettes* tan alapján bűnök lehetnek a szabályozási törvények tudatos vagy 'szándékos' megsértésében. ").

153. A Nemzeti Közúti Közlekedésbiztonsági Hivatal nyitottnak tűnik egy hasonló funkcionális érvelésre. *Lásd* Natl Highway Traffic Safety Admin., Opinion Letter on Applicability of Federal Motor Vehicle Safety Standards to Google's Self-Driving Vehicles (2016. február 4.), <https://research.nhtsa.gov/files/Google%20-%20compiled%20response%20to%2012%20Nov%20%2015%20interp%20request%20-%204%20Feb%2016%20final.htm> [<https://perma.cc/7MQF-YD95>] ("Ha a járműben nem tartózkodó ember vezetheti ténylegesen a járművet, akkor ésszerűbb a sofőr útgy azonosítani, mint bármi (szemben azzal, aki vezeti). ").

2020]

KITERJESZTETT VÁLLALATI ELME⁹¹⁷

vállalati kontextus: ha a világ egy része olyan vállalati folyamatként működik, amelyet, ha azt egy munkavállaló végezné, a jog a vállalati tudat részeként fogadna el, akkor a világnak ez a része a vállalati tudat része.¹⁵⁴ Lényeges, hogy a világnak ez a "része" lehet egy intelligens algoritmus, amely a vállalati műveleteket végzi. A vállalati elme ilyen kiterjesztése az emberi alkalmazottakról az automatizált algoritmusokra sok informatikusnak teljesen természetesnek tűnik. A mesterséges intelligencia klasszikus definíciója funkcionális: minden olyan algoritmus, amely "olyan viselkedésre készítet egy gépet, amelyet intelligensnek neveznénk, ha egy ember viselkedne így."¹⁵⁵

A kiterjesztett elme tézis melletti érvelés a vállalatok esetében még egyszerűbb lehet, mint az emberek esetében. A vállalati elmék, akár csak maguk a vállalatok, társadalmilag konstruált objektumok.¹⁵⁶ Kizárólag a¹⁵⁷ jogi végzésnek köszönhető, hogy a vállalatoknak egyáltalán van elméjük, vagy hogy osztoznak az alkalmazottaik mentális állapotában.¹⁵⁸ Ezzel elhárulnak a kiterjesztett elme tézisének legfőbb intuitív akadályai. Az emberi lények esetében természetes alternatívája van annak, hogy azt mondjuk, az elme magában foglalja az összes funkcionálisan integrált kognitív segédeszközt. Ez az alternatíva a józan biológián alapul - az elmét az agy korlátozza.¹⁵⁹ A vállalati elme esetében azonban sokkal nehezebb egy hihető korlátozó elvet találni.¹⁶⁰ A vállalatoknak nincs saját agyuk.¹⁶¹

E rész hátralévő része azt tárgyalja, hogyan lehet a kiterjesztett elme tézisént integrálni a társasági felelősség jogába. Két beszámolót követően a

154. *Lásd* Clark & Chalmers, *Supra* 43. lábjegyzet, 8. pont.

155. John McCarthy et al., Proposal for the Dartmouth Summer Research Project on Artificial Intelligence 7 (Aug. 31, 1955) (kiadatlan kézirat), <http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf> [<https://perma.cc/H9L6-S8D7>]. *Lásd általában* A. M. Turing, *Computing Machinery and Intelligence*, 59 MIND 433 (1950) (tágra tárgyalja, hogy a gépek képesek-e gondolkodni).

156. *Lásd* Trs. of Dartmouth Coll. v. Woodward, 17 U. S. (4 Wheat.) 518, 636 (1819) ('A corporation is an artificial being, invisible, intangible, and existing only in contemplation of law.').

157. És talán némi pszichológiai kivétel a részünkről. *Lásd* Diamantis, *Corporate Criminal*, 35. lábjegyzet, 2077-80. o.

158. A különböző alternatív megközelítések, amelyeket a törvények a vállalati elme meghatározására alkalmaznak, a respondeat superior metafizikai önkényességét tükrözik. *Lásd a Crim. Code Act of 1995* (Cth) div 12.3(2)(c) (Austl.) (a vállalati ethosz megközelítést ismerteti); *United States v. Bank of New England, N.A.*, 821 F.2d 844, 853, 856 (1st Cir. 1987) (a kollektív tudás doktrínáját ismerteti); MODEL PENAL CODE § 2.07(1)(c) (AM. LAW INST., Proposed Official Draft 1962) (kifejtve, hogy egy társaság elítélhető az igazgatótanács vagy egy magas rangú vezető tisztségviselő munkakörében eljáró döntései vagy intézkedései alapján).

159. Ennek a kihívásnak egy kifinomult változata az úgynevezett "kapcsolási-alkotmányos tévedés". *Lásd* FREDERICK ADAMS & KENNETH AIZAWA, *THE BOUNDS OF COGNITION* 76-105 (2010) (ebook).

160. Hasonló okokból a kiterjesztett elme tézisével szembeni ellenvetés, miszerint az emberek nem férnek hozzá közvetlenül az információkhoz a kiterjesztett rendszerekben, *lásd általában* John Preston, *The Extended Mind, The Concept of Belief, and Epistemic Credit*, in *THE EXTENDED MIND*, 109. lábjegyzet, 355. o. (a hozzáférés, a tekintély és a hit tárgyalása a kiterjesztett elme elméletével összefüggésben), nem vonatkozik a vállalatokra.

161. *Lásd* *United States v. Ladish Malting Co.*, 135 F. 3d 484, 492 (7th Cir 1998) ('A vállalatok nem az idegpályákon rögzítik a tudást; hanem irattárakban (és egyre inkább számítógépes lemezekben).').

A javasolt doktrínát illetően ez a rész olyan politikai megfontolásokat vesz sorra, amelyek pragmatikus korlátokat szabhatnak annak, hogy meddig terjedjen a vállalati elme.

A. *Tanulmányi javaslat*

Milyen feltételek mellett ismeri egy vállalat az algoritmusokba vagy big data rendszerekbe ágyazott információkat? Az évszázados jogi doktrína szerint a vállalatok ismerik az alkalmazottak agyában tárolt dolgokat. A kiterjesztett elmeelmélet keretet kínál arra, hogy ezen a hagyományos kereten túlmutatva a digitálisan tárolt információkra is kiterjedjen. Az emberi elmék tárgyalásához hasonlóan az egyes esetek értékelésére kétféle megközelítés létezik: ad hoc analógiák felállítása és általánosított kritériumok alkalmazása. Mindkét megközelítésnek nagyjából ugyanazokat az eredményeket kell hoznia, bár az egyik vagy a másik hasznosabb lehet a kontextustól függően.

1. Az analóg megközelítés

Az előző részben Alice és Barry tárgyalása jól szemlélteti az analóg megközelítést.¹⁶² Elvontan megfogalmazva, a megközelítés összehasonlít egy érdekes személyt, P_1 , aki funkcionális kapcsolatban áll bizonyos információval, I_1 , egy másik személlyel, P_2 , aki relevánsan hasonló funkcionális kapcsolatban áll bizonyos hasonló információval, I_2 . Ha egyértelmű, hogy P_2 ismeri I_2 , akkor P_1 -nak ismernie kell I_1 .¹⁶³ Ez a mentális állapotok funkcionális szerepük alapján történő meghatározásának eredménye - azonos funkcionális kapcsolat, azonos mentális állapot.

Az analógiás megközelítés végrehajtása a vállalati emberek esetében kissé másképp működik, mint a természetes emberek esetében. Ha P_1 egy természetes személy (mint Barry), akkor a legkönnyebben összehasonlítható P_2 egyértelmű ismeretekkel rendelkező személy, aki (mint Alice) az információt az agyában tárolta. Ha azonban P_1 egy vállalati személy, akinek digitálisan tárolt I_1 van, az analóg megközelítés másfajta P_2 összehasonlítható igényel. A respondeat superior szerint a vállalati tudás egyértelmű esete az, amikor az információt az *alkalmazott* agyában tárolják. A P_2 tehát egy olyan vállalatnak kell lennie, amely hasonlóan viselkedik, mint a P_1 , de ahol egyértelmű, hogy egy alkalmazott ismeri az I_2 -t. Az analógia sikere két tényezőtől függ:

(1) a P_1 és I_1 , valamint a P_2 és I_2 közötti funkcionális kapcsolatok relevánsan hasonlóak-e, és (2) mennyire nyilvánvaló, hogy P_2 (azaz a P_2 egyik alkalmazottja) ismerté I_2 .¹⁶⁴

162. *Lásd fentebb a III. részt.*

163. A gondatlanságot tárgyaló egyes tudósok különböző normákat javasoltak a munkavállalókra és a mesterséges intelligenciára, például azt, hogy a mesterséges intelligenciára vonatkozó gondatlansági normáknak kétszer olyan szigorúnak kell lenniük, mint az emberi egyénekre vonatkozó normáknak. *Lásd Geistfeld, Supra 17. lábjegyzet, 1679. o.* (érvelése szerint az automatizált autóknek 'legalább kétszer olyan biztonságosnak' kell lenniük, mint az ember által működtetett járműveknek, mielőtt forgalomba helyezik őket).

164. Az elemzés feltételezi, hogy van egy működőképes elméletünk arra vonatkozóan, hogy a mesterséges intelligencia viselkedése mikor tulajdonítható a vállalatoknak. Amennyire én tudom, nincs. Több lehetőség is van. A hozzárendelés függhet attól, hogy a vállalat tulajdonában van-e az adott szoftver. Vagy attól, hogy a vállalat tulajdonában van-e a szoftvert futtató hardver. Vagy attól, hogy a vállalat utólag jóváhagyta-e a viselkedést. Míg én azt mondanám, hogy

2020]

KITERJESZTETT VÁLLALATI ELME919

Az analóg megközelítés könnyen adaptálható a tárgyalás ténymegállapítási folyamatára. A hipotetikus és összehasonlító érvelés már most is alapvető jellemzője az ítélkezési folyamatnak.¹⁶⁵ Valóban, más tudósok hangsúlyozták az összehasonlító érvelés fontosságát a vállalati mentális állapotok értékelésében.¹⁶⁶ Az ilyen érvelés során a felperes vagy a vádló feladata az összehasonlító eset felépítése. A védelem feladata a hasonlóságuk megkérdőjelezése. A ténymegállapító feladata pedig az összehasonlítás meggyőző erejének eldöntése.

Hangsúlyozni kell, hogy az analóg megközelítést alkalmazó polgári felpereseknek és büntetőügyészeknek továbbra is meg kell felelniük a vonatkozó bizonyítási tehernek: a bizonyítékok túlnyomó többségének¹⁶⁷ és az ésszerű kétséget kizáróan¹⁶⁸. Nem lenne elég, ha az összehasonlító társaságban, P_2 , egyszerűen megállapítanák, hogy egy alkalmazott tudott I_2 . Ez csak azt bizonyítaná, hogy *van egy* hasonló P_2 , aki tudott I_2 . A legerősebb következtetés, amelyet a ténymegállapítók egy ilyen esetből levonhatnának, az a pusztán *lehetőség*, hogy a társasági alperes, P_1 , tudott I_1 . Stratégiaileg a felpereseknek és az ügyészeknek ehelyett egy olyan P_2 -ot kellene bemutatniuk, amely hozzáféréssel rendelkezett a vonatkozó információkhoz, hasonlóan viselkedett, mint P_1 , és ezt csak alkalmazottak segítségével tette. A felpereseknek és az ügyészeknek nem kellene kikötniük, hogy a P_2 egyik alkalmazottja ténylegesen ismerte az információt. A felelősségre vonás bizonyítási terheinek teljesítéséhez a ténymegállapítóknak arra kellene következtetniük, hogy a P_2 (azaz a P_2 egy alkalmazottja) nagy valószínűséggel (polgári jog) ismerte, vagy (büntetőjog) ismernie kellett a vonatkozó információt. Ez nem csak azt mutatná, hogy *van egy* analóg P_2 , amely ismerte az információt, hanem azt is, hogy *bármelyik* analóg P_2 valószínűleg ismerte. Ebből az következne, hogy P_1 valószínűleg szintén tudta.

Formálisan megfogalmazva ez az érvelési stílus bonyolultnak tűnhet, de valójában egy intuitív folyamatról van szó, amelyet a bírák és az esküdtek folyamatosan alkalmaznak. Ennek során annak valószínűségét kell értékelni, hogy egy bizonyos tény igaz marad-e egy olyan hipotetikus esetben, amelynek tényei hasonlóak a tényleges esethez. A polgári jogban a *res ipsa loquitur* érvek - azaz, hogy a szóban forgó baleset nem szokott gondatlanság nélkül bekövetkezni¹⁶⁹ - azonos logikai felépítésűek. A felperes gyakorlatilag azzal érvel, hogy bármely relevánsan hasonló esetben a balesetet okozó személy, aki a balesetet okozta.

próbaképpen azt javasolják, hogy a releváns változó a vállalat által a mesterséges intelligencia felett gyakorolt ellenőrzés szintje legyen, az olvasóknak a vállalati viselkedés saját preferált elmélete szerint kell eljárniuk.

165. *Lásd általában* California v. Carney, 471 U.S. 386, 389, 393 (1985) (összehasonlító érvelés); Scott Brewer, *Exemplary Reasoning: Semantics, Pragmatics, and the Rational Force of Legal Argument by Analogy*, 109 HARV. L. REV. 923, 925-26 (1996) (kifejtve, hogy az analógia a jogi érvelésben gyakran alkalmazott technika).

166. *Lásd pl.* William S. Laufer, *Corporate Bodies and Guilty Minds*, 43 EMORY L.J. 647, 701, 704 (1994) [a továbbiakban Laufer, *Corporate Bodies*].

167. *Lásd* Herman & MacLean v. Huddleston, 459 U. S. 375, 387 (1983) (Egy tipikus pénzbeli kártérítési perben a felperesnek a bizonyítékok túlnyomó többségével kell bizonyítania az ügyét.).

168. *Lásd In re* Winship, 397 U.S. 358, 361 (1970) (a büntetőügyekben szükséges bizonyítási szintet tárgyalva).

169. *Res Ipsa Loquitur*, BLACKS LAW DICTIONARY, 61. lábjegyzet.

valószínűleg gondatlanul viselkedett.¹⁷⁰ Hasonló érvelés jelenik meg a társasági büntetőjogban. A vállalati ügyészek bebizonyíthatják, hogy a kötelességszegés körülményei alapján valamelyik alkalmazottnak vétkes tudással kellett rendelkeznie, még akkor is, ha az ügyész nem tudja bizonyítani, hogy ki volt az.¹⁷¹ Ennek az érvelésnek a felépítése megköveteli annak bizonyítását, hogy bármely relevánsan hasonló esetben egy alkalmazottnak vétkes tudással kellett rendelkeznie.

Nézzük meg, hogyan működne az analóg megközelítés a HealthCo hipotetikus esetben. Annak bizonyítására, hogy a HealthCo tudta, hogy a nyomtatványok hamisak, a vádnak egy hipotetikus összehasonlító esetet kellene bemutatnia, amelynek három alapvető jellemzője van:

1. A vállalat hasonlóan viselkedett (azaz hamis visszatérítési nyomtatványokat nyújtott be); és
2. a vállalat ezt alkalmazottak segítségével tette (azaz nem egy olyan algoritmus segítségével, mint a FormBot);
3. akik hozzáférhettek (anélkül, hogy kikötötték volna, hogy tudtak) hasonló információkhoz (azaz azokhoz az információkhoz, amelyeknek szerepelniük kellett volna a nyomtatványokon).

A vád érvelésének erőssége az összehasonlító eset hitelességétől, a vádlott és a vállalati vádlott közötti funkcionális hasonlóságának közelségétől, valamint attól függ, hogy mennyire valószínű, hogy az összehasonlító eset valamely alkalmazottnak ismernie kellett az információt.

Mint oly sok minden a tényfeltárási folyamatban, az elemzés a legtöbb esetben nem lesz egyértelmű. A HealthCo feltételezett ügyében több tényre lenne szükség, mielőtt a határozat kikristályosodna. Kiderülhet például, hogy a HealthCo nyomtatványain szereplő téves állítások mind viszonylag kis jelentőségűek voltak. Azokban a hipotetikus összehasonlító esetekben, amikor egy alkalmazott töltötte ki a nyomtatványokat, ez összhangban lehet azzal, hogy az alkalmazott inkább kerekítési hibákat követett el, mint hogy olyan nyomtatványokat töltött volna ki, amelyekről tudta, hogy hamisak. Az is kiderülhet azonban, hogy a kerekítési hibák mindig a HealthCo javára történtek. Ez az alkalmazott által végrehajtott programszerű magatartás inkább megfelelne a tudatos hamisításnak.

2. Általánosított kritériumok használata

Bizonyos helyzetekben hasznos lehet, ha általános kritériumokkal rendelkezünk annak értékelésére, hogy egy vállalat ismeri-e az algoritmusába ágyazott információkat. A kritériumoknak jellemezniük kell a funkcionális kapcsolatot

170. Egy másik polgári jogi példa a termékfelelősségből származik, ahol a ténymegállapítók a termék hibájára következtethetnek a termék olyan teljesítményéből, amely "olyan jellegű volt, amely általában termékhiba következtében fordul elő". "RESTATEMENT (THIRD) OF TORTS: PRODS. LIABILITY § 3(a) (AM. LAW INST. 1997).

171. Lásd: *Fejlemények a jogban - Vállalati bűnözés: A vállalati magatartás büntetőjogi szankciókkal történő szabályozása*, 92 HARV. L. REV. 1227, 1248 (1979) (idézi a *United States v. T.I.M.E.-D.C., Inc.*, 381 F. Supp. 730, 735 n.7, 739 (W.D. Va. 1974)).

2020]

*KITERJESZTETT VÁLLALATI ELME*⁹²¹

a vállalatok olyan információkhoz, amelyeket a jelenlegi jog egyértelműen úgy kezel, mintha tudnának. A *Respondeat superior* meghatározza, hogy melyek ezek a helyzetek. Ha a kritériumok pontosan megragadják ezt a funkcionális kapcsolatot, a funkcionalizmus azt diktálja, hogy minden olyan társaság, amely bizonyos információk tekintetében megfelel a kritériumoknak, az adott információ ismeretének minősül. Ez attól függetlenül igaz lenne, hogy az információt hogyan vagy hol tárolták.

A "respondeat superior" alapkövetelménye, hogy valamelyik alkalmazott ismerje az információt.¹⁷² Mivel ez az első követelmény végső soron a természetes emberek (az alkalmazottak) ismeretére vonatkozik, a kiterjesztett elme teoretikusai által az emberekre már felállított kritériumok kissé módosított változatainak kell működniük:

1. Az információ rendelkezésre áll, és a munkavállaló/algorithmus (a vállalat nevében) jellemzően lekérdezi azt;
2. az alkalmazott/algorithmus (a vállalat nevében) többé-kevésbé automatikusan jóváhagyja az információt a visszakereséskor; és
3. az alkalmazott/algorithmus (a vállalat nevében) könnyen hozzáférhet az információkhoz.

Ezek a kritériumok együttesen intuitív módon megragadják azt a funkcionális kapcsolatot, amelyet a vállalat az alkalmazottak agyában tárolt vállalati információkkal ápol - az alkalmazottak könnyen fel tudják idézni ezeket az információkat, és munkájuk elvégzéséhez használják azokat. Az "alkalmazott" kifejezés "algorithmussal" való helyettesítésével a kritériumok könnyen alkalmazkodnak annak megválaszolásához, hogy a vállalat valamelyik algoritmusán keresztül ismeri-e az információkat. A legtöbb olyan vállalati algoritmus, amely könnyen hozzáfér a vállalati műveletek irányítására használt információkhoz, megfelel a kritériumoknak.

A *respondeat superior* két további követelményt támaszt a munkavállalók tudásának a vállalatoknak való tulajdonításához: először is, hogy a munkavállaló a munkaviszonya keretében ismeri az információt, másodsor pedig, hogy azt a munkáltatója javára történő felhasználás szándékával használja fel.¹⁷³ Ezt a két követelményt azonban a bíróságok annyira meggyengítették, hogy kérdéses, hogy az ismeretekre itt kidolgozott általánosított kritériumoknak figyelembe kell-e venniük őket. A munkavállaló ténylegesen a munkaviszonya keretében végzett munkának minősül, amikor a munkahelyén tartózkodik, még akkor is, ha a munkáltatója utasításaival ellentétesen cselekszik.¹⁷⁴ És a munkavállaló akkor is úgy tekintendő, mint aki a munkáltatója javát akarja szolgálni, ha ez a szándék másodlagos,¹⁷⁵ hipotetikus,¹⁷⁶ vagy hatástalan.¹⁷⁷

172. *Lásd fentebb az I. részt.*

173. *Lásd* 18B AM. JUR. 2D *Corporations* §§ 1812, 1817, Westlaw (2020. februárban frissített adatbázis).

174. *Lásd pl.* *United States v. Hilton Hotels Corp.*, 467 F.2d 1000, 1004 (9th Cir. 1972).

175. *Lásd* *United States v. Automated Med. Labs., Inc.*, 770 F.2d 399, 407 (4th Cir. 1985).

176. *Lásd* *United States v. Sun-Diamond Growers of Cal.*, 138 F.3d 961, 970 (D.C. Cir. 1998), *megegerősítve*, 526 U.S. 398 (1999).

177. *Lásd* *Old Monastery Co. kontra Egyesült Államok*, 147 F.2d 905, 908 (4th Cir. 1945).

A "respondeat superior" munkaviszonyra vonatkozó és a haszonszerzési szándékra vonatkozó követelményeinek intuitív célja az, hogy a vállalatok két esetben megússzák a felelősségre vonást: amikor egy alkalmazott csak magánjelleggel tud valamit ("Apámnak valójában nem volt ilyen mütétje."), vagy amikor a munkavállaló a tudását csak arra használja fel, hogy meghiúsítsa a vállalati célokat ("Így tudnám elsikkasztani a vállalati vagyont."). Mivel az algoritmusoknak nincs magánéletük, az első forgatókönyv irreleváns. Az algoritmusok azonban előmozdíthatják vagy akadályozhatják a vállalati célokat. Egy elromlott algoritmus éppúgy áldozatul eshet egy vállalatnak, mint a csaló alkalmazottak.¹⁷⁸ A következő negyedik kritériumnak meg kell felelnie a respondeat superior szándékos haszonszerzésre vonatkozó követelménye mögött álló érdekeknek:

4. Az algoritmusnak továbbá úgy kell felhasználnia az információt, hogy az a vállalat számára valamilyen (esetleg jogtalan vagy csekély) előnyt jelentsen.

A törvény jelenlegi, a haszonszerzési szándékra vonatkozó követelményéhez hasonlóan az esetek túlnyomó többsége könnyedén megfelel ennek a negyedik kritériumnak.

Hogyan vonatkoznának a kritériumok a HealthCo esetében? A HealthCo formanyomtatvány-bejelentő algoritmus, a FormBot egyértelműen megfelel a negyedik kritériumnak (a vállalat haszna), mivel a hamisított formanyomtatványok több millió dollárt hoztak a HealthCo számára. A többi kritérium értékeléséhez több technikai részletre lenne szükség arról, hogy a FormBot hogyan fér hozzá az információkhoz és hogyan használja fel azokat. Valószínűleg feltételezhető, hogy a FormBot megfelel az első (jellemzően az információra való hivatkozás) és a harmadik (az információ automatikus jóváhagyása) kritériumnak. Nincs ok arra, hogy a HealthCo mérnökei a FormBotot úgy tervezzék meg, hogy az akadályozottan férjen hozzá a követelésekkel kapcsolatos információkhoz, vagy hogy a FormBot jellemzően ne hívja fel ezeket az információkat a követelések benyújtásakor. A második kritérium (az információ automatikus jóváhagyása) tekintetében nagyobb a mozgástér. A mérnökök úgy is tervezhetnék volna a FormBotot, hogy kritikátlanul elfogadja és felhasználja a kapott kárigényinformációkat. Ez megerősítené azt az állítást, hogy a FormBot teljesíti a második kritériumot, és következésképpen azt, hogy a HealthCo ismerte az információt. Alternatív megoldásként a HealthCo mérnökei tervezhetnék volna a FormBotot úgy is, hogy szkeptikusabb legyen, beépített ellenőrzési rendszerrel, amely az információk felhasználása előtt ellenőrzi és validálja azokat. Ebben az esetben a FormBot kevésbé valószínű, hogy teljesíti a második kritériumot, és a HealthCo kevésbé valószínű, hogy úgy tekintené, hogy ismeri az információt.

178. *Lásd Sun-Diamond*, 138 F.3d at 970 (megjegyezve, hogy a munkavállaló inkább a munkáltatóját akarta becsapni, mintsem hasznot húzni belőle). Ha egy algoritmus helytelenül diszkriminál a hitel- vagy álláskeresőkkal szemben, az a vállalat is veszít, amelynek nevében dolgozik, mert elesik a jövedelmező hitelezési lehetőségektől vagy a minőségi jelöltektől. *Lásd* Margareta Drzeniek-Hanouz, *Why Discrimination Is Bad for Business*, WORLD ECON. F. (2015. március 6.), <https://www.weforum.org/agenda/2015/03/why-discrimination-hurts-competitiveness/> [<https://perma.cc/H7R9-BH6F>].

2020]

KITERJESZTETT VÁLLALATI ELME923

Érdemes röviden megállni, hogy megemlítsük, hogy a négy kritérium hogyan hoz vonzó politikai eredményeket. Működőképes és elméletileg megalapozott alapként szolgálnak a kiterjesztett elméletnek a társasági jogba való bevezetéséhez. Mint ilyenek, lehetőséget kínálnak arra, hogy a HealthCo-hoz hasonló vállalatokat felelősségre lehessen vonni, ha algoritmusai megszegik a törvényt. A kritériumok nem írják elő automatikusan felelősséget minden olyan esetben, amikor egy vállalati algoritmus kárt okoz. Van egy olyan feltétel, amely társadalmilag előnyös ösztönzőket ad a vállalatoknak. Például az ex ante perspektívából nézve a kritériumok lehetővé tették volna a HealthCo számára, hogy csökkentse a felelősségre vonás lehetőségét azáltal, hogy a FormBotba további minőségellenőrzéseket épít be. A felelős algoritmusfejlesztés ösztönzése pontosan az, amit a jognak tennie kellene. A cikk most erre és más szakpolitikai megfontolásokra tér ki.

B. Szakpolitikai alapú kifogások és korlátozások

A vállalati elme kiterjesztése az imént ismertetett két megközelítés valamelyikének alkalmazásával jó első lépés lenne az algoritmikus vállalati helytelen viselkedés problémájának megoldására. Bár a javaslatok biztos előrelépést jelentenek a jelenlegi joghoz képest - amely számos algoritmikus károkozás esetén hatékonyan védi a vállalatokat a felelőség alól -, a további finomítások még jobban előmozdíthatják a társasági jog céljait. Az alábbi vita a büntetőjogi politikára összpontosít; a kapcsolódó megfontolások polgári jogi kontextusban is felmerülnek.

A vállalati büntetőjog legfontosabb politikai céljai ismerősek a büntetőjogból szélesebb értelemben, nevezetesen a¹⁷⁹, azaz a megtorlás¹⁸⁰ és az elrettentés.¹⁸¹ A megtorlás kezdetben furcsának tűnhet a társasági büntetőjogban, mivel a társaságok nem hétköznapi erkölcsi szereplők.¹⁸² A megtorló érzelmek azonban a társasági büntetőjog egyik erős mozgatórugója.¹⁸³ A megtorlás elméletének különböző változatai léteznek. A társasági büntetőjoghoz legjobban illeszkedő változat a büntetőjogi felelőséget arra kívánja felhasználni, hogy igazolja a közvélemény intuícióit azzal kapcsolatban, hogy a társaságok mikor érdemelnek erkölcsi elítélést.¹⁸⁴ Még ha a vállalatok nem is valódi erkölcsi szereplők, szociálpszichológiánkban olyan helyet foglalnak el, amely természetessé és ellenállhatatlanná teszi a róluk alkotott erkölcsi ítéleteket.¹⁸⁵ Jól szemlélteti ezt az egyértelmű

179. Lásd Meir Dan-Cohen, *Sanctioning Corporations*, 19 J. L. & POLY 15, 22 (2010) (a politikai célok felsorolása).

180. Regina A. Robson, *Bűn és büntetés: Rehabilitating Retribution as a Justification for Organizational Criminal Liability*, 47 AM. BUS. L.J. 109, 110 (2010).

181. Daryl K. Brown, *Street Crime, Corporate Crime, and the Contingency of Criminal Liability*, 149 U. PA. L. REV. 1295, 1325 (2001).

182. Alschuler, *Two Ways To Think*, *supra* 68. lábjegyzet, 1392. o. (kiemelve azt a sajátosságot, hogy egy vállalatot vagy képzeletbeli személyt úgy büntetnek, mintha valódi személy lenne); Baker, *supra* 67. lábjegyzet, 350. o. (lásd: Baker, *supra* 67. lábjegyzet, 350. o.).

183. Lásd Baer, *Supra* 58, 621. o. (a társasági büntetőjog retributivista érzelmeinek megvitatása).

184. Lásd PAUL ROBINSON, INTUITIONS OF JUSTICE AND THE UTILITY OF DESERT 176-88 (2013); Dan M. Kahan & Martha C. Nussbaum, *Two Conceptions of Emotion in Criminal Law*, 96 COLUM. L. REV. 269, 352 (1996).

185. Lásd Diamantis, *Corporate Criminal*, 35. lábjegyzet, 2077-80. o.

igazságszolgáltatásra szólítanak fel azokkal a vállalatokkal szemben, amelyek függőséget okozó opioidokat nyomtak a kétségbeesett fogyasztókra¹⁸⁶ vagy elpusztították az érzékeny környezeti élőhelyeket.¹⁸⁷ A felhívás hevedése arra figyelmezteti a büntető igazságszolgáltatási rendszert, hogy ne hallgasson süket fülekre.

Az elrettentő elmélet szerint a büntetőjogi felelősségnek a vállalati visszaélések megelőzésére kell törekednie azáltal, hogy növeli a törvénysértés költségeit.¹⁸⁸ A jog ezt úgy érheti el, hogy a szankciókkal való fenyegetés révén arra ösztönzi a vállalatokat, hogy körültekintőbben működtessék a vállalkozásaikat. Ez általában azt jelenti, hogy a vállalatok további megfelelési programokat vezetnek be, például az alkalmazottak képzésére és ellenőrzésére.¹⁸⁹ Az algoritmusok esetében az óvatosság azt jelenti, hogy olyan algoritmusokat kell tervezni, amelyek kisebb valószínűséggel szegik meg a törvényt. Bár semmi sem garantálhatja, hogy egy gépi tanuló algoritmus mindig betartja a törvényt (ahogy azt sem, hogy az alkalmazottak mindig betartják a törvényt), a¹⁹⁰ szoftvermérnökök lépéseket tehetnek annak érdekében, hogy csökkentsék az algoritmusok helytelen viselkedésének valószínűségét.¹⁹¹ Ezek a lépések a következők: az algoritmusokat író mérnökök körének diverzifikálása,¹⁹² gondosabb kezdeti programozás,¹⁹³ a gyakorlóadat-készletek körültekintőbb kiválasztása,¹⁹⁴

186. *Lásd pl.* Maia Szalavitz, *Big Pharma's Opioid Greed Was Even Worse than We Thought*, VICE (2018. szeptember 13., 15:41), https://www.vice.com/en_us/article/7xj97q/big-pharmas-opioid-greed-was-even-worse-than-we-thought [<https://perma.cc/2ZKH-VSWL>] (Mint az apokrif gyermek, aki megölte a szüleit, majd együttérzésért könyörgött, mert árvává vált, a Purdue először nyereségesen nyomott egy függőséget okozó gyógyszert, majd nyilvánvalóan még több pénzt akart keresni az általa okozott függőségek kezelésével.).).

187. *Lásd pl.* George Monbiot, *Shell Is Not a Green Savior. It's a Planetary Death Machine*, GUARDIAN (2019. június 26.), <https://www.theguardian.com/commentisfree/2019/jun/26/shell-not-green-saviour-death-machine-greenwash-oil-gas> [<https://perma.cc/JP2Y-XPHT>] ("A természetes ökoszisztémákba való befektetésének hangoztatása számomra úgy tűnik, mint egy eszköz, amellyel fenntartja a társadalmi engedélyt [sic] az életünket elpusztító gáz és olaj kitermelésére.).).

188. *Lásd* Harvey M. Silets & Susan W. Brenner, *The Demise of Rehabilitation: Sentencing Reform and the Sanctioning of Organizational Criminality*, 13 AM. J. CRIM. L. 329, 367 (1986).

189. *Lásd* Brent Fisse, *Reconstructing Corporate Criminal Law: Deterrence, Retribution, Fault, and Sanctions*, 56 S. CAL. L. REV. 1141, 1204-06 (1983).

190. *Lásd* Irwin Schwartz, *Toward Improving the Law and Policy of Corporate Criminal Liability and Sanctions*, 51 AM. CRIM. L. REV. 99, 112 (2014).

191. *Lásd általában* William D. Smart, Cindy M. Grimm & Woodrow Hartzog, *An Education Theory of Fault for Autonomous Systems*, 2017 PROC. WEROBOT 24-27, <http://people.oregonstate.edu/~smartw/library/papers/2017/werobot2017.pdf> [<https://perma.cc/8EPL-WART>]. Az algoritmusokban előforduló elfogultságok kialakulásának részletes kezelését lásd: Nizan Geslevich Packin & Yafit Lev-Artez, *Learning Algorithms and Discrimination*, in RESEARCH HANDBOOK ON THE LAW OF ARTIFICIAL INTELLIGENCE 88 (Woodrow Barfield & Ugo Pagallo eds., 2018) (ebook).

192. *Lásd* Kate Crawford, *Artificial Intelligence's White Guy Problem*, N.Y. TIMES (2016. június 25.), <https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html> [<https://perma.cc/FSSW-R5HL> (sötét archívum)].

193. *Lásd* Geistfeld, *Supra* 17. lábjegyzet, 1634-36. o. (a programozási hibák által okozott algoritmushibákról).

194. *Lásd* Barocas & Selbst, *supra* 29. lábjegyzet, 680-81. o.; Oscar H. Gandy, Jr., *Engaging in Rational Discrimination: Exploring Reasons for Placing Regulatory Constraints on Decision Support Systems*, 12 ETHICS & INFO. TECH. 29, 31 (2010) (annak tárgyalása, hogy a rossz adatok hogyan torzíthatják az automatizált rendszereket).

2020]

KITERJESZTETT VÁLLALATI ELME925

szélesebb körű bevezetés előtti tesztelés,¹⁹⁵ rendszeres bevezetés utáni minőségellenőrzések,¹⁹⁶ rutinszerű futásidejű megfelelőségi rétegek,¹⁹⁷ hatékony felügyelet,¹⁹⁸ és folyamatos szoftverfrissítések a felmerülő problémák kezelésére.¹⁹⁹ A programozóknak is vannak eszközeik, amelyekkel bebizonyíthatják (önmaguknak vagy másoknak), hogy egy algoritmus következetesen alkalmazza a szabályait.²⁰⁰ Ezen óvintézkedések mindegyike olyan költségekkel jár, amelyeket a vállalatok mindent összevetve inkább elkerülnének. A büntetőjogi felelősségre vonás a szankciók fenyegetése révén olcsóbbá teheti az elővigyázatosságot, mint a jogsértés megkockáztatása.

Mivel e cikk középpontjában a társasági felelősség hatályának kiterjesztése áll, a legsürgetőbb politikai aggály az, hogy a javaslatok nem mennek-e túl messzire. Az itt javasolt megközelítés például nem követeli meg, hogy az algoritmikus helytelen magatartáson kívül a vállalat részéről bármiféle jogellenes cselekedetet kövessen el. Ez a közvetett felelősség megtorló jelleggel nem megfelelő formájának tűnhet²⁰¹: Hogyan érdemelhet büntetést egy vállalat, ha nem követett el semmi rosszat?²⁰²

195. Lásd Geistfeld, *Supra* note 17, 1623, 1651-54; lásd még DAVE CLIFF & LINDA NORTHOP, GOV'T OFFICE FOR SCI., A GLOBÁLIS PÉNZÜGYI PIACOK: AN ULTRA-LARGE-SCALE SYSTEMS PERSPECTIVE 19-20 (2012) (a kereskedési algoritmusok szimulációkkal történő tesztelésének szükségességét tárgyalva).

196. Lásd B. Bodo et al., *Tackling Algorithmic Control Crisis-The Technical, Legal, and Ethical Challenges of Research into Algorithmic Agents*, 19 YALE J. L. & TECH. 133, 142-44 (2017) (az algoritmusok ellenőrzésének leírása); James Guszcza et al., *Why We Need To Audit Algorithms*, HARV. BUS. REV. (2018. november 28.), <https://hbr.org/2018/11/why-we-need-to-audit-algorithms>. Lásd általában Shlomit Yanisky-Ravid & Sean K. Hallisey, *Equality and Privacy by Design: A mesterséges intelligencia adatátláthatóságának új modellje az auditálás, a tanúsítás és a biztonság kikötő rendszerek révén*, 46 FORDHAM URB. L. J. 428, 434-35 (2019) (az adatok és a mesterséges intelligencia auditálásának és tanúsításának rendszerét támogatja az átláthatóság ösztönzése érdekében).

197. Lásd Louise Dennis et al., *Formal Verification of Ethical Choices in Autonomous Systems*, 77 ROBOTICS & AUTONOMOUS SYSS. 1, 1-2 (2016) (a formális verifikáció tárgyalása); Felipe Meneguzzi & Michael Luck, *Norm-Based Behavior Modification in BDI Agents*, 2009 PROC. 8TH INTL JOINT CONF. ON AUTONOMOUS AGENTS & MULTI-AGENT SYSS. 177, 177.

198. King et al., *Supra* note 81, 110-12. o. (az algoritmusok négy lehetséges felügyeleti mechanizmusát tárgyalja).

199. Lásd általában NAT'L HIGHWAY TRAFFIC SAFETY ADMIN., FEDERAL AUTOMATED VEHICLES POLICY: ACCELERATING THE NEXT REVOLUTION IN ROADWAY SAFETY 16 (2016),

<https://www.hsd1.org/?view&did=795644>[<https://perma.cc/6BWR-YTTR>](azönvezető autók gyártóinak rendszeres szoftverfrissítéseik képzelik elbiztonság javítása érdekében).

200. Kroll et al., 117. lábjegyzet, 662-72. o. (a rendelkezésre álló mechanizmusok leírása). E mechanizmusok nélkül az utólagos ellenőrzés számos technikai és jogi okból nehézségekbe ütközhet. Lásd általában Amanda Lewindowski, *How Copyright Law Can Fix Artificial Intelligence's Implicit Bias Problem*, 93 WASH. L. REV. 579, 589-90 (2018) (annak vizsgálata, hogy a tisztességes felhasználás doktrína lehetővé teheti-e a mesterséges intelligencia alkotóinak, hogy elkerüljék a szerzői jog által okozott következtelenségeket).

201. Az Association for Computing Machinery valami hasonlót javasolt az egyének számára, nevezetesen, hogy az általuk használt algoritmusok által hozott döntésekért felelősségre kell vonni őket. ASSN FOR COMPUTING MACH., U. S. PUB. POLICY COUNCIL, STATEMENT ON ALGORITHMIC TRANSPARENCY AND ACCOUNTABILITY 2 (2017), https://www.acm.org/binaries/content/assets/public-policy/2017_usacm_statement_algorithms.pdf [<https://perma.cc/YVT6-5GSK>].

202. Lásd Lemley & Casey, *Supra* 17. lábjegyzet, 1313-15. o. (hasonló kérdés vizsgálata).

1. Felelősség a hibás algoritmusokért

Ez a kihívás nem egyedülálló a kiterjesztett elmeelmélet és az algoritmikus félvezetés esetében. Úgy tűnik, hogy minden vállalati felelősség, beleértve a jelenlegi jog szerinti büntetőjogi felelősséget is, helyettesítő jellegű. A vállalatok nem tudnak önállóan cselekedni;²⁰³ az őket működtető alkalmazottakon és (most már) algoritmusokon keresztül kell cselekedniük. A respondeat superior a vétkességet az alkalmazottokról a vállalatokra hárítja, méghozzá anélkül, hogy a vállalat részéről további vétkességet, például gondatlan felvételi gyakorlatot vagy a kötelességszegés utólagos jóváhagyását követelné meg.²⁰⁴ A fent javasolt doktrínák nem lennének ettől eltérőek. Ami a megtorló igazságtalanság vádját illeti, a felröhatósági elmélet és a kiterjesztett elmélet nevében a válasz ugyanaz: a kifogás fogalmi tévedésre támaszkodik. A válasz azt feltételezi, hogy van egy szakadék aközött, hogy a vállalat valamit rosszul csinál, és aközött, hogy az alkalmazottai vagy algoritmusai valamit rosszul csinálnak. Az alkalmazottak és az algoritmusok azonban a vállalatok részei.²⁰⁵ Tehát az alkalmazottak és az algoritmusok hibái vállalati hibák. A felelősség végül is nem igazán helyettesítő jellegű.²⁰⁶

Van egy elrettentésen alapuló logika is, amely a "respondeat superior" és a kiterjesztett vállalati elméletet is vezérli. Bár a vállalatok soha nem tudják garantálni, hogy alkalmazottaik a munkahelyükön megfelelően fognak viselkedni, a²⁰⁷ vállalatok vannak a legjobb helyzetben ahhoz, hogy csökkentsék a helytelen viselkedés kockázatát.²⁰⁸ Azzal, hogy a vállalatot büntetéssel fenyegetik, ha valamelyik alkalmazottja valami rosszat tesz, a respondeat superior arra ösztönzi a vállalatokat, hogy olyan megfeleléségi protokollokat vezessenek be, mint például²⁰⁹ további szűrés a felvételnél, jobb munkavállalói képzés, hatékonyabb munkavállalói ellenőrzés, nyitottabb jelentési csatornák és szigorúbb fegyelmi intézkedések.²¹⁰ Ugyanez igaz a vállalatok felelősségre vonására az algoritmikus visszaélésekért. A vállalatok felelősségre vonása a következőképpen

203. Laufer, *Corporate Bodies*, *supra* note 166, 652. o. ('Egy társaság csak egy ügynökön keresztül járhat el. . . .').

204. Egyes teoretikusok azt javasolták, hogy az ilyen típusú vállalati hiányosságok alternatív keretet képezhetnek a vállalati hiba megállapítására. *Lásd pl.* Fisse, *Supra* note 189, 1200. o. (a reaktív és a proaktív vétkesség tárgyalása).

205. *Lásd* N. Y. Cent. & Hudson River R. R. v. United States, 212 U. S. 481, 492-93 (1909) (megjegyezve, hogy a társaságot a tisztviselői és ügynökei alkotják).

206. *Lásd* Am. Med. Assoc. v. United States, 130 F. 2d 233, 253 (D. C. Cir. 1942) ('Ha egy vállalat bűnös egy bűncselekményben, akkor az egy vállalati cselekmény, egy vállalati szándék miatt van.

társaság csak emberi szereplők által cselekedhet, lényegtelen. '), *megeősíve*, 317 U. S. 519 (1943).

207. *Lásd* Cindy R. Alexander & Mark A. Cohen, *Why Do Corporations Become Criminals? Ownership, Hidden Actions, and Crime as an Agency Cost*, 5 J. CORP. FIN. 1, 6 (1999) ('[A] bűncselekmények előfordulása reálisan a vezetésen kívül eső számos befolyástól függ. '); Schwartz, *supra* note 190, 112. o., 112. o.

208. Jennifer Arlen, *The Failure of the Organizational Sentencing Guidelines*, 66 U. MIAMI L. REV. 321, 332-33 (2012) ('[C]orporations often are the low-cost providers of many forms of policing.').

209. *Lásd id.* 332-34.

210. *Lásd* Tanina Rostain, *General Counsel in the Age of Compliance*: GEO. J. LEGAL ETHICS 465, 466-67 (2008).

2020]

KITERJESZTETT VÁLLALATI ELME927

az általuk használt algoritmusokat, a törvény arra ösztönözheti a vállalatokat, hogy jobb munkát végezzenek algoritmusaik megtervezésében, nyomon követésében és javításában.

Aggódhatunk amiatt, hogy ugyanez az érvelés nem fog működni, ha egy vállalat - mint az gyakran előfordul - egy tapasztaltabb technológiai céget bíz meg az algoritmusok megtervezésével.²¹¹ Ebben az esetben a bűncselekmények megelőzésének legközvetlenebb módja az lehet, ha a technológiai cég ösztönzőit vesszük célba. A vállalati végfelhasználók felelősségre vonása azonban ugyanezt az eredményt érheti el. A vállalatok kétségtelenül az algoritmikus visszaélések költségeit a technológiai cégekre hárítják át kártérítési megállapodásokon keresztül, így a technológiai cégek kénytelenek internalizálni a visszaélések kockázatát.²¹² Ha ők viselik a pénzügyi kockázatot, amikor algoritmusaik rosszul viselkednek, a technológiai cégek hatékonyabb óvintézkedéseket fognak tenni termékeik tervezése és tesztelése során. Adminisztratív szempontból egyszerűbb a vállalati végfelhasználókat felelősségre vonni, mint közvetlenül a technológiai cégeket. Ez megszabadítja a bíróságokat attól a zűrös üzleti²¹³ feladattól, hogy a felelősséget a technológiai cégek (a tervezési hibákért)²¹⁴ és a vállalati végfelhasználók (a felhasználói hibákért) között osszák fel.²¹⁵

2. Mások információiért való vikárius felelősség

Van egy másik szempont is, amelyben a kiterjesztett elmélet a vállalati felelősség túlzott mértékű kiterjesztéséhez vezethet. Emlékezzünk arra, hogy a kiterjesztett tudatelmélet szerint, ha egy személy megfelelő funkcionális kapcsolatban áll bizonyos információkkal, akkor azt úgy kell tekinteni, mintha ismerné azokat, függetlenül attól, hogy hol vagy hogyan tárolják az információt. Egy digitálisan összekapcsolt világban ez végül *nagyon is* átfogó lehet. Gondoljuk meg, mit jelent ez a természetes emberek számára. Barry a naplójában lévő információt ismerőnek számít, mert könnyen használhatja azt arra, hogy megtalálja

211. *Lásd* Lemley & Casey, *Supra* 17. lábjegyzet, 1352. o. ('A robotok sok összetett alkatrészből állnak ... gyakran különböző vállalatok tervezik, működtetik, lízingelik vagy birtokolják őket.').

212. *Lásd* Pamela H. Bucy, *Vállalati etosz: A Standard for Imposing Corporate Criminal Liability*, 75 MINN. L. REV. 1095, 1146 (1991) ('[M]ost corporations use at least one form of compensation, indemnification, in a way that encourages corporate crime'); David R. Cohen & Roberta D. Anderson, *Insurance Coverage for "Cyber-Losses"*, 35 TORT & INS. L. J. 891, 926 (2000) ('igazgatók és tisztviselők [(D&O) biztosítási kötvényei fedezetet nyújtanak a vállalati kötelezettségeik teljesítésével kapcsolatban beperelt igazgatók és tisztviselők védelmi és kártalanítási költségeire. Egy tipikus D&O biztosítási kötvény biztosítást nyújt minden olyan 'vesztés'ellen, amely 'jogellenes cselekményből' származik.').

213. *Lásd* Lemley & Casey, *Supra* 17. lábjegyzet, 1352-53. o. ('robotok tervezői, tulajdonosai, üzemeltetői és felhasználói természetesen azon fognak vitatkozni, hogy ki viseli a valódi jogi felelősséget azért, hogy a robot úgy viselkedett, ahogyan viselkedett. És ezek az összetett megkülönböztetések még csak nem is számolnak a harmadik felek szerepével, akik a robotokat káros viselkedésre készítik').

214. Feltételezve, hogy a hiba a tervezési oldalon, még mindig nehéz lehet megtalálni, hogy hol. *Lásd* David C. Vladeck, *Gépek elvek nélkül*: WASH. L. REV. 117, 125 (2014) (a mesterséges intelligencia hibáiért felelős személyek meghatározásának nehézségeit tárgyalja).

215. *Lásd* MATILDA CLAUSSEN-KARLSSON, ARTIFICIAL INTELLIGENCE AND THE EXTERNAL ELEMENT OF THE CRIME 22 (2017), <http://oru.diva-portal.org/smash/get/diva2:1115160/FULLTEXT01.pdf> [<https://perma.cc/9F29-WQNR>] (a mesterséges intelligencia felhasználóinak és gyártóinak részvételét tárgyalja); McAllister, *Supra* note 123, 2550. o.; McAllister, *Supra* note 123, 2550. o.

a kávézó. De mi lenne, ha Barry ahelyett, hogy a naplójába írta volna az útvonalat, inkább a mobiltelefonján nézte volna meg, miután elindult. Feltételezve, hogy Barry még mindig megfelelő funkcionális kapcsolatban állt az információval, úgy számítana, hogy ismeri azt. Mivel a telefonján keresztül mindig könnyen hozzáférhetett az információhoz (bármikor megnézhetné volna), úgy kell tekinteni, hogy már akkor is tudta, amikor még nem vette ki a telefont a zsebéből. Az a tény, hogy az adatokat valójában egy több száz mérföldre lévő szerverfarmon tárolták, nem befolyásolná az elemzést. Ami igaz az irányadatokra, az ugyanúgy igaz lehet bármely más, Barry adatkapcsolatán keresztül elérhető információra is. Például Barry, mobiltelefonnal a kezében, úgy is számíthat, hogy mindent tud a Wikipédiáról, még a még el nem olvasott cikkeket is. Ez nagyon furcsa eredményekhez vezethet, például ahhoz, hogy amit Barry tud, az folyamatosan változik a tudta nélkül, mivel a Wikipédia szerkesztői hozzáadnak, törölnek és módosítanak tartalmakat.

Ez még a kiterjesztett elmék sok rajongója számára is túl messzinek tűnik.²¹⁶ A kiterjesztett elme tézisének korlátozására egyes teoretikusok egy további kritériumot javasolnak annak értékelésére, hogy egy alany külső információ ismeretének számít-e: az alanynak korábban jóvá kell hagynia az információt.²¹⁷ Ez gyakorlatilag az alany által megismerhető információk körét azokra az információkra korlátozná, amelyek korábban rajta keresztül (azaz az agyán keresztül) jutottak el hozzá. Ezzel az új korlátozással Barry még mindig úgy számítana, hogy ismeri a naplójában szereplő útbaigazításokat, mivel azokat a számítógépén olvasta, majd maga írta le. De már nem számítana úgy, hogy mindent tud a Wikipédiáról, mert korábban a legtöbbről nem volt tudomása.

Hasonlóan korlátozó kritérium lehet megfelelő a vállalati kontextusban. Feladataik ellátása során a vállalati alkalmazottak gyakran használnak harmadik felek tulajdonában, fenntartásában és üzemeltetésében lévő, védett adatbázisokat.²¹⁸ Ugyanez igaz a vállalati algoritmusokra is.²¹⁹ A hitelengedélyezési platformok például automatikusan támaszkodnak a hitelminősítő intézetek által fenntartott adatbázisokra.²²⁰ Ha a harmadik fél adatbázisaiban található információk kellőképpen beépülnek a vállalat algoritmikus döntéshozatalába, akkor a vállalat úgy minősülhet, mintha ismerné azokat. És az, hogy a vállalat mit ismer, állandóan változik, mivel az adatbázist fenntartó harmadik fél megváltoztatja annak tartalmát. Ez az eredmény a büntető igazságszolgáltatás szempontjából aggasztó lehet. Mivel a vállalatok nem tudják közvetlenül ellenőrizni a harmadik fél szerverein lévő információkat, a tartás

216. *Lásd* Clark & Chalmers, *Supra* 43. lábjegyzet, 17. pont.

217. *Lásd id.*

218. *Lásd* Pittman, *Supra* note 9, 768-79. o.; *lásd még* Hillary Hellmann, *A fenyegetés elismerése: 36 ENERGY L.J.* 157, 161 (2015).

219. *Lásd* Pittman, *Supra* Note 9, 767-69. o. (a harmadik féltől származó forrásokról általánosságban).

220. *Lásd pl. Comprehensive Credit Decisioning Software*, ACTICO, <https://www.actico.com/solutions/loan-origination-decisioning/credit-decision-platform/> [<https://perma.cc/Y47T-D2QP>] ('A Credit Bureau Gateway zökkenőmentes integrációt biztosít külső adatszolgáltatókkal, például hitelinformációs irodákkal.')

2020]

*KITERJESZTETT VÁLLALATI ELME*929

vállalatok elszámoltatása igazságtalannak tűnhet. A kockázatkerülő vállalatok esetleg tartózkodnak attól, hogy algoritmusaiknak hozzáférést biztosítsanak harmadik féltől származó információkhoz, még akkor is, ha a hozzáférés egyébként költséghatékony lenne.

Feltételezve, hogy kívánatos lenne korlátozni a kiterjesztett testületi elme hatályát, számos korlátozó elv áll rendelkezésre. Az egyik nagyon korlátozó feltétel megkövetelné, hogy az adatrendszerben tárolt információkat, illetve az információkból következtethető tényeket olyan munkavállaló helyezze el, aki korábban jóváhagyta az információkat. Ez közvetlenül párhuzamba állítható a kiterjesztett elme teoretikusai által a természetes személyekre javasolt korlátozással:²²¹, ami gyakorlatilag azt a követelményt jelenti, hogy a vállalat a felettes felelőssége alapján úgy minősüljön, mint aki egyszer már ismerte az információt. Egy sor gyengébb feltétel függhet attól, hogy a vállalati alkalmazottak milyen mértékben ellenőrzik az információt, az agresszív folyamatos ellenőrzéstől az alkalmi minőségellenőrzési mintavételig bármit megkövetelhet. További kérdés lenne, hogy a társaságnak ténylegesen gyakorolnia kell-e ezt az ellenőrzést, vagy elegendő, ha a társaságnak csupán (jogi, gyakorlati vagy egyéb) lehetősége volt rá.

Jó okunk lehet azonban arra, hogy lemondjunk a vállalati ismeretekre vonatkozó további korlátozó feltételekről. A korlátok keresését az egyéni kontextusban vezérlő intuíciók gyengébbek, amikor a vállalatokról van szó. Van egy intuitív értelmezése annak, hogy mit jelent az emberi alany számára az információ "előzetes jóváhagyása". Az emberi alanyok térben és időben korlátozott biológiai egységek. A kiterjesztett elme tézis által felvetett kérdés az, hogy a megismerésük nem terjedhet-e túl a tér-időbeli korlátokon. A vállalatok azonban természetüknél fogva téridőben elosztott alanyok. A vállalati koponyának nincs megfelelője, amelyre rámutathatnánk. Mivel nincs erős intuíció a vállalati elme más entitások által kezelt távoli adatrendszerekre való kiterjesztése ellen, a korlátozó feltételre vonatkozó retributív érv gyenge.

A korlátozó feltétel elutasítása mellett erős elrettentő érv is szólhat. E cikk egyik mozgatórugója az, hogy a jelenlegi jog szerint a vállalatok elszigetelhetik magukat a felelősség alól azáltal, hogy a műveleteket az alkalmazottakról algoritmusokra hárítják át. Ennek az aggodalomnak egy változata azonban megfordulhat, ha a vállalatok kihasználhatják a korlátozó feltételt azáltal, hogy egyes algoritmusoktól és adatbázisoktól másokra hárítják a műveleteket. Tegyük fel például, hogy egy korlátozó feltétel megköveteli az információk valamilyen munkavállalói felügyeletét, például szűrőpróbaszerű minőségellenőrzési auditokat. A vállalat úgy kezelheti kötelezettségeit, hogy az információkat egy másik szervezet által fenntartott távoli adatbázisokba helyezi át. Ezen túlmenően sok esetben az lehet a leghatékonyabb, ha a vállalatokat minden olyan információért felelősségre vonják, amelyhez rutinszerűen hozzáférnek és amelyet rendszeresen felhasználnak. Ez a

221. *Lásd* Clark & Chalmers, *Supra* 43. lábjegyzet, 17. pont.

ösztönözné a vállalatokat a minőségellenőrzés fenntartására, és arra, hogy a harmadik fél információmegőrzőkre is nyomást gyakoroljanak, hogy ugyanezt tegyék.²²² "[A]z ellátás biztosításának legbiztosabb módja, ha a kockázatot arra a személyre hárítjuk, aki eldönti, hogy milyen óvintézkedéseket kell tenni."²²³ Természetesen a kockázatok elosztása tranzakciós költségekkel és egyéb potenciális akadályokkal jár az üzleti kapcsolatokban, amelyek mind az innovációt és a gazdasági fejlődést visszafogják. Nyitott empirikus kérdés, hogy a negyedik követelményhez hasonló valami végül is segítene vagy ártana a vállalati kontextusban. Legalábbis ott, ahol a potenciális társadalmi tét nagy - mint gyakran a tudáson alapuló polgári és büntetőjogi jogsértések esetében -, talán a jognak meg kellene követelnie a vállalatoktól, hogy igényesek legyenek az információ minőségével kapcsolatban, függetlenül attól, hogy honnan származik.

KÖVETKEZTETÉS

Sok vállalat számára az automatizálás jelenti a jövőt. Ez a jövő gyorsabbá és olcsóbbá teszi a vállalatokat, de nem szünteti meg a vállalati károkat. A jog a jelenlegi formájában hamarosan olyan eszközök nélkül marad, amelyekkel a vállalati visszaélések széles körét kezelni lehetne. A legtöbb vállalati felelősség vállalati mentális állapotokat igényel - például a valótlanág ismeretét vagy a csalási szándékot -, amelyeket a törvény jelenleg az alkalmazottak mentális állapotai alapján határoz meg. Amikor azonban algoritmusok irányítják a vállalatot, az alkalmazottak mentális állapota, és így a vállalati felelősség is kikerül a képből. Ez a cikk olyan megoldást javasol, amely kihasználja a társasági felelősség jelenlegi keretrendszerét. A kortárs filozófia és a kognitív tudomány témáira támaszkodva megmutatja, hogy az elmét nem korlátozzák a hagyományosan feltételezett határok. Egy sor külső kognitív segédeszköz tölti be az emberi agyhoz hasonló szerepeket. A kiterjesztett elme tézis szerint ezek a segédeszközök az emberi elme részét képezik. Hasonlóképpen, egy sor algoritmikus segédeszköz fog a vállalatoknál munkavállalóhoz hasonló szerepet betölteni. Ennek megfelelően ez a cikk diszkrét jogi reformokat kínál annak elismerésére, hogy a vállalati elme kiterjedhet ezekre is. Az alap gondolat az, hogy azokat a vállalatokat, amelyek algoritmusokat használnak munkavállalói szerepek betöltésére, úgy kell kezelni, mint amelyek ugyanolyan mentális állapotokkal rendelkeznek, mint azok a vállalatok, amelyek munkavállalókat használnak e szerepek betöltésére. Ez a reform megakadályozná, hogy az opportunisták társaságok a felelősségi kockázatukat azáltal korlátozzák, hogy a műveleteket az alkalmazottakról algoritmusokra hárítják át.

Most már világosnak kell lennie, hogy a javaslat megfelel e cikk minimalista célkitűzésének - a hatályos jogban nagyon kevés változtatásra lenne szükség. A javaslat nagymértékben támaszkodik a társasági jog jelenlegi felelősségi keretrendszerére. Valójában a társasági személyiség fikciója az egyik legfontosabb motiváció a kiterjesztett felelősségre vonás kiigazítására.

222. Az indoklás itt párhuzamos azzal az esettel, amikor az önvezető autók gyártóit felelőssé teszik a harmadik fél által okozott sérülésekért. *Lásd* Geistfeld, *Supra* 17. lábjegyzet, 1690. lábjegyzet.

223. OLIVER WENDELL HOLMES, JR., *THE COMMON LAW* 117 (Little, Brown & Co. 1923) (1881).

2020]

*KITERJESZTETT VÁLLALATI ELME*⁹³¹

az elmeelméletet a természetes emberektől a vállalati kontextusig. Ez a lépés csak azon a fiktív feltételezésen alapul, hogy a vállalatok olyan emberek, akiknek olyan elméjük van, mint nekünk. Ugyanilyen fontos, hogy ez a cikk sehol sem feltételezi - ahogy a törvény sem -, hogy az algoritmusoknak van elméjük, vagy hogy felelősek lehetnek. A kiterjesztett elme tézis szerint a hipotézis nem az, hogy a külső kognitív segédeszközöknek saját, független mentális állapotuk van.²²⁴ Barry naplója nem tudta, hogyan kell eljutni a kávézóba, még akkor sem, ha Barry (a naplójába írt útbaigazítással) tudta. Analóg módon itt az állítás az, hogy az algoritmusok a vállalati elme részét képezhetik, nem pedig az, hogy saját elmével rendelkeznek. A vállalatok közvetlenül felelősek lehetnek azokért a dolgokért, amelyeket döntenek és tesznek, még akkor is, ha mesterséges intelligenciát használnak a döntések meghozatalához és a cselekvésekhez.

224.

Andy Clark, *Coupling, Constitution, and the Cognitive Kind*, in *THE EXTENDED MIND*, *supra* 109. megjegyzés, 81., 83. pont.

932NORTH

CAROLINA LAW REVIEW

[Vol. 98