

ZÖDI ZSOLT*

JOG ÉS JOGTUDOMÁNY A BIG DATA KORÁBAN**

A „Big Data” kifejezés az Interneten keletkező hatalmas adatmennyiségre utal, a „Big Data korszak” pedig arra, hogy ez az adattermelődés új társadalmi jelenségeket indukál már jelenleg is, és még komolyabb változásokat indukálhat a jövőben. A cikk ennek a jelenségnek a jogra és a jogtudományra gyakorolt hatásával igyekszik számot vetni.

Egyelőre természetesen nem mérhető fel minden hatás, de az már most látható, hogy a Big Data jelenség – főként a rá alapozható döntések, a predikciós (jóslási) potenciálja miatt már rövid távon teljesen felforgathatja az adatvédelem és az üzleti döntéshozatal világát, de középtávon hatással lehet a jogalkalmazásra és a jogalkotás (szabályozás) tervezésére, hosszabb távú hatásai pedig szinte felmérhetetlenek. Az már most látszik, hogy a Big Data birtokában olyan információs aszimmetria alakul ki, amely feltétlenül szabályozást fog igényelni, és az is biztos, hogy a Big Data-ra alapuló automatikus, gépi döntések átláthatósága a közeljövő egyik legégetőbb problémája lesz. Ezenfelül a Big Data igen komoly morális problémákat is felvet, részben azért is, mert úgy tűnik, hogy a jövő mesterséges intelligenciái nagyrészt Big Data alapúak lesznek.

A cikk négy területet vizsgál: a Big Data-t mint szabályozandó társadalmi jelenséget (mit és hogyan kellene és egyáltalán kell-e szabályozni), a Big Data jelenséget mint a jogalkotás és a jogalkalmazás egy lehetséges új, hatékony eszközét, végül a Big Data jogtudományra gyakorolt hatásait és a Big Data használatát mint a jogtudományi kutatás egy lehetséges új módszerét.

A Big Data jelenség¹ (a továbbiakban BD, vagy BD-jelenség) elsősorban a számítástechnika, a szociológia és az üzlet világában tűnt fel, és fut be máig tartóan nagy karriert. A jogot és a jogra reflektáló jogtudomány világát bár nem hagyta épp érintetlenül, a reflexiók száma jóval csekélyebb és szűkebb körű, mint például az üzle-

* PhD, egyetemi docens, Budapesti Corvinus Egyetem Gazdálkodástudományi Kar, 1093 Budapest, Fővám tér 13–15; tudományos munkatárs, MTA Társadalomtudományi Kutatóközpont Jogtudományi Intézet, 1097 Budapest, Tóth Kálmán u. 4.
E-mail: zodi.zsolt@tk.mta.hu.

** Az írás „Az internetes forgalomirányító szolgáltatások szabályozási kérdései” című, 116979 számú OTKA-kutatás keretében készült.

¹ Hagományosan, eredetileg az Eaton és szerzőtársai könyvében – Chris EATON [et al.]: *Understanding Big Data; Analytics for Enterprise Class Hadoop and Streaming Data* (New York: McGraw Hill 2012) 5. – található definíció alapján három, majd később különböző forrásokban (pl. az IBM 'Big Data Hub'-ján: www.ibmbigdatahub.com/infographic/four-vs-big-data) négy jellegzetességet, „V-t” említenek a Big Data definíciójaként. Ez az a „négy V”, amely a 'Big Data'-t megkülönbözteti az „egyszerű” adatoktól. Ezek: a Volume (nagy mennyiség), Velocity (gyors változás), Variety (nagy heterogenitás) és Veracity (ellenőrizhetetlen megbízhatóság). Ugyanakkor az a benyomásom, hogy a „négy V” korántsem konjunktív feltétele annak, hogy 'Big Data'-ról beszélhessünk. A jogi tárgyú írások gyakran emlegetik – sőt igazság szerint a leggyakrabban ezt emlegetik

ti tudományokban,² vagy épp a politikai szférában.³ Ez a cikk hozzájárulás szeretni lenni a BD-jelenség és a jog kapcsolatának megértéséhez.

Nem kísérlem meg ennek a szerteágazó jelenségnek a pontos definícióját megadni. Nagyjából mindenki számára érthető, hogy miről is van szó: az Internet és a részben az Internetre kötött eszközök (szenzorok és a „Dolgok Interneté”, „*Internet of Things*”, „*IoT*”) elterjedésével óriási mennyiségű adat képződik minden pillanatban, amely közvetve a társadalomról, rólunk, emberekről szolgáltat információkat. Igen sokféle eszközre és keletkezési módra lehet gondolni, a mobiltelefonok által generált cellainformációktól és helyadatoktól, a keresőmotorokba begépelte keresési kifejezésektől kezdve az online pénztárgépek és az okos-mérőórák által nap mint nap termelt adatokon, valamint a közigazgatás, vagy a bírósági rendszer által termelt szöveges dokumentumhalmazokon keresztül egészen a kültéri kamerák által rögzített képekig, vagy bizonyos tudományos projektekben generált adatokig. A jelenség nyilván nem elsősorban a számítástechnikai, vagy adattárolási problémák miatt érdekes, hanem a társadalmi hatásai miatt. Ez a hatalmas mennyiségű adat ugyanis nemcsak másfajta tárolást, kezelést és értelmezést igényel, és másfajta, egészen újszerű módokon használható, mint akár a hagyományos offline világban keletkezett adatok, akár a kontrollált módon épített és „táplált” adatbázisokban tárolt adathalmazok, hanem a jelenségnek nyilvánvalóan társadalmi következményei is vannak. Ezek egy része már most is érzékelhető, de a jövőbeli hatásai egyelőre szinte beláthatatlanok.

– az egészségügyben elektronikusan képződő betegadatokat és az ebből a speciális algoritmusok segítségével levonható következtetéseket és előrejelzéseket mint 'Big Data' jelenséget. Ezekben az esetekben inkább az adatnak a felvételtől célhoz képest *eltérő célra történő használata*, valamint az adatokból automatikusan, algoritmusok segítségével *levonható következtetések*, és az adatokon alapuló *jóslási (predikciós) képesség* az, ami miatt a 'Big Data'-jelenséghez sorolhatjuk ezt az adathalmazt.

² Összehasonlításképpen, csak egyetlen kiadónál, a Wiley-nél 30(!)-féle könyv kapható jelenleg, amely az üzlet és a Big Data kapcsolatát elemzi – meg sem merem becsülni hány könyv született összességében. A jog és a BD kapcsolatát tudtommal monográfia eddig egyetlen egy sem és – a HeinOnline szakcikkbázisból történt lekérdezés után – cikk is csak 180 elemzi.

³ Az USA kormánya igen aktív volt az Obama-adminisztráció idején a témakörben, és ennek nyomai az azóta archivált weboldalakon is megtalálhatók. Az elnök hivatala 2012-ben indította útjára a Big Data Kutatási és Fejlesztési Kezdeményezést (Big Data Research and Development Initiative – obamawhitehouse.archives.gov/blog/2012/03/29/big-data-big-deal). A program a részben már ekkor is futó kormányzati kezdeményezések összehangolására, áttekintésére is vállalkozott (obamawhitehouse.archives.gov/the-press-office/2014/05/01/fact-sheet-big-data-and-privacy-working-group-review). 2014-ben jelent meg a „Big Data: Seizing Opportunities, Preserving Values” című tanulmány (obamawhitehouse.archives.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf), amely a BD-jelenséget először próbálta megérteni a politika és a szabályozás szemszögéből. Ezt 2015-ben a Nemzeti Tudományos Alapítvány (*National Science Foundation*) Big Data központjainak (*Big Data Hubs*) felállítása követte (obamawhitehouse.archives.gov/blog/2015/11/04/big-announcements-big-data és Big Data Hubs NFS Program Solicitation [2015], www.nsf.gov/pubs/2015/nsf15562/nsf15562.htm). Végül 2016 májusában készült el egy átfogó elemzés, amely a BD és egyes társadalmi szférák kapcsolódásait vizsgálja, tudományos alaposággal. Magam is sokat támaszkodom erre az anyagra (Big Data: A Report on Algorithmic Systems, Opportunity and Civil Rights. Executive Office of the President. May 2016, obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf).

Ebben a cikkben a társadalmi hatások két aspektusát, a BD-jelenség jogra (a jogi szabályozásra és a „jogászkodásra”), valamint a jogtudományra gyakorolt hatását szeretném megvizsgálni. Az itt található fejtegetések és következtetések egy része természetesen egyszerűen a jelenlegi folyamatok extrapolálásával felállított hipotézis. Mint minden ilyen jellegű fejtegetésnél, a hiba lehetősége nagy. Ugyanakkor ezeket a jóslatokat nem szerettem volna kihagyni a tanulmányból, hiszen a feldolgozott irodalmak egy része is hasonlóan jár el, és a téma miatt ez csaknem elkerülhetetlen.

A BD mint fogalom már 2010 előtt feltűnt, igazán azonban a 2010-es évek elején kezdték el komolyan elemezni – egyáltalán felismerni, hogy a tömeges adattermelődés mélyebb változásokat jelez és indukál a társadalomban. Ekkortól több száz közlemény jelenik meg a világ vezető *jogi* folyóirataiban is a BD és a jog kapcsolatáról.⁴ A cikkek egy része a jog és a BD általános viszonyát elemzi. Olyan kérdéseket, mint a Big Data által felvetett etikai problémák,⁵ vagy a BD hatása általában az adatvédelemre.⁶ Más írások a BD és a jog egyes részterületeinek kapcsolatát vizsgálják, sokszor egészen konkrét problémákra és esetekre utalva.⁷

A BD és a jog kapcsolata sokféleképpen tematizálható, de én itt nem a fenti, szintén logikus felosztást (általános viszony és konkrét problémák) követem, hanem azt a felosztást alkalmazom, amely a kölcsönhatásnak, közös metszeteknek *két területét* és ezen belül *két szintjét* különbözteti meg. Ez a felosztás azon az egyszerű feltevésen alapszik, hogy mind a jog, mind a jogra kívülről reflektáló jogtudomány egyszerre tekinti a BD-jelenséget *tárgynak* (szabályozási tárgynak és vizsgálati tárgynak), és használja azt *eszközüül* (a jogalkotáshoz és a jogalkalmazáshoz, illetve a tudományos kutatáshoz). A két szempont alapján az alábbi táblázat rajzolható fel:

1. ábra. A jog(tudomány) és a Big Data lehetséges kapcsolódási pontjai

	Jog	Jogtudomány
Tárgy	① A BD által felvetett adatvédelmi, adatbiztonsági, alkotmányos stb. kérdések szabályozása.	⑤ A BD mint társadalmi és jogtudományi jelenség – például a lehetséges szabályozás alternatívái, módozatai, morális dilemmái stb.
Eszköz	② A BD használata a jogalkotásban – tudományos jogalkotás, adatalapú „ <i>social engineering</i> ”. – A BD használatának lehetőségei a jogalkalmazásban: BD-alapú elemző-, kereső- és predikációs eszközök.	④ A BD mint kutatási eszköz: a BD használata például nagy tömegű bírósági ítélet, hatósági határozat stb. statisztikai módszerekkel történő elemzéséhez; <i>natural language processing</i> és tartalomelemzési módszerek használata, stb.

⁴ Tudomásom szerint az első, azóta is sokat idézett tanulmány a jogi területen Tene és Polonetzky [2012] írása volt (Omer TENE – Jules POLONETSKY: „Privacy in the Age of Big Data: A Time for Big Decisions” *Stanford Law Review Online* 2012/February. 63–69.), amely azonban többször hivatkozza az *Economist* 2010. február 25-én megjelent különszámát, amely teljes egészében a „nagy adat” témaköréről szól.

⁵ Neil M. RICHARDS – Jonathan H. KING: „Big Data Ethics” *Wake Forest Law Review* 2014/2. 393–432.

⁶ TENE–POLONETZKY (4. l.).

⁷ Pl. Andrew Guthrie FERGUSON: „Big Data and Predictive Reasonable Suspicion” *University of Pennsylvania Law Review* 2015/2. 327–410, vagy Jin-ho JEON – Seung-Ryul JEONG: „Designing a Crime-Prevention System by Converging Big Data and IoT” *Journal of Internet Computing and Services* 2016/3. 115–128. (www.jksii.or.kr/upload/1/1212_1.pdf) a BD-alapú bűnmegelőzés (előrejelzés) vonatkozásában.

Ez a cikk műfajilag a jobb felső (harmadik) cellába tartozik, azaz (részben) a jogtudomány módszereivel igyekszik kijelentéseket tenni a Big Data és a jog világának interferenciájáról. A jogtudományi vizsgálat természetesen önmagát is tárggyá teheti, de vizsgálhatja a másik három kapcsolódási pontot is. Én pontosan így fogok tenni: a cikkben röviden igyekszem mind a négy területet érinteni.

Az 1. és a 2. területeken keletkezik a legtöbb irodalom. Itt emiatt inkább egyfajta panorámát igyekszem felvázolni, hiszen nem tudok róluk kimerítő képet adni. A célom itt tehát a BD-jelenség beemelése a magyar jogtudományi diskurzusba. A 3. és 4. pont (a BD és a jogtudomány) kapcsolata már korántsem olyan sokrétűen tárgyalt probléma, mint az első két terület, így itt inkább az első benyomásaimat fogom ismertetni, részben egy általam lefolytatott kutatás kapcsán is, amelyben némi gyakorlati tapasztalatot is szereztem a BD(-szerű) adatok használhatóságáról a jogtudomány területén.

1. A BD MINT A JOGI SZABÁLYOZÁS TÁRGYA

A legtriviálisabb kapcsolódás tehát az, hogy a Big Data *szabályozandó társadalmi jelenség*. Mint minden új társadalmi jelenség esetén, a BD-vel kapcsolatban is az az első, közhelyszerű megállapítás tehető, hogy egyelőre a szabályozóknak nincsen ötletük, hogy hogyan viszonyuljanak hozzá. Világosan látható ugyanis, hogy egyszerre rejt óriási lehetőségeket, és egyben hatalmas veszélyeket is. És ahogyan ez lenni szokott, már abban bizonyos tanácstalanság mutatkozik, hogy miben is lenne megragadható ennek az új társadalmi jelenségnek a lényege, egyáltalán szükség van-e bármilyen szabályozásra, és különösen *új szabályozásra*.

Ami a kockázatokat illeti, a szakirodalom a következő, már most is látható és egymással is összefüggő problémákat emlegeti a BD-jelenséggel kapcsolatban.

1. Az első az, hogy az egyébként hozzájárulással gyűjtött személyes, sőt a nem személyes jellegű adatokból a BD-analízisek segítségével olyan következtetések vonhatók le, amelyek *messze túlmutatnak* a hagyományos személyes adatkezelés veszélyein és karakterén. A klasszikus példa a sokak által ismert *Target*-eset. A Target amerikai áruházlánc egy, a vásárlói szokásokból kialakított BD-alapú predikciós algoritmust használva fedte fel egy 16 éves lány terhességét, majd küldött neki gratuláló levelet és baba-kellékeket tartalmazó katalógust, melyet persze a fiatal hölgy apja bontott azután ki.⁸

2. A második esetkör ehhez hasonló, de mégis kicsit más karakterű. Arról van szó, amikor a BD az egyébként anonimizált adatbázisok „visszaszemélyesítését” teszi lehetővé.⁹ Ez egy még tágabb horizontról azt veti fel, ahogyan Tene és

⁸ Az ügy alapos, elmélyült feldolgozását adja Charles DUHIGG: „How Companies Learn Your Secrets” *New York Times* 2012. 02. 19., www.nytimes.com/2012/02/19/magazine/shopping-habits.html?pagewanted=1&_r=1&hp.

⁹ Itt a klasszikus példa az AOL-eset, amikor az internetszolgáltató kutatási céllal megnyitotta 2006-ban a keresési stringeket tartalmazó adatbázisát, amelynek segítségével kiszűrve leleplezték egy 62 éves, Thelma Arnold nevű hölgyet, aki pl. olyan zavarba ejtő kereséseket folytatott, mint

Polonetzky,¹⁰ valamint Crawford és Schultz¹¹ is rámutatnak, hogy a Big Data korában a „hagyományos” adatvédelem szinte minden kategóriája megkérdőjeleződik és átalakul. A személyhez fűződő adat, az arányosság, a célhoz kötöttség, és a hozzájárulás jelentősége egyaránt más értelmet nyernek. Az első problémakörre példaként azt említik, hogy újabban több esetben is bebizonyosodott, hogy a BD-elemzésekre támaszkodva a magánszemély és az adat között felszámolt összefüggés helyreállítható,¹² például az eddig nem védett metaadatok használatával.¹³ A második problémakörből pedig elég csak a hagyományos adatvédelem „hozzájárulás”-centrikusságát kiemelni. A BD korában egyszerűen annyi adat képződik, hogy azt egy magánszemély képtelen kontrollálni. Végül a célhoz kötöttség elvének szigorú betartása a BD korszakában megbénítaná az üzleti életet, amely igen széles körben kezdte különböző következtetések levonására, személyre szabásra, célzott marketingre és predikcióra használni ezeket az adatokat.

3. A harmadik probléma szintén a BD azon karakterével van összefüggésben, hogy segítségével egy üzleti vállalkozás sokkal mélyebbre képes látni, mint a hagyományos információgyűjtéssel, és ezzel például kirekesztő, bizonyos társadalmi csoportokat megkülönböztető gyakorlatokat alakíthat ki. A probléma az amerikai *Federal Trade Commission* (FTC) figyelmét is felkeltette, és 2014-ban egy konferenciát is rendezett a témakörben.¹⁴ A FTC egy másik jelenségre is felfigyelt: az egyéniesített BD-alapú információkból egyéni, olykor diszkriminatív árazást lehet kialakítani,¹⁵ amelyet korábban csak emberi (értékesítői) döntéssel lehetett. Ilyen egyedi árazás például az Uber árazási metódusa, amelyről lentebb beszélünk majd.

Kérdés, hogy mennyiben más az a helyzet, amikor valakinek a BD-hez van hozzáférése, mint ha egyszerűen nagy mennyiségű másféle adathoz. Először is azt érdemes tisztázni, hogy bár a BD gyakran nem személyes adat (legfeljebb újfajta módszerekkel visszaállítható a személyes jellege), ugyanakkor a fenti példák-ból egyértelmű, hogy *információs túlhatalmat* hozhat létre annál, aki rendelkezik vele, főként azzal, hogy *következtetések* vonhatók le belőle, *sőt az előrejelzés* is lehetővé válik általa.¹⁶ Nagy adathalmazokból az előrejelzés és a következte-

hogyan „kutya, amely mindent levizel”. Paul OHM: „Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization” *UCLA Law Review* 2010/6. 1701–1778. 1718.

¹⁰ TANE–POLONETZKY (4. lj.).

¹¹ Pl. Kate CRAWFORD – Jason SCHULTZ: „Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms” *Boston College Law Review* 2014/1. 93–128, 94.

¹² OHM (9. lj.) 1717–1722.

¹³ Big Data: Seizing Opportunities, Preserving Values (3. lj.) 49.

¹⁴ Big Data: A Tool for Inclusion or Exclusion? [2014] Konferencia a Federal Trade Commission szervezésében, 2014. szeptember 14-én, www.ftc.gov/news-events/events-calendar/2014/09/big-data-tool-inclusion-or-exclusion.

¹⁵ The White House Council of Economic Advisers. „Big Data and Differential Pricing” February 2015. obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/docs/Big_Data_Report_Nonembargo_v2.pdf, valamint lásd a 14. lj.-ben szereplő konferencia anyagait.

¹⁶ Big Data: Seizing Opportunities, Preserving Values (3. lj.) 50, vagy Phil SIMON: *Big Data Lessons from Netflix*, www.wired.com/insights/2014/03/big-data-lessons-netflix/. A profilizálásról és a nagy pontossággal személyre szabott reklámokról van szó.

tés eddig is lehetséges volt,¹⁷ a BD-korszak ezt illetően az adatok képződésének módjában (spontán, folyamatos adatképződés), mennyiségi növekedésében (sokkal több dologról áll rendelkezésre adat), és az adatokhoz történő sokkal szélesebb körű hozzáférésben (rengeteg magánszervezet is folyamatosan gyűjt és elemez adatokat) különbözik. A BD prediktív ereje szinte minden tanulmányban középponti helyet foglal el, és nem véletlenül. Aki a jövőt tudja, az olyan hatalommal rendelkezik, amely szinte isteni jelleggel ruházza fel őt.¹⁸ A helyzetet még bonyolultabbá teszi az, hogy az előrejelzést algoritmusok és gépek végzik, amelyek egy egyszerű ember számára átláthatatlanul működnek, sőt, olykor még a következtetések is alig érthetőek meg „józan ésszel”. Ezért – úgy tűnik – hogy a BD-vel kapcsolatos szabályozási törekvések többsége ennek a szabályozási túlhatalomnak a megszüntetésére, enyhítésére, kompenzálására, valamint az algoritmusok transzparenssé tételére irányul. Többen például azzal érvelnek, hogy minden eszközzel ösztönözni kell a BD-tulajdonosokat, hogy ezeket az adatszetteket hozzák nyilvánosságra.¹⁹

A BD azért is tűnik újfajta megközelítést igénylő jelenségnek, mert bár a hagyományos adatvédelmi rezsimek alapelvei ebben a világban is számon kérhetők és elvileg betarthatók is lennének (teljes és mindenre kiterjedő hozzájárulás az érintett részéről, a személyes adat személyes jellegének megtartása addig, ameddig bármilyen módon helyreállítható a kapcsolat, a különböző adatok átadásának és összekapcsolásának teljes tilalma stb.), csakhogy ebben az esetben a BD-t teljesen megfosztanánk az értékteremtő erejétől. Itt látom a BD-jelenség másik új tulajdonságát, hogy a korábbi, szinte teljes egészében „érintett”-központú, az érintettek védelmére egyoldalúan koncentráló adatvédelmi szabályozás helyett a legtöbb írás nagyon komolyan a védelemmel és a veszélyekkel legalább megegyező súllyal említi, hogy a BD óriási *lehetőségeket rejt*, amelyeket nem szabadna kiaknáztatlanul hagyni. E lehetőségek között a legnagyobb súllyal az szerepel, hogy a mesterseges intelligenciák szinte kivétel nélkül BD-alapúak.²⁰

A téma egyik szerzője, Byers például egy empirikus kutatásra alapozva öt területet jelöl meg, ahol a BD értékteremtő ereje érvényesül.²¹

¹⁷ Ez a „következtető statisztika” (*inferential statistics*) lásd pl. James W. GRIMM – Paul R. WOZNIAK: *Basic Social Statistics and Quantitative Research Methods* (Belmont: Wadsworth 1990) 271–336. Többen jelezték a cikk olvasása után, hogy a legtöbb probléma és módszer, amelyet itt ismerek nem BD-specifikus, mert jellemző a hagyományos statisztikai elemzésekre és kvalitatív módszerekre is. Ez igaz. A BD használatá során hasonló matematikai módszereket használnak, mint a hagyományos kvalitatív statisztikai kutatásokban. A különbségeket fentebb taglalom.

¹⁸ A BD predikciós képességéről általában Eric SEGEL: *Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die* (New Jersey: Wiley 2013). A vásárlási szokásokat illetően Veronica MCGREGOR – Sophia Helena CALDERON – Roberta TONELLI: „Big Data and Consumer Financial Information” *Business Law Today* 2013/11. 1–4. A rendőrségi munkával kapcsolatban FERGUSON (7. lj.).

¹⁹ Pl. Michael MATTIOLI: „Disclosing Big Data” *Minnesota Law Review* 2014/2. 535–584, 537.

²⁰ Csak gondoljunk a Google Translate projektjére, vagy az önzetető autók által használt tudásbázisokra.

²¹ Angela BYERS: „Big Data, Big Economic Impact” *I/S: A Journal of Law and Policy for the Information Society* 2015/3. 757–764.

1. A BD a piacon, de akár egy cégen belül is az addig *láthatatlan teljesítménykülönbségek mérését* teszi lehetővé. BD alapon mérhetők az egyes alutjeljesítő résztelemek, BD alapon hasonlíthatók össze a piacon szolgáltatások és árak, sőt – tesszük hozzá – a BD az egyébként nehezen mérhető közszolgáltatások (például az egészségügyi szolgáltatások, vagy a hivatali ügyintézés) minőségének a mérésére, de legalábbis hatalmas mintákon történő összehasonlítására is alkalmas.

2. A BD az élet sok területén *kísérletezésre és szimulációkra* ad alkalmat. Például weboldalakat lehet kétféle verzióban elindítani, hatalmas adatállományokban lehet változókat kicserélve a végeredményt megfigyelni. Ezek a kísérletek jóval pontosabb eredményeket adnak, mintha kisebb mintákon végeznénk, vagy egyszerűen spekulálnánk.

3. A BD az élet szinte minden területén alkalmas bizonyos populációk szegmentálására, és így *egyenre szabottabb akciók* megtételére, ajánlatok és szolgáltatások kínálására, döntések meghozatalára. A marketing kézenfekvő területe ennek, de elég csak a biztosítás területét (például kik a gyakoribb károkozók, melyek a veszélyes helyzetek stb.), vagy a prediktív rendőrségi cselekvést²² említeni.

4. A korábbi területekkel összefügg, hogy a BD sokkal kiterjedtebben lehetővé teszi az *automatizált döntéshozatalt*. Egy olyan vállalatnál, amelyben a gyártás nagymértékben automatizált, és a gépek működése szenzorokon alapszik, az ezekből származó információkból egy jó algoritmus gyorsabban tud jobb döntéseket hozni, mint egy ember, ami végső soron az erőforrások (így a természeti erőforrások) sokkal optimálisabb kihasználásához vezet.

5. Végül a BD új üzleti modellek megszületését is lehetővé teszi. Ha az információk valós időben rendelkezésre állnak, akkor például a pillanatnyi kereslet-kínálatot figyelembe vevő árazást lehet kialakítani, ahogyan azt az *Uber surge pricing-ja* esetében láthatjuk.²³

Ugyanígy, az EU például valósággal eufóriában van a BD-vel kapcsolatban, a Tanács „Egy virágzó adat-alapú gazdaság felé” című dokumentuma szinte kizárólag a BD előnyeire koncentrál olyan területeken, mint „az egészségügy, élelmiszerbiztonság, klímavédelem és energiahatékonyság, intelligens szállítási rendszerek és okos városok”.²⁴ Jóllehet az anyag az adatalapú gazdaságról általában (így nem csak a BD-jelenségről) szól, a négy szabályozási terület, amelyet említ (adat-

²² JEON–JEONG (7. l.).

²³ Pl. Jonathan HALL – Cory KENDRICK: „The Effects of Uber’s Surge Pricing” *UBER Newsroom* 2015. szeptember 17., newsroom.uber.com/the-effects-of-ubers-surge-pricing/. Azért tegyük hozzá, hogy ez az eset azt is mutatja, hogy a BD olyan lehetőségeket nyit meg, amelyek talán a gazdasági növekedés és a szabad piac szempontjából nézve előnyösek, ugyanakkor más csoportok számára hátrányosak lehetnek. Bizonyos esetekben és csoportok számára a taxi hatósági ára kiszámíthatóbb, mint a szituatív *surge pricing*, vö. Utpal M. DHOLAKIA: „Everyone Hates Uber’s Surge Pricing – Here’s How to Fix It” *Harvard Business Review* 2015. december 21., hbr.org/2015/12/everyone-hates-ubers-surge-pricing-heres-how-to-fix-it.

²⁴ Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Towards a Thriving Data Driven Economy. Towards a thriving data-driven economy (COM/2014/0442 final), eur-lex.europa.eu/legal-content/EN/TXT/?qid=1404888011738&uri=CELEX:52014DC0442.

és fogyasztóvédelem, adatbányászat, adatbiztonság, adattulajdonlás és -átadás)²⁵ részben a BD világát is érinti. Az eufória az USA kormányzatát is jellemzi.²⁶

Mindezek alapján nehéz válaszolni arra, hogy vajon a BD esetében elég lesz-e a hagyományos adatvédelmi szabályozás átalakítása, esetleg a személyiségi jogvédelmi (*privacy*) szabályozás átértelmezése, kiterjesztése, vagy épp szűkítése ahhoz, hogy kezelje azt. Úgy tűnik, hogy néhány új szabályra szinte bizonyosan szükség lesz; a kérdés csak az, hogy ezek a szabályok beilleszthetők lesznek-e a „személyes adatok védelmének”, esetleg az „elektronikus kereskedelem” szabályainak a keretei közé, vagy új *keretekre* is szükség lesz. A hagyományos adatvédelem is információs túlhatalmaktól védte az állampolgárt, de a BD információs túlhatalma ezt az aszimmetriát új dimenzióba emeli. Ugyanakkor ennek legfőbb eleméről, a varázslatos predikciós erőről egyelőre *szinte semmit nem tudunk*: egy-két esetben ez egészen félelmetes, ijesztő volt, de senki sem tudja biztosan, hogy vajon ezek nem csak elszigetelt, és véletlenek által segített esetek-e? Vajon az egyedi vásárlásainkkal, letöltéseinkkel, kereséseinkkel kapcsolatos adathalmazok bizonyos matematikai algoritmusok segítségével valóban olyan előrejelzéseket tesznek-e majd lehetővé, amelyek olyasmiről árulkodnak, amit magunk sem tudtunk saját magunkról, s ezáltal a nagy adathalmazokkal rendelkező cégek, szervezetek legbelsőbb titkainkhoz, szinte a lelkünkhöz, agyunkhoz férnek majd hozzá? És a másik oldalról nézve ugyanilyen súlyos, és egyelőre megválaszolhatatlan kérdés, hogy ezek az adatok és algoritmusok a gazdasági növekedésnek – egyáltalán a gazdaságnak és a szolgáltatásoknak – valóban új, virágzó szakaszát nyitják-e majd meg? Amíg ezeket nem tudjuk, arra a kérdésre is lehetetlen válaszolni, hogy a régi szabályozás átalakítása, vagy teljesen új szabályok kellenek-e.

Végül itt kell megemlékeznünk arról, hogy a fenti dilemmát tükrözi, hogy az EU és a BD viszonya is ellentmondásos. Mint az ismert, az EU még a 2010-es évek elején kezdte el a régi adatvédelmi irányelvének felülvizsgálatát. A paradox az EU-s szabályozással kapcsolatban az, hogy amikor a munka elindult, a Big Data még egyszerűen nem volt a közbeszéd része. Az adatvédelemről készült speciális Eurobarometer-jelentésekben²⁷ – sem 2010-ben, sem 2015-ben – nem fordul elő a Big Data kifejezés. A kifejezés mellőzése ellenére persze a jelenség néhány aspektusára a kérdőív rákérdezett. Így például az európaiak 71%-a gondolja a személyes adatok átadását a modern élet egyre gyakoribb velejárójának,²⁸ sőt elkerülhetetlennek, és ez a szám az iskolázottsággal egyenes arányban,²⁹ az életkorral fordított arányban³⁰ növekszik. Ugyanakkor ez az anyag és a többi előkészítő anyag is mindvégig abban a fogalmi keretben mozog, amelynek az a lényege, hogy a személyes adatok védelmét fokozni kell, mert csak így lehet a fogyasztók *bizalmát*

²⁵ 24. lj. 11–12.

²⁶ Lásd 3. lj.

²⁷ Special Eurobarometer 431. Data Protection (adattfelvétel: 2015 március, publikálás dátuma: 2015 június), ec.europa.eu/public_opinion/archives/ebs/ebs_431_en.pdf.

²⁸ 27. lj. 28.

²⁹ 27. lj. 23, 33.

³⁰ 27. lj. 33.

elnyerni. A bizalom egyre intenzívebb használatot fog eredményezni, amely vég-ső soron segíti a gazdasági növekedést, az online magán- és kormányzati szolgáltatások kiépülését is.³¹ A másik érték, amelyre az irányelv épül, az adatok szabad áramlásának elősegítése, és az adatvédelmi szintek kiegyenlítése. Nincsen tehát szó explicite a BD-jelenségről, ugyanakkor az EU is érzékelt, hogy valahogyan reflektálni kellene rá: egy közelmúltbeli *policy paper* a BD és a Rendelet viszonyát próbálja tisztázni,³² és ez is azzal érvel, hogy a szigorú adatvédelmi szabályok által felkeltett és fenntartott fokozott fogyasztói bizalom versenyelőnyt fog biztosítani az európai vállalkozásoknak.

Nekem kétségeim vannak azzal kapcsolatban, hogy az adatvédelmi szttenderdek folyamatos emelése önmagában, a bizalom emelkedése miatt élénkebb adat-ipart, és adathasználatot fog eredményezni. A rendelet más részei nem BD-barát szabályozást mutatnak. A rendelet a BD-t ugyan nem, de a BD legfontosabb *felhasználási területét*, a profilalkotást és az előrejelzést említi, méghozzá meglehetősen negatív kontextusban „automatikus döntéshozatalként”.³³ A rendelet ezzel kapcsolatban alapvetően azt a szabályt fekteti le, hogy a magánszemélyeknek joguk van arra, hogy rájuk ilyen automatikus döntéshozatal *egyáltalán* ne terjedjen ki, illetve, amennyiben kiterjed, joguk van tudni a mögöttük meghúzódó logikákat, algoritmusokat, és bármikor „emberi beavatkozást” kérni, álláspontjukat kifejezni, és kifogást benyújtani. Nem nehéz elképzelni, hogy ezeket a szabályokat mennyire nehéz lesz betartani, és azt is, hogy ezek a szabályok mennyire nem azoknak a cégeknek kedveznek, akik szeretnék a BD előnyeit akár a saját tevékenységük körében, akár más vállalkozások számára szolgáltatásként használni. Az, hogy a BD-vel előrejelzéseket lehet tenni, és döntéseket automatikusan lehet hozni, a BD értékteremtő erejének lényege. Az ilyen szabályok elriasztják ezeket a cégeket, így el sem indítják azokat a szolgáltatásokat, amelyek magas adatvédelmi szintje a vevőket utóbb ezek igénybevételére csábítaná. Később majd ismertettem például a jogi információszolgáltatás területén tevékenykedő két amerikai cég BD-alapú, hihetetlenül innovatív online szolgáltatását. Ezeket a szolgáltatásokat Európában jelenleg egyszerűen *el sem lehetne indítani*.

³¹ Vö. pl. Commission Staff Working Paper. Impact Assessment, Accompanying the document (...) General Data Protection Regulation (...) Brussels, 25. 01. 2012, ec.europa.eu/justice/data-protection/document/review2012/sec_2012_72_en.pdf.

³² The EU Data Protection Reform and Big Data (Factseet), 2016. március, ec.europa.eu/justice/data-protection/files/data-protection-big-data_factsheet_web_en.pdf.

³³ Általános adatvédelmi rendelet – Az Európai Parlament és a Tanács (EU) 2016/679. rendelete (2016. április 27.) a természetes személyeknek a személyes adatok kezelése tekintetében történő védelméről és az ilyen adatok szabad áramlásáról, valamint a 95/46/EK rendelet hatályon kívül helyezéséről (általános adatvédelmi rendelet), preambulum, 72. pont, 13. cikk (2) f), 14. cikk (2) g) és 15. cikk (1) h) és 22. cikk.

2. A BD MINT ESZKÖZ A JOGALKOTÓK ÉS A JOGALKALMAZÓK KEZÉBEN

A második cella a táblázatban azt mutatja, hogy a jog nemcsak szabályozhatja a BD-t, hanem támaszkodhat is rá, eszközként is használhatja. Ha a jog világát vesszük, itt újból két szintet érdemes megkülönböztetni: egyfelől a BD a jogalkotás egyik új, fontos eszköze lehet, másfelől a BD a jogalkalmazásban is helyet kaphat.

Ami ez elsőt illeti, az az egyszerű feltételezésem van – és szemmel láthatóan az EU-nak is³⁴ –, hogy a BD segítségével jobb jogszabályok alkothatók, de legalábbis a BD új perspektívákat nyithat meg a jogszabály-előkészítésben és a hatásvizsgálatban. Az közismert, hogy a jogalkotás már most is alkalmaz statisztikai adatokat, hiszen például egy áfával kapcsolatos jogszabály-változtatás, vagy társasági törvény módosítás aligha volna lehetséges, ha nem tudjuk, hogy mennyi a kiskereskedelmi forgalom árukategóriánként, illetve hány kft. van az országban. A BD azonban messze túlmutat ezeken az adatokon. Itt arról van szó, hogy a társadalmi jelenséget reprezentáló adat legtöbbször *a maga teljességében* rendelkezésre áll, és nemcsak valamilyen közvetítő médiumon keresztül, szűrt, vagy kisebb mintán felvett formájában mutatkozik. Az online pénztárgépek adatai például a maguk teljes részletességében mutatják az áruk és szolgáltatások forgalmát, igaz, ömlesztett formában. A mobiltelefonok cellainformáció-adatai az emberek valós mozgását regisztrálják, nem torzítja őket az utólag visszaemlékező ember memóriája. A közösségi oldalakon zajló kommunikáció közvetlenül az emberi interakciókat mutatja, és nem utólagos visszaidézése azoknak.

A jogalkotást a BD többféleképp segítheti. Egyrészt egy-egy jogalkotási aktus, döntés hatásai a múltban mérhetők azokon a kimeneteken, amelyek egy életviszony (például a gazdaság) jellegzetességeit, adatait egyébként mikroszinten tartalmazzák. (Például egy gazdaságra vonatkozó jogalkotói döntés hatása a cégnyilvántartás változásaiban, vagy a társaságok letéti mérlegeiben.) Ezekből azután felállíthatók olyan összefüggések, amelyek a jövőbeli jogalkotás alapjául szolgálhatnak. Másrészt egy-egy társadalmi jelenség szabályozásakor a kiinduló adatok – amelyeket jelenleg vagy csak a múltra vonatkozó statisztikákból, vagy egyáltalán nem vesznek figyelembe a jogalkotáskor – a valós idejű és teljes formájukban vehetők figyelembe. (Például a kiskereskedelmi forgalom az online pénztárgépek adataiból, vagy a gépjárműforgalom az autópálya-kamerák adataiból.³⁵) Végül harmadszor: ahogyan azt fentebb már a téma egyik szakértőjének cikkére utalva jelez-

³⁴ A H2020 program egyik alprogramjaként született kiírás a „*Policy-development in the age of big data: data-driven policy-making, policy-modelling and policy-implementation*” című. Lásd ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/co-creation-06-2017.html.

³⁵ Vö. GILYÁN Csaba: Online pénztárgépadatok hivatalos statisztikai célokra való felhasználásának kérdései, tapasztalatai. Előadás a KSH és az MTA TK közös szervezésében megrendezett Big Data Workshopon, 2016. május 27-én, és Lovics Gábor: A turisztikai határfelvételek NÚSZ Zrt.-től érkező járműforgalmi adatok felhasználásával kapcsolatos tapasztalatok (előadás ugyanezen a rendezvényen).

tem, a BD rengeteg kísérletezésre, szimulációra ad lehetőséget.³⁶ Olyan modelleket lehet felállítani, amelyek nagy pontossággal képesek előre jelezni egy szabályozás hatásait, sőt olyan szabályozásokat lehet korlátozottabb kis mintákon (például csak egy kisebb közösség, vagy város lakosságán) tesztelni, amelyek azelőtt azért voltak lehetetlenek, mert az output (hatás) oldalán nem álltak rendelkezésre információk. Ma már olyan egészen „puha” hatások is mérhetők (például egy intézkedés bejelentésének társadalmi fogadtatása a közösségi oldalakon tett bejegyzések szöveg-elemzésével), amelyek korábban elképzelhetetlenek lettek volna.

Ugyanez igaz a jogalkalmazásra is, de a jogalkalmazás segítése már nemcsak elvi lehetőség, és ahogyan az lenni szokott, az üzleti szféra hamarabb látta meg a lehetőséget a BD-ben, mint a kormányzatok.

2008-ban a Stanford Egyetemen indult egy projekt, amelynek célja egy olyan rendszer kiépítése volt, amely a különböző nyilvános adatbázisokban található jogi dokumentumokat (javarészt döntéseket)³⁷ gyűjt össze és dolgoz fel szoftverek segítségével. A rendszer 2010-ben a *Lex Machina* nevet kapta,³⁸ majd 2015-ben a LexisNexis (a multinacionális Relx kiadóvállalat része) felvásárolta.³⁹

A rendszer, hasonlóan a versenytársához, a szintén stanfordi diákok által alapított *Ravel*-hez,⁴⁰ tehát nagy mennyiségű – elsősorban szöveges – adat elemzésére, és abból következtetések levonására képes. Az AALL (az Amerikai Jogi Könyvtárak Szövetsége) így lelkendezett, amikor nemrég 2016 legjobb termékének választotta a *Ravel* „Bíróelemző” algoritmusát: „a Bíróelemző felismeri azokat az eseteket, bírói fórumokat és bírákat, amelyeket egy meghatározott [a rendszer által elemzett – Z. Zs.] bíró a legmeggyőzőbbnek tart, és azonosítja azokat a szabályokat és speciális nyelvezetet, amelyeket a bíró általában idéz.”⁴¹ Ugyanez a termék ajánlásokat is tesz a leghatékonyabb perstratégiára, érvelésre. A *Lex Machina*t pedig azzal a kissé marketingizű szöveggel dicséri a *technolawyer* blog,⁴² hogy ha azzal dolgozunk, olyan, mintha légy volnánk az ellenérdekű fél képviselője irodájának falán, vagy a bíró szobájában. A *Lex Machina* képes elemezni a peradatokat bíróság, bíró, vagy az ellenfél képviselője szemszögéből. „Ez a tudás megfizethetetlen, ha a perstratégiáról, *forum shopping*ról, vagy egy fellebbezés megfogalmazásáról van szó” – mondja a reklámanyag. Az elemzés nemcsak táblázatok és grafikonok felrajzolását jelenti, hanem adott esetben a saját ügyünk kimenetelének százalékos arányban történő megjóslását a történeti adatok alapján. Az adatokat a

³⁶ BYERS (21. l.).

³⁷ Az amerikai District Courtok nyilvános online irattárából, a PACER-ből (www.pacer.gov/), az USA szabadalmi hivatalának adatbázisából, www.uspto.gov/ és az USA Nemzetközi Kereskedelmi Bizottságának (www.usitc.gov/) online dokumentumrendszeréből, az EDIS-ből, [edis3-external/page.svc?page=edis3Central%3AHome](http://edis.usitc.gov/edis3-external/page.svc?page=edis3Central%3AHome).

³⁸ Lásd lexmachina.com/.

³⁹ Lásd lexmachina.com/media/press/lexisnexis-acquires-lex-machina/.

⁴⁰ Lásd ravellaw.com/.

⁴¹ Lásd www.aallnet.org/hc/NewsCallout/Ravel-Law-Wins-2016-AALL-New-Product-Award.pdf.

⁴² Jennifer DIXON: „Review of Legal Analytics Platform” *Technolawyer blog* (Litigation World rovat) 2016. szeptember 23. Az írás a blogon közvetlenül nem érhető el, azt a *Lex Machina* oldala közli az alábbi linken: lexmachina.com/wp-content/uploads/2016/10/LitigationWorld-Review-2016.pdf.

rendszerek persze mindig „jogász-kompatibilis” formában találják: látványos vizualizáció (gráfok, háromdimenziós hálózatok, színes, de egyszerű grafikonok, diagrammok) segítségével.

A BD használata a jogalkalmazás bizonyos szféráiban tehát már megkezdődött. Ez a terület – a predikció lehetőségén kívül, amely mint mondtam, a BD általános jellegzetessége – szintén felvet egy izgalmas kérdést, mégpedig az adatok *értelmezésével* kapcsolatban. Már az egyszerű statisztikai adatok értelmezése is vehet fel számos kérdést, és ezeknek a kérdéseknek egy része érvényes a BD esetén is.⁴³ A BD új kérdései azzal függenek össze, hogy az adatok spontán, nem kutatási céllal képződnek. A statisztika ugyanis legtöbbször tudatos, előre megtervezett adatfelvételekkel operál. Így ott az adott társadalmi jelenség értelmezésének legnagyobb része már az adatfelvétel előtt eldőlt, hiszen el kell dönteni, miről és hogyan veszünk fel adatot. Ráadásul ez az adatfelvétel általában illeszkedik egy ismert, magasabb szintű narratívába. (Például az igazságügyi statisztika a bíróságok ügyszakai szerint, a Btk. büncselekményei szerint, azaz legtöbb esetben a jogszabályok által kijelölt normatív-fogalmi térben mozog.) A BD-jelenség esetén azonban az adatok mikroszinten, valós időben, spontán módon termelődnek hatalmas mennyiségben, és gyakran anélkül, hogy valaha valaki ezeknek az adatoknak az elemzés szempontjából történő feldolgozására gondolt volna. A BD-hez a magyarázatok már az adatok keletkezése után fűződnek hozzá, és ez felértékeli a BD-hez utólag hozzáfűzött narratívákat. Amikor a Ravel a bíró számára „nyerő” érveléseket, azaz bizonyos ügytípusok nyerteseinek érvelési mintáiban az ismétlődéseket kilistázza, ez egyben azt is feltételezi, hogy van egy olyan magasabb szintű narratívánk, amely szerint az érvelési mintázatok és a pernyertesség között valamilyen összefüggés mutatkozik. Ez ugyanúgy egy narratíva, mint az, hogy bizonyos ügyvédek bizonyos ügytípusokban sikeresebbek, mint más ügyvédek, jóllehet ennek lehet egy olyan egyszerű (másik) oka is, hogy a sikeres ügyvéd nem vállal el csak biztosnak látszó ügyeket.

Voltaképp ugyanennek a kérdéskörnek az egyik leágazása az is, amelyet a statisztika a korreláció fogalmában ragad meg, de a BD világában szintén kicsit másképp jelentkezik. A statisztikában is hozzá lehet bármely adathoz bármely más adat változását kapcsolni. Ugyanakkor a BD-korszakban annyiféle adat keletkezik, hogy egyre nagyobb a csábítás, hogy egymástól teljesen távol eső adathalmazok összefüggéseit vizsgáljuk meg. A fentebb már emlegetett *Target*-esetben éppen ez történt, és teljesen váratlan, és – legalábbis egy esetben – valós összefüggések mutatkoztak meg. De említhetnénk azt a kutatást, amely bebizonyította, hogy a konzervatív beállítódású emberek hajlamosabbak könnyebben undorodni bizonyos dolgoktól,⁴⁴ így az undorérzet és a politikai beállítódás között összefüg-

⁴³ Ismert az egyedi történés oksági és a nagyobb sokaságok statisztikai összefüggéseinek különbözősége: ha a fővárosba történő mozgás befelé hétköznapokon nagyobb, mint ünnepnapokon, akkor ezt legfőképp úgy lehet értelmezni, hogy „az emberek általában dolgozni járnak a fővárosba”, de természetesen ez nem jelenti azt, hogy hétköznap reggel a főváros felé tartó minden autóban kizárólag a fővárosban dolgozók ülnek.

⁴⁴ Yoel INBAR [et al.]: *Disgust Sensitivity, Political Conservatism and Voting*, *spp.sagepub.com/content/3/5/537.abstract*.

géseket találhatunk. Vajon hány ilyen rejtett összefüggésre fog még fény derülni az elkövetkező években, akár a jog területén is? De ez a téma már átvezet bennünket a következő fejezethez.

3. A BD MINT A JOGTUDOMÁNY TÁRGYA

Ahogy az a cikk elején említettem, a jogtudományi folyóiratok százas nagyságrendben közöltek az elmúlt években cikkeket a BD témakörében.⁴⁵

Mivel elemezhető jogi szabályozás még nincsen, a jogtudomány az etikai kapcsolódási pontokra igyekszik reagálni, alapelveket keres, illetve megpróbál ésszerű keretek között belelátni a jövőbe. Az etikai vizsgálatra a legjobb példa Richards és King cikke,⁴⁶ amely az egyik legtöbbször hivatkozott írás a Big Data tematikában, így érdemes röviden rekapitulálni a benne leírtakat.

A szerzők a cikkükben négy „magas szintű” alapelvet fektetnek le, amelyet szerintük a Big Data korszakában is követni kell. 1. Szerintük először is a „privacy” a mai korszakban az „információ kezelésére vonatkozó szabályrendszer” (*information rules*). Eszerint le kell számolni azzal a szabályozási előfeltevéssel, hogy magánszemély képes kontrollálni a személyes adatainak az áramlását. Ettől még nem lesz a „privacy halott”, ahogyan ezt sokan gondolják – egyszerűen csak arról van szó, hogy a személyes adatok védelme helyett talán inkább a személyes adatok kezelésének szabályain lesz a hangsúly. 2. A megosztott információ még maradhat bizalmas, a megosztás és a privacy között nincsen szükségszerű ellentmondás. A jelenlegi szabályrendszer bináris jellegű. A Big Data korszakában szakítanunk kell ezzel a bináris szemlélettel: ahhoz, hogy használhassuk és kihasználhassuk a BD előnyeit, az adatok megosztását elő kell segítenünk, és olyan szabályozásra van szükségünk, amely ezt úgy teszi lehetővé, hogy közben az adat bizalmas jellege megmarad. 3. További etikai alapelv, hogy a Big Data-nak, és főként az adatok mozgásának mindvégig transzparensnek kell maradnia – tudnunk kell, hogy mi történik az egyes adatok átadásakor, hogy a titkos adatmozgatásokkal ne tudjanak ellenőrizetlen hatalmat gyakorolni felettünk. 4. Végül a negyedik elv az identitás elve. A BD számos helyzetben képes korlátozni az identitás-meghatározási képességünket. Megmondhatja – úgymond –, hogy kik vagyunk, még azelőtt, hogy mi magunk megtehetnénk ezt. A legrosszabb, legveszélyesebb formáit ezeknek a gépi döntéseknek egyszerűen meg kell tiltanunk, a kevésbé rossz formáit pedig szabályoznunk kell – mondják a szerzők. Az utóbbi két pont kapcsán már csak az a kérdés – teszem hozzá már én –, hogy hogyan lehet ezeket a hatalmas adathalmazoz-

⁴⁵ Néhány cikk címe: „Big Data és a tisztességes eljárás: egy elméleti keret alapvonalai a prediktív személyes adatsértések kompenzálásához” [CRAWFORD–SCHULTZ (11. lj.)]; „Privacy a Big Data korszakban: itt az idő a nagy döntésekre” [TENE – POLONETSKY (4. lj.)]; „A Big Data és a prediktív ésszerű gyanú” [FERGUSON (7. lj.)]; „A betegek személyes adatainak védelme a Big Data korszakban” (NICOLAS P. TERRY: „Protecting Patient Privacy in the Age of Big Data” *UMKC Law Review* 2012/2. 385–416.) stb.

⁴⁶ RICHARDS–KING (5. lj.).

kat, valamint a bonyolult elemző és döntő algoritmusokat transzparenssé tenni egy egyszerű állampolgár számára.

A fentebbi gondolatokat kiegészíteném még kettővel, amelyek már kicsit a jóslás kategóriájába tartoznak. Mindkettő a BD predikciós képességével függ össze. Az első az, hogy a BD olyan predikcióra tehet alkalmassá bennünket, amely a jelenlegi világgépünk szerint jogilag és morálisan *igazolhatatlan következményekhez vezet*. Képzeld el azt a helyzetet, hogy bebizonyosodik, hogy a 46-os lábú, fekete hajú, 190 centiméter magas 30 éves férfiak nagyon nagy eséllyel követnek el erőszakos bűncselekményeket. (Szándékosan írtam ilyen abszurd paramétereket.) Az persze elképzelhetetlen, hogy preventív céllal ezeknek az embereknek a személyes szabadságát korlátozzuk, de egy idő után nagy lesz a csábítás, hogy legalábbis valamilyen módon például megfigyeljük őket. De még a megfigyelés sem lesz a hagyományos igazolási technikákkal igazolható. Hogyan magyarázom azt meg, hogy az intézkedés tartalmát semmilyen szempontból nem érintő paraméter az intézkedés indoka? Ez a gondolkodásmód teljes szakítás lenne a felvilágosodás eszményeivel: például azzal a megfontolással, hogy csak olyasmért vagyok felelőssé tehető, amelyen magam is képes vagyok változtatni. (Azon, hogy nagy a lábam, nem tudok változtatni.) Vagy képzeld el a nevezetes terroristadilemmát⁴⁷ a BD világban, úgy, hogy a gép 90% feletti valószínűséggel mondja majd rá valakire, hogy súlyos terrorcselekményt fog elkövetni. Maradjon tétlen a jogi gépezet ebben az esetben? Vagy tegyen valamit? És mégis, mit?

Ha jól belegondolunk, az állami garanciarendszertől távol már most is számos példa van a statisztikai összefüggéseken alapuló – nem feltétlenül, bár gyakran BD-alapú – döntésekre. Például ilyen a bankok részben „titkos” algoritmusok alapján végzett *scoring-ja*, azaz hitelengedélyezési eljárása, modelljei.⁴⁸ A bankok természetesen nem indokolnak a hitelkérelem elutasításakor, és részben épp azért, mert az indokok egy része elfogadhatatlan lenne a „felvilágosult” hitelfelvevőknek. (Például az, hogy ők kockázatos kategóriába tartoznak. Hiszen egyébként talán úgy érzik, hogy a kategóriájuk kockázatos, azonban ők maguk nem azok.) Ugyanígy: a biztosítótársaságok, amikor kockázati besorolást végeznek, akkor például a fizetendő díjat is determinálják ezzel. Én lehetek egy csendes, szelíd 20 éves fiatalember, aki sportkocsiját visszafogottan és lassan vezeti, a rendszer a külső paramétereim alapján mégis a legkockázatosabb kategóriába fog besorolni.

A másik érdekes lehetséges jövőbeli fejlemény szintén a BD prediktív erejéhez kötődik. A BD a jog hagyományos magatartásbefolyásoló eszközei helyett „irracionalis” magatartásbefolyásoló eszközöket hozhat létre: a „*social engineering*” új változatait. A *Target*-esetben az öntudatlan vásárlási minták „bukttatták le” a kis-

⁴⁷ A terroristadilemma lényege, hogy megmenthetünk-e több millió embert egyetlen ember megkínzása árán. Magyar nyelven a legjobb feldolgozását lásd JAKAB András: „Jogállamiság és terrorfenyegetés. Az alkotmány normatívitásának és az életmentő kínzás megengedhetőségének kérdése” in FEKETE Balázs – HORVÁTHY Balázs – KREISZ Brigitta (szerk.): *A világ mi magunk vagyunk...* *Liber Amicorum Imre Vörös* (Budapest: HVG-ORAC 2014) 253.

⁴⁸ PL. ORAVECZ Beatrix: „Credit scoring modellek és teljesítményük értékelése” *Hitelintézeti Szemle* 2007/6. 607–627.

mamát. Már korábban is sokszor bebizonyosodott, hogy mindennapi döntéseink 45%-át egyszerűen a megszokás és nem a tudatos döntéshozatal jellemzi,⁴⁹ ugyanakkor ezeket a rejtett döntéshozatal tényezőket ez ideig tudatosan inkább csak a marketing, a média világa használta ki. De a Big Data alapú döntések az üzleti és a politikai döntések világában is hamarosan bizonyára megjelennek. És a kérdés az, hogy a jog intakt maradhat-e ebben a világban, vagy lehetséges-e, hogy egyszer a jog is elfogad majd ilyen döntéseket? A jog egyelőre racionális motiváló eszköz, amely a szigorú erkölcsi individualizmus és a „morális okozatosság” elvén alapszik, de a kérdés az irracionális döntések előrelátása és befolyásolása lesz.

Nem lehet valakit előzetesbe tenni, vagy akár csak megfigyelni önmagában azért, mert egy predikációs szoftver 99%-ra teszi a bűnelkövetési hajlandóságát. Ugyanakkor például gondoljunk bele abba, ha a BD viselkedéstudományi alapon képes pontosan megmondani, hogy a nő terhes, vagy hogy engem fog-e érdekelni egy film, akkor például a büntetéskiszabás jelenlegi „önkényes” rendszerét nem lehetne-e BD-alapú, adatok milliárdjaiból dolgozó algoritmusokra bízni, és teljesen egyediesíteni? Vagy például sok polgári ügyben, ahol a bíró mérlegelési kategóriák alapján dönt (például a sérelemdíjak megállapításánál), miért ne lehetne olyan BD-algoritmusokat használni, amelyek az adott társadalmi jelenség környékéről, és a korábbi ítélkezési gyakorlatokból, esetleg a megállapított büntetések utóéletéről adatokat gyűjtve, pontosabban és jobban meg tudják mondani, mi az „épp jó” mérték.

Ha ez az utópisztikus világ megvalósul, annak több nagyon érdekes következménye lehet. Ezek közül kettőt emelek ki. Képzeld el, hogy a büntetéskiszabás BD-alapú. Ez esetben azonnal előtérbe kerül például az az egyszerű kérdés, hogy a speciális, vagy a generális prevenció a fontosabb. Hiszen lehetséges, hogy a számítások azt igazolják majd, hogy a speciális prevencióhoz (amely az elkövető adataalapú profilján nyugszik) egy egészen csekély büntetés is elegendő, míg a generális prevenció szempontjai azt mutatják, hogy a veszélyeztetett elkövetői rétegek számára egy jóval nagyobb büntetés bír visszatartó hatással. És adott esetben mindkét verziót teljesen egzakt számok támasztják majd alá. Egyszóval a BD-alapú döntések meghozatalához használt algoritmusok explicitté fogják tenni *az addigi implicit értékválasztásainkat*. A másik következménye pedig az lehet a BD-alapú döntéseknek, hogy mivel azok a normál ítélkezéshez képest jóval több – adott esetben az ember számára átláthatatlanul sok – paraméter alapján születnek, és így a hétköznapi és morális narratíváinktól eltérő döntéseket is eredményezhetnek, adott esetben például az ítélkezés *széttartó* voltát fogják majd eredményezni, és épp nem annak egységességét. Hiszen a BD-algoritmusok sokkal több egyéni körülményt figyelembe fognak tudni majd venni, mint a „normál” ítélkezés. Ugyanakkor a „jogegység” eszméje feltételezi a valamilyen szempontból homogénnek tekintett perszoportokat. Újra beleütközhetünk egy komoly filozófiai és értékdilemmába, amely szerint az igazságos döntés egyszerre reflektál az eset *egyedi és megismételhetetlen* vonásaira, és ugyanakkor ezzel egy időben engedelmeskedik a „hasonló ese-

⁴⁹ DUHIGG (8. l.).

teket hasonlóan” parancsának. A BD-döntések a „minden eset egyedi” predikción fognak alapulni. Ezek a predikciók talán olykor érthető (például az Amazon ajánlóalgoritmus), olykor pedig szinte érthetetlen (például a *Target* esete) narratívák, amelyeket egyedül a predikciós erejük igazol. Ugyanakkor az ember a saját, „leegyszerűsítő” morális narratívái között él, és e megélt narratívák alapján akar magyarázatokat a világ dolgaira. Egyelőre elfogadhatatlannak látszik, hogy a gép „jobban tudjon” valamit, és elfogadjunk egy általa hozott és a hagyományos morális mércéinkkel abszurdnak tűnő döntést.⁵⁰ Nem kizárt azonban, hogy ezek a narratívák lassan megváltozzanak, ahogyan például a tudomány új narratívát teremtett a döntésekhez – hiszen az egyre gyakrabban kirendelt szakértők épp ezt a narratívát erősítik. A „tudomány azt mondja, hogy [...]” – halljuk ma már szinte végső érvként. Ez valójában szintén ellenőrizhetetlen narratíva, és ezért a szakértő is kettős igazolási helyzetben van, hiszen nemcsak a kollégái, hanem a bíró és a felek előtt is igazolnia kell az álláspontját. Nem tudjuk mi lesz akkor, ha egy BD-ből leszűrt algoritmus hoz meg majd egy döntést, és az igazolása csak annyi lesz majd, hogy ez a megoldás az „esetek elsöprő többségében működik”.

4. A BD MINT A JOGTUDOMÁNYI KUTATÁSOK EGY LEHETSÉGES ÚJ MÓDSZERE

A BD bizonyára hatással lesz a jogtudományi kutatásra is, de az a feltételezésem, hogy kicsit másképp, mint a társadalomtudományokra általában. A jogtudományt ugyanis szokásosan két részre szokták bontani: a szűkebb értelemben vett „hagyományos”, vagy dogmatikai jogtudományra,⁵¹ és az empirikus, társadalomtudományi jellegű jogtudományra.⁵² Utóbbit nem is feltétlenül lehet egyértelműen „jogtudománynak” nevezni, hiszen valójában a szociológia, a statisztika, a politikatudomány stb. eszköztárával vizsgálja a társadalom egyik alrendszerét, vagy ennek egy darabját – a jogot, s így nyugodtan sorolható a politikatudomány, vagy a szociológia körébe. Erre a – gyakorta empirikus eszközökkel operáló, és leegyszerűsítve, összefoglalóan jogszociológiának nevezett – tudományra a BD-jelenség pontosan ugyanúgy fog hatni, mint a bármilyen más (empirikus) társadalomtudományra. Új kihívásként jelentkezik például, hogy az adatok előzetes tervezés nél-

⁵⁰ A dilemma a mesterséges intelligenciák által hozott (egyébként gyakran szintén BD-alapú) döntéseknél is nagyon hasonló. Ezekre a dilemmákra irányítja rá a figyelmet az MIT „morális gépe”, amely egy kérdőív, ahol egy önvezető autót irányító mesterséges intelligencia számára kell olyan preferenciákat megadnunk előre, amelyeket kiélezett válsághelyzetekben használhat majd a gép. Ilyen pl. hogy mi legyen a döntés, ha pl. az autó elkerülhetetlenül el kell, hogy üssön valakit, de dönthet, hogy egy gyereket vagy egy felnőttet üssön inkább el (*moralmachine.mit.edu*).

⁵¹ Lásd pl. BÓDIG Mátys: „A jogtudomány módszertani karaktere és a dogmatikai tudomány eszméje” in BÓDIG Mátys – ZÓDI Zsolt (szerk.): *A jogtudomány helye, szerepe és haszna; tudomány-módszertani és tudományelméleti írások* (Budapest: MTA TK Jogtudományi Intézet – Opten 2016) 86–109. vagy SZABÓ Miklós – JAKAB András: „A jogdogmatikai kutatás” in JAKAB András – MENYHÁRD Attila (szerk.): *A jog tudománya* (Budapest: HVG-ORAC 2015) 51–78.

⁵² Vö. pl. FLECK Zoltán – GAJDUSCHEK György: „Empirikus kutatás a jogban” in JAKAB–MENYHÁRD (51. l.) 101–131.

kül, spontán módon, nagy tömegben termelődnek és utólag kell hozzájuk narratívákat, magyarázó értelmezési kereteket keresni, valamint az, hogy az adatok közötti előzetes szelektálásra és az adatok validálására sokkal nagyobb hangsúlyt kell fektetni, mint eddig.

Számomra itt izgalmasabb kérdés az, hogy hogyan fog hatni a BD-jelenség a hagyományos, dogmatikai jogtudományra. Bár vitatott, hogy ennek a szűkebb értelemben vett jogtudománynak inkább a fogalmak köre szervezettsége vagy inkább az „interpretív” természete a meghatározója,⁵³ az biztos, hogy a jogtudománynak ez annyiban máig a „szíve”, amennyiben ez az, amely a legszorosabb kapcsolatban van a gyakorló jogászok munkájával. A dogmatikai tudomány egyszerre elemzi a gyakorlatot egy sajátos belső nézőpontból, azaz nem társadalomtudományi kategóriák, hanem jogi fogalmak segítségével, és szállít ugyanakkor érveket is a gyakorlat felé. Ennek a dogmatikai jogtudománynak tehát eltéphetetlenül szoros kölcsönkapcsolata van a joggyakorlattal. Ugyanez az empirikus jogtudományi kutatásokról nem mondható el.

Ez a hagyományos jogtudomány elsősorban autoritatív szövegekkel foglalkodik. A BD-jelenség ott fogja érinteni, hogy ezek a szövegek és a hozzájuk kapcsolódó, a jog mindennapjaiban keletkező másféle jogi szövegek korábban elképzelhetetlen mennyiségben fognak termelődni és rendelkezésre állni, és ezeknek a szövegeknek a feldolgozása, megértése, rendszerezése, vizsgálata a hagyományos jogtudományi módszerek mellett újfajta – főként gépi – módszerekkel is meg fog indulni. Statisztikai jellemzők, ismétlődések, szövegmintázatok, fogalmak és érvelési minták eloszlásai mentén már korábban is elemeztek jogi szövegeket,⁵⁴ azonban a BD-jelenség azt eredményezi, hogy ezek a szövegek mennyiségükben és típusaikban is jóval nagyobb számban fognak rendelkezésre állni, és a kutatásuk szinte kizárólag számítógépes eszközök, algoritmusok segítségével lesz lehetséges.

Nézzünk egy egyszerű példát: tegyük fel, hogy az a kutatási kérdésünk, hogy „hogyan változott a bírói gyakorlat az orvosi műhibaperek kapcsán”? A hagyományos jogtudomány módszereivel a témáról írott tanulmányokat és a bírói gyakorlat – nagyrészt felsőbbirósági részének – legfontosabb döntéseit dolgoznánk fel. Természetesen lehetőségünk lenne empirikus kutatást is folytatni, például ügyvédek, vagy bírák körében kérdőíves felmérést folytatni. Ebben a kutatásban akár kvalitatív elemek is lehetnének. Ami új, hogy a BD korszakában lehetőség lesz olyan dokumentumok (periratok és ítéletek *teljes* szövegállománya, kórházi dokumentációk teljes állománya stb.) gépi elemzésére is, amelyek eddig kiestek ezekből a vizsgálatokból. (Most elméletileg tekintsünk el a példában felvetődő adatvédelmi aggályoktól).

Hogyan folyhat egy ilyen kutatás a gyakorlatban? A módszer alapja csaknem mindig szöveges dokumentumok statisztikai alapú elemzése lesz. Az „elemzés” ez esetben ismétlődések, minták, szabályszerűségek feltárását jelenti, mégpedig – a

⁵³ Előbbit vallja pl. Jakab András, utóbbit Bódig Máttyás.

⁵⁴ Itt pl. a hivatkozás-statisztikai kutatásokra gondolok, amelyeknek a rövid történetét egy korábbi írásomban már áttekintettem. ZÓDI Zsolt: „Citations of Previous Decisions and the Quality of Judicial Reasoning” *Acta Iuridica Hungarica* 2015/2–3. 129–148.

kutatási célkitűzéshez képest – „strukturálatlan” adat- és dokumentumhalmazokon. Azonban ezek a szabadszöveges dokumentumok csak látszólag „strukturálatlanok”, hiszen valójában minden szöveg tartalmaz szabályszerűségeket (ha mást nem, egy nyelv szabályait), és a hivatalos iratok meglepően sok számszerűsíthető formában is kifejezhető, adattá konvertálható struktúrát tartalmaznak. Például egy sor számot (iktatószám, dátumok, címek), hivatkozást (jogszabály-hivatkozások, jogeset hivatkozások), különböző, a valóságban azonosítható személyeket és szervezeteket (magánszemélyek, intézmények neveit), eseményleírásokat (tényállási elemeket), utasításokat, deontikus mondatokat (a hatóság, a hivatal valamire utasítja a feleket vagy más entitásokat). A fenti, szövegesen is azonosítható entitások száma a végtelenségig szaporítható, és a BD-kutatások képesek lesznek majd ezeknek az entitásoknak és a hagyományos, dogmatikai tudomány tárgyát képező fogalmi és érvelési hálóknak a kölcsönös viszonyát és összefüggéseit feltárni.

Ezt a módszert nevezhetjük „szövegempiriának” is, és már korábban is ismert volt, elsősorban olyan formában, hogy egy nagyobb dokumentumhalmazban egy meghatározott szempontrendszer alapján emberi megértés, elolvasás után rögzítettek adatokat táblázatokban, rendszerint zárt lehetőséglistákból választva ezeket.⁵⁵ A BD korszakában egyre több dokumentumot fogunk így feldolgozni, immár gépekkel: a fenti példában a vizsgált iratokból kiemelhetünk olyan adatokat, mint hogy melyek az érintett intézmények, orvosok, mennyi ideig tartanak átlagosan ezek a perek, kik a jellemző pernyertesek és pervesztések, milyen betegség/kezeléstípusok esetén fordulnak elő és milyen gyakorisággal, stb. A dologban igazán izgalmas, hogy olyan összefüggéseket is találhatunk, amelyre a kutatás kezdetén még csak nem is gondoltunk.

Mivel magam is végeztem ilyen „szöveg-empirikus” kutatást egy nagyobb dokumentum-(ítélet)halmazon, itt a cikk végén az ebből levonható – és egyáltalán nem biztosan minden területre általánosítható – következtetéseimet szeretném megosztani.⁵⁶

Kutatásom 2012-ben 62 000 magyar bírósági ítéletben található korábbi jogesetekre történő hivatkozás elemzése volt. Ez mind a dokumentumok, mind a hivatkozások számát, mind pedig a vizsgált jelenséget összességében leíró adatok számát illetően valójában inkább „*small data*”-kutatás volt.⁵⁷ Ami miatt mégis érdekes lehet a tapasztalatokról beszélni, az az, hogy ez volt az első olyan vizsgálat, amelynek a *teljes* közzétett bírósági határozathalmaz volt a tárgya.

⁵⁵ Ilyen a híres *Spaeth*-adatbázis, az amerikai Legfelsőbb Bíróság ítéleteit 247 paraméter alapján feldolgozó adatbázis (The Supreme Court Database [2016]), *scdb.wustl.edu*), amely az amerikai politikatudományi irodalomban az utóbbi évtizedekben szinte megkerülhetetlen lett.

⁵⁶ A kutatás eredményei megtalálhatók Zódi Zsolt: „A korábbi esetekre történő hivatkozások mintázatai a magyar bíróságok ítéleteiben” *MTA Law Working Papers* 2014/1, jog.tk.mta.hu/uploads/files/mtalwp/2014_01_Zodi_Zsolt.pdf.

⁵⁷ Az összes vizsgált *hivatkozó* dokumentum száma 61 500 db volt. A hivatkozó dokumentumokat leíró adatok száma 6 metaadat volt (bíróság, ügyszak, évszám, eljáró bíró és ügy tárgya). A hivatkozott dokumentumok száma 35 118 db volt, ezek 3 metaadattal voltak leírhatók (évszám, sorszám, típus). Az összes hivatkozó dokumentum száma 19 933 db, az összes hivatkozás száma pedig 41 678 db volt. Ezek a mennyiségek bőven kezelhetők még akár egy Excel-táblázatban is.

Az első tapasztalatom az volt, hogy a legszabályosabbnak tűnő adathalmaz is tartalmaz teljesen váratlan eltéréseket, szabálytalanságokat. Ha egy teljes dokumentumhalmazt vizsgálunk, mindig lesznek benne megmagyarázhatatlan előfordulások, kivételek, anomáliák. Ezeket a kivételeket, anomáliákat nagyon óvatosan kell kezelni, mert előfordulhat, hogy épp ezek mutatnak valamilyen mélyebb szabályszerűséget. A másik tapasztalat azt mutatja, hogy minden jól magyarázható trend mellett mutatkoznak ezekben a kutatásokban teljesen megmagyarázhatatlanok is. Ilyen volt például a kutatásban a menekültügyi közigazgatási perekben tapasztalható „precedenshasználat”, amely egy sablon-ítéletsszövegeknek volt köszönhető.⁵⁸ Végül a legfontosabb tanulság az volt, hogy egy hagyományos „kvitatív” kutatásnak mindig ki kell egészítenie a kvantitatív kutatási részeket, és nemcsak, sőt nem is elsősorban azért mert a kvantitatív „tévedne”, hanem elsősorban amiatt, mert csak a kvitatív, azaz a hagyományos fogalmi keretekben mozgó, a nyers számokat és összefüggéseket a hagyományos narratívákba visszahelyező eredményekkel tudunk egyáltalán valamit kezdeni. A saját kutatásomra utalva, annak szinte minden száma afelé mutatott, hogy a korábbi ítéletekre történő hivatkozások száma növekszik, így a magyar jogrendszer valamiféle precedens-rendszer felé halad. Ugyanakkor az ítéletek véletlenszerű elolvasása mutatta meg, hogy ez csak egyszerű mennyiségi növekedés, mert az ítéletek érvelésében nem mutatkoznak azok a mintázatok, amelyek a precedensjogokra jellemzők. (Az ún. tény-intenzív⁵⁹ érvelési mintázatok.)

Részben ugyanez a kutatás, illetve ennek egy szála mutatta meg azt is, hogy a bíróságok érvelései tele vannak olyan mintázatokkal, amelyeknek nincsen jogszabályi alapjuk, hanem valamiféle közös előismeret-készlet, a „népi dogmatika” részei.⁶⁰ Úgy vélem, hogy ezen a ponton kapcsolódhat össze a dogmatikai tudomány a BD-alapú vizsgálatokkal. A BD-alapú kutatások a szövegekben tükröződő jogi mindennapokat egészen pontosan képesek megmutatni, hiszen nem viszi el azokat az adatfelvétel során óhatatlanul bekövetkező torzítás. Ugyanakkor a jogtudósok a mérhető paraméterekkel operáló kutatásokat még mindig ellenérzéssel figyelik.⁶¹ Ennek az egyik oka biztosan az, hogy az adatok alapvetően tekintélyellenesek, míg a dogmatikai érvelésnek mindig is fontos eleme volt a tekintélyre hivatkozás. Az elkövetkező években bizonyosan több adat-, és „nagy” adat alapú kutatás is le fog zajlani a jogban, amelyek a dogmatikai fogalmak és konstruk-

⁵⁸ ZÓDI (56. l.) 39.

⁵⁹ A „tény-intenzív” érvelés az alapul fekvő és a precedensekben található tényállások hasonlóságainak és különbségeinek gondos összevetésén, ezek használatán alapszik. Lásd JAN KOMÁREK: „Reasoning with Previous Decisions: Beyond the Doctrine of Precedent” *The American Journal of Comparative Law* 2013/1. 149–171.

⁶⁰ ZÓDI Zsolt: „A jogirodalomra történő hivatkozás a bíróságok ítéleteiben; a jogtudomány és a népi dogmatika” in BÓDIG–ZÓDI (51. l.).

⁶¹ Ennek okait részletesen elemzi JAKAB András: „Jogi érvelések és dogmatikai jellegzetességek összehasonlítása számokkal – jog-összehasonlítás, matematika és jogbölcsélet” in HEGYI Szabolcs – VINNAI Edina – ZÓDI Zsolt (szerk.): *Orbis Iuris – Ünnepi tanulmánykötet Szabó Miklós 65. születésnapjára* (Miskolc: Bíbor 2016). Ugyanakkor van egy nagyon is erős „empirikus” szála a jogtudománynak, sőt ez erősödni is látszik. Lásd pl. a Wiley kiadó által megjelentett *Journal of Empirical Legal Studies* c. szaklap újabb cikkeit.

ciók valóságos használatát és hatékonyságát, sikerességét fogják mérni. Ez pedig azt fogja eredményezni, hogy ezek a kutatások nem felváltani fogják a hagyományos dogmatikai jellegű tudományos teljesítményeket, hanem az egyes dogmatikai konstrukciók, a fogalmak értelmezésének valódi, statisztikai, adatalapú használatáról lesznek képesek visszajelzéseket adni. Hogy ezt a dogmatikai jogtudomány hogyan lesz képes beépíteni, és ezek majd oldják-e a jogászok ellenérzéseit, az már egy másik kérdés, és nem ennek a tanulmánynak a tárgya.