

Bátorfy Attila

## Adat, társadalom és politika

### Bevezetés a kritikai adatkutatásba

**Absztrakt:** A tanulmány célja az, hogy megismertesse a magyar olvasót a több mint egy évtizede létező, a Big Data kihívásaira válaszul létrejött kritikai adatkutatás (*critical data studies*) történetével, alapvető kérdéseivel és problémafelvetéseivel, módszereivel, alkalmazhatósági területeivel, valamint ellentmondásaival. Ezen túl a szöveg egy nagyobb, általában véve a társadalomtudományok kvantifikálásával kapcsolatos kérdéskörbe illeszti be a kritikai adatkutatás problémafelvetéseit.

**Kulcsszavak:** kritikai adatkutatás, Big Data, datafikáció, ismeretelmélet, adatmitológia

## Bevezetés

Miközben a *Replika* kritikai adatkutatásról<sup>1</sup> (*critical data studies*) szóló számát szerkesztjük, a világot már két éve a mesterséges intelligencia (MI) körül kialakult viták izgatják. Ez már abból a szempontból is fontos számunkra, hogy miközben az MI kritikái részben a kritikai adatkutatás által felvetett problémákban és állításokban gyökereznek, utóbbiak szinte észrevétlenül maradtak a hazai tudományos irodalomban<sup>2</sup> és közbeszédben is, jóllehet a felvetett kérdésekkel az adatokkal foglalkozó társadalomtudósok a munkájuk során megannyiszor találkozhattak. Röviden tehát, ez a tanulmány, ahogy maga a kötet is, mint olyan sok minden más is Magyarországon, jelentősen megkésett. Ennek okai összetettek, és nem is tartoznak a tanulmány szorosán vett témájához. Ugyanakkor késni még mindig jobb, mint el sem indulni, ezért meg sem érkezni, ebből fakadóan ennek a tanulmánynak a célja az, hogy magyar nyelven először adjon átfogó képet a Big Data jelenségre reagáló kritikai adatkutatás rövid történetéről, alapvető megállapításairól, irányairól, módszereiről és ellentmondásairól. Mindezt azonban annak figyelembe vételével, hogy a Big Data egy olyan szociokulturális környezetbe robbant be, amelyben egyfelől a politika és a társadalom is John Ziman (1996) által megfogalmazott poszt-akadémiai tudást, tehát az üzleti élet által is finanszírozott, megoldásközpontú, profitorientált, rövid távú tudástermelést várta el a tudományos intézményrendszertől, másfelől „a mérés zsarnokságának” (Muller 2018) fennhatóságát már évtizedek óta a természettudományokon kívüli területekre is ki akarta terjeszteni. Erre azért fontos felhívni a figyelmet, mert a Big Data csak egy következő, bár jelentős léptékű váltás az emberi tevékenység adatosításának történetében, és a vele szembeni fenntartások számos esetben korábról is ismerős ismeretelméleti, hatalom- és intézménykritikai, valamint tudományszociológiai felismerések. Tisztázandó tehát, hogy ebben az összefüggésben a tanulmány két síkon mozog: apró-pója konkrétan a Big Data mint a mindennapi életünket is alakító tudományos és politikai paradigma társadalomkritikai elemzéseinek összegzése. Ugyanakkor a fentiekből adódik, hogy a kritikai adatkutatás egy létező, sokszor nem is kritikai hagyományt gördít tovább, így a tanulmány szeretné egy nagyobb történeti elbeszélésbe is beilleszteni ezt a megközelítést.

---

1 A *critical data studies*-t a *cultural studies* kritikai kultúrákutatásra való fordításának (Vörös és Nagy 1995: 153–154) mintájára fordítjuk kritikai adatkutatásra. A cikk címe is Vörös és Nagy akkori címére való utalás.

2 A *Replika* 2015-ben egy külön blokkot szentelt a Big Data társadalomtudományi hasznosíthatóságáról szóló hazai reflexióknak. Ezek között Z. Karvalics László, Németh Renáta és Székely Iván cikkei is foglalkoznak azokkal a problémákkal, amelyek a kritikai adatkutatás problémafelvetéseihez is kapcsolódnak (Dessewffy és Láng szerk. 2015).

## Előzmények: adatkritika a kritikai adatkutatás előtt

A kritikai adatkutatás, amely diszciplína keletkezése a Big Data révén időhöz kötött<sup>3</sup>, jóval korábban létezett annál, minthogy két geográfus, Craig Dalton és Jim Thatcher először nevet adott volna neki (Dalton és Thatcher 2014). Azok az ismeretelméleti problémák<sup>4</sup>, amelyek a világ megméréséből fakadnak, már a felvilágosodás során artikulálódtak, ám ezek az episztemológiai, továbbá a keletkezett adatok hasznosításából származó szerteágazó tudományfilozófiai és társadalmi problémák sokáig rejtve maradtak, és részben olyan tudásarcheológusoknak kellett ezeket újra felszínre hozniuk, mint Michel Foucault, Mary Poovey, Lorraine Daston, Theodore Porter, David Wootton vagy Daniel Rosenberg. Mondhatnánk azt is, hogy a hasznosításukból származó hatalmas és gazdasági előnyök ezeket a problémákat eleve elrejtették, és talán nem is járnánk messze az igazságtól. Főként, ha figyelembe vesszük, hogy a modernitás a mérésekből keletkező adatoknak köszönhetően olyan új, gyakran ünnepezt szavakkal és fogalmakkal ajándékozta meg az emberiséget, mint a növekedés, a sztenderdizáció, a tervezés, a hatékonyság, a menedzsment, a racionalizálás, automatizálás, az átlag, a bürokrácia vagy éppen a normalitás. A felvilágosodás kezdeti ígérete Daniel Kehlmann bogaras és házsártos Humboldtja<sup>5</sup>, aki az ész elleni merényletnek tartotta, ha egy dombról nem lehetett tudni, hogy milyen magas. Ezzel szemben a felvilágosodás realitása Adolphe Quetelet *l'homme moyenje*, vagyis átlagembere.<sup>6</sup> Az eredetileg csillagász Quetelet a társadalom törvényszerűségeit a fizika törvényeihez hasonlóan nem csupán vizsgálni, hanem megteremteni is akarta. A társadalmi fizika (*physique sociale*), ahogy a dühös Comte-tól kölcsönvéve elnevezte a matematikai képleteken és az adatok elemzésén alapuló társadalomstatisztikát, célja Quetelet és követői számára nem volt más, mint egyfelől a különféle testi mérések és társadalmi statisztikák által megismerni a statisztikai hibáktól, vagyis a fizikai és mentális extremitásoktól, devianciáktól, abnormalitástól mentes embert, másfelől pedig ezen ismérvek alapján a továbbiakban előállítani őket. Quetelet számára ez a biopolitikailag meghatározott átlagos ember a (nemzet)állam szervezésének, irányításának és tervezésének kulcsa, és a különféle társadalmi vadhajtások, mint például az *erkölcstelenség* lenyesegetésének alapja (Porter 1986: 52–55). Ennek az átlagra és normalitásra épülő antropometrikai elvnek – vagy James Vincent (2022: 222) megfogalmazásával élve, „statisztikai Frankensteinnek” – lesz később az eredménye Francis Galton szörnyszülöttje, a normál eloszlásgörbére épülő

3 Bár újabban minden sokkal korábban kezdődött, mint gondolnánk. Asheesh Kapur Siddique például legújabb könyvében a Big Data-alapú állam alapjait a brit gyarmatbirodalom adminisztrációjának megszervezésében látja (Siddique 2024). Frédéric Kaplan és Isabelle di Lenardo még tovább mennek, amikor azt mondják, hogy az adat nagysága relatív, ezért a történelem során Mezopotámiától kezdve számos „Big Data momentuma” volt az emberiségnek, amelyek jellemzője, hogy a korábbiakhoz képest rendkívüli mértékben és gyorsasággal nőtt meg az akkumulált adatmennyiség (data acceleration regimes) (Kaplan és di Lenardo 2017).

4 És nem csak ismeretelméleti problémák, hanem konkrét szóhasználati, fogalmi és alkalmazhatósági problémák is, lásd (Dusek 2024).

5 Daniel Kehlmann magyarul *A világ fölmérése* címmel megjelent könyve a német természettudós Alexander von Humboldt és a matematikus Carl Friedrich Gauss fiktív, bár valós elemeket is ötvöző, szórakoztató életrajzi regénye.

6 A francia *moyen* valójában közepet jelent, lásd például: *moyen âge* = középkor. A statisztikai szaknyelv az átlagot középértéknek is hívja, az egyszerű átlagot számtani/aritmetikai középnek.

eugenika is (Cryle és Stephens 2017: 212–260; Vincent 2022: 229–234), valamint számos egyéb, az ember és teljesítményének félreméréséből következő borzalom és félreértés (Gould 1996[1981]; Muller 2018).

De miért tettünk egy kiterőt Quetelet-nél? Az átlagember, akár fizikai, akár *morális* értelemben használjuk, statisztikai módszerekkel való megmérése, valamint előállítás és műveltető igekézpövel: előállítása sokat elárul számunkra a különféle adattípusok természetéről. Idézzük fel ismét Kehlmann Humboldtját, mint Quetelet képzeletbeli antagonistáját. Kehlmann regényében Humboldt, akármennyire is karikatúraszzerű, a megismerés örömeinek metaforikus alakja. Humboldt a nap huszonnegy órájában kísérleteket és méréseket végez, árammal rázatja meg magát, kiköti magát a hajóorra, bemászik a vulkán kürtőjébe, fura és veszélyes dolgokat nyel le, és szinte elalél a párizsi méter bevezetésétől. Eszközei különféle mérőeszközök, amelyeket nem csupán a pontosabb mérések miatt hoztak létre, hanem azért is, hogy kiküszöböljék az ember érzékszerveinek tapasztalati csorbitásait. Ezzel szemben a német statisztika (*statistik*), vagy akkori angol megfelelője, a politikai aritmetika (*political arithmetic*)<sup>7</sup> mint eredetileg az állam tudománya<sup>8</sup>, mesterséges, a természetben „nem létező”, a kérdező pozíciójától, érdeklődésétől és érdekétől függő adatokat hoz létre az emberi tevékenység kvantifikálásával. Éppen ezért nem mindegy, hogy ki kérdez, milyen szándékkal kérdez, mit kérdez, mit tart fontosnak tudni, mit és hogyan használ fel ebből a tudásból, és ebből mit tesz közkinccsé és mit titkol el?

Az ezekről a kérdésektől szóló tudományos és népszerű diskurzus mindig is szélsőséges volt, és számos egyéb társvitához kapcsolódott. A Big Data megjelenése előtti nagyívű elbeszélések azonban, ha különböző mértékben is, de mintha mind kissé csalódtak lettek volna. Ha *A felvilágosodás dialektikájának* adatfóbiáját egyfajta kezdővonalnak tekintjük, akkor a következő nemzedékek sorra kitermelték a maguk kritikus olvasatát az adatokra építő társadalommérnökösödésről. Horkheimernél és Adornónál a felvilágosodás logikájából adódó adatok nyelve a homogenizáció, az uniformizáció, az atomizáció, a kommercializáció, a szómágia és – végső soron – a totalitarizmus nyelve.

Azoknak az adatoknak a vaksága és némasága, amelyekre a pozitívizmus a világot redukálja – írják – átterjed magára az adatok regisztrálására szorítókozó nyelvre. Így áthatolhatatlanná válnak maguk a megnevezések is, olyasfajta ütőerőre, a vonzás és taszítás olyan vonzerejére tesznek szert, amely hasonlóvá teszi őket extrém ellentétükhöz, a varázsigékhez (Horkheimer és Adorno 2020: 203–204).

Az olyan, kettejüknél kevésbé borútapasztalt szerzők, mint Foucault, James R. Beniger, Theodore Porter, a nemrég elhunyt James C. Scott vagy Susan Buck-Morss történeti elemzései a felvilágosodás és a modernitás vészjóslóan ugyanazon irányba tartó ösvényeire kalauzolnak el bennünket: az ellenőrző-tervező államot kiszolgáló egyszerűsítő,

7 A német(es) statisztika és az angol politikai aritmetika között lényegi eltérés volt. Míg előbbi az állam általános leírását tekintette céljának, utóbbi a különféle, először lokális demográfiai adatok matematikai elemzését. Érdemes megjegyezni, hogy noha a 19. század első felére Európa nagy részén a politikai aritmetika kiszorította a német államleíró statisztikát, a politikai aritmetika viszont pont a statisztika javára kiszorult a szóhasználatból (Desrosières 1998: 16–19).

8 A statisztika szó eredetileg az állam tisztségviselőire használt olasz *statista* szóból származik. A tudomány létrejötté eredetileg nem a társadalom megismerésének vágyából fakadt, hanem az állam javainak, gazdagságának felmérési igényéből és az állam jobb megszervezésének céljából.

sztenderdizáló, normalizáló-normativizáló, homogenizáló tudások, rendszerek és osztályozások felé, melyeket sűrűn vett adatok és információk szolgálnak ki, és amelyek az emberi élet minden aspektusát behálózzák az igazságosságnak, hatékonyságnak maszkírozott kontroll révén. Ez Foucault *Felügyelet és büntetésének* alap gondolata (1990 [1975]), a mások megfigyelésére épülő, egymást erősítő tudomány és hatalom párosából fakadó panoptikus és normalizáló felügyelet, mint hatalomtechnológiai rendszerének eredője. Beniger részben ugyanezt a folyamatot elemzi a másokhoz képest visszafogott *Az irányítás forradalmában* (2004[1988]) az információs társadalom kialakulásával párhuzamosan, rámutatva az államirányítás kríziseire adott információtechnológiai válaszok kétarcúságára. Beniger 19. százada, ugyan más elemzési síkon, de hasonlít Theodore Porter 19. századára. Porter (1995) szerint az objektivitást mániákusan hajszoló, a számok iránt vakbuzgó bizalmat tanúsító 19. század eredménye a biológiai és adminisztrációs sztenderdizáció, a számvitel-kultúra, a bürokraták rendje és a társadalomtudományokra egyre jobban ránehezülő kvantifikációs nyomás. James C. Scott a *Seeing Like a State*-ben (1998) azt írja, hogy minden modern állam célja a társadalom és az állampolgárok sematikus és leegyszerűsített olvashatóvá tétele, mivel ez az államszervezés és tervezés alapja. Ennek ellenére Scott történelmi példáinak (erdők és földművelési területek parcellázása, szovjet kolhozok és tervgazdálkodás, a racionális várostervezés) tanulsága az, hogy minél nagyobb a társadalom és az állampolgárok igényeinek sematikus absztrahálása, vagyis az olvashatóságuk érdekében tett leegyszerűsítése, a hozzá rendelt „jobbító”, társadalommérnöki szándékok annál nagyobb katasztrófát idézhetnek elő. Susan Buck-Morss (1995) szerint pedig az adatok és azok képei, vagyis a grafikonok, diagramok, adattérképek a 18. századtól kezdve a kapitalizmus ideológiájának ágensei, és mint ilyenek, maguk is hozzájárulnak például a társadalmi és regionális egyenlőtlenségek újratermeléséhez.

Tulajdonképpen meglepő, hogy a kritikai adatkutatás körébe sorolható szövegtermés a fentebbi előzményekről mintha tudomást se venne, noha problémáinak egy jelentős része azonos. Mindazonáltal érthető is, hiszen a kritikai adatkutatás a fentebbieken túl egy nagyon is konkrét jelenségre, a Big Datára, vagyis a hatalmas és gyorsan növekvő méretű, sokféle, mindenre kiterjedő, részletes, nagy felbontású, számítógépekkel és szoftverekkel feldolgozható, kezelhető, kereshető, elemezhető és egymással összeköthető, relációs adataszsemblázsok politikai, társadalmi, kulturális kockázataira (Kitchin és Lauriault 2014; Iliadis és Russo 2016) adott reakció.

### A kritikai adatkutatás színre lép: a Big Data ígéretei

A kritikai adatkutatás nyitószövege danah boyd és Kate Crawford 2012-es *Critical Questions for Big Data* vitaindító esszéje.<sup>9</sup> A szerzők ebben a cikkben lényegében a nem sokkal később kritikai adatkutatásnak elnevezett részdiszciplína már minden fontosabb kérdését felteszik. boyd és Crawford összes Big Datára vonatkozó kérdése minden tudományos-technológiai innováció alapkérdése is: kiknek lesz jó, kiknek válik hasznára, kik

---

9 Magyar nyelven *Az adatrengeteg kínos kérdései: Vitaindító egy kulturális, műszaki és tudományos jelenségről* címmel is megjelent Balogh Dániel fordításában (boyd és Crawford 2012a, 2012b).

fognak belőle profitálni? A szerzők a Big Data jelenségét három dimenzió, a technológia, az analízis és a mitológia kölcsönhatásaiként elemzik, és több provokatív, részben gyakorlati és episztemológiai, részben etikai állítást fogalmaznak meg.

Itt érdemes egy időre a mitológiánál megállnunk. Ahogy Horkheimer és Adorno rámutat az adatokról szóló beszéd nyelvet és elmét megbabonázó paradoxonára, a Big Data örömhírének terjesztői sem fukarkodtak sohasem a mágikus ígéretekkel. Chris Anderson, a *Wired* geek-magazin akkori nagy hatású főszerkesztője<sup>10</sup>, nem mellékesen jelenleg egy drónirányítási szoftvereket gyártó cég alapító-ügyvezetője, 2008-ban azt ünnepelte, hogy az adatáradat az elméletek, valamint a tudományos módszerek végét fogja jelenteni. Az akkor, ma már régimódinak tetsző *Petabyte Age*-nek hívott korszak jelentőségét Anderson abban látta, hogy szimpla statisztikai és matematikai elemzési modellek alkalmazásával teljesen lényegtelenek lesznek az ok-okozatokra, jelentésekre vonatkozó kérdések, és az emberi viselkedésre, nyelvhasználatra, társadalomra vonatkozó elméletekre nem lesz többé szükség.

Felejtjük el a taxonómiát, az ontológiát és a pszichológiát. Ki tudja, miért csinálják azt az emberek, amit csinálnak? A lényeg az, hogy csinálják, mi pedig soha korábban nem tapasztalt pontossággal tudjuk követni és mérni. Ha elég adatunk van, a számok magukért beszélnek (Anderson 2008) – írta.

Anderson ugyanígy feleslegesnek jövedőlt a hipotézis-modell-teszt szentháromságra épülő absztrakt tudományos módszereket is. Anderson szerint ugyanis a tudomány nagy kérdései egymással versengő modellekké és elméletekké váltak, amelyek nem egyebek, mint adatokra éhező „csodás történetek”, de bizonyíthatatlanok és falszifikálhatatlanok. Anderson szerint azonban ez a szöszmötölés értelmetlenné válik, mert érdektelen azt a kérdést feltenni, hogy *miért* van dolgok között összefüggés, elég csak megállapítani, hogy összefüggés van közöttük. A nagy adat elemzése ugyanis ott is kimutatja ezeket az összefüggéseket, mintázatotokat, ahová a tudományos módszertannak esélye sincs bepillantani. Anderson 2008-as cikkét a kritikai adatkutatás általában malíciával idézi fel, mint a 2000-es évek végi hurraóptimizmus különösen gögös példányát. De hasonló áhitattal beszélt akkoriban a nagy adatról a digitális bennszülött (*digital native*) fogalmat is megalakító Marc Prensky (2009) is. A Big Data fogalmának sikeréhez természetesen az üzleti és technológiai sajtó metaforái is hozzáadták a magukét (Puschmann és Burgess 2014). A rajongás a 2010-es évek elejére olyan hevültté vált, hogy az adatok mindenhatóságát hirdető, már-már a miszticizmus határát súroló filozófiát David Brooks 2013-ban dataizmusnak, követőit pedig dataistáknak nevezte (Brooks 2013). Igaz, a Big Data-forradalom hívei maguk is szívesen nevezik szemléletüket dataizmusnak (Lohr 2015).

Boyd és Crawford sem hagyta annyiban Anderson és a hozzá hasonlók ortodoxia felé mutató arroganciáját, amelyet David M. Berry (2011: 12) ontoteológiának nevez, utalva arra, hogy a Big Data apologetái egy olyan digitális számszerűsíthetőségben (*computationality*) hisznek, amely révén olyan destabilizáló mennyiségű tudás és

---

<sup>10</sup> Chris Anderson amerikai újságíró és vállalkozó, 2001 és 2012 között az amerikai *Wired* magazin főszerkesztője volt. A 2006-os bestsellere, a *The Long Tail: Why the Future of Business is Selling Less of More* magyarul *Hosszú farok: a végtelen választék átírja az üzlet szabályait* címmel jelent meg 2007-ben.

információ termelődik, amelyből hiányzik a filozofikus gondolkodás szabályozó ereje. Éppen ezért Boyd és Crawford fentebb említett provokatív pontjai azokkal az állításokkal szembesíti a Big Data-kultúra veretes állításait, amelyek jobb esetben félrevezetők, rosszabb esetben igazolhatatlanok. A Big Data Nagy Ígérete ugyanis az (volt) – mintha csak Comte, Lord Kelvin vagy Karl Pearson a pozitív tudomány objektivitasáról szóló áhítatos ódáit olvasná az ember –, hogy az elérhető nagy mennyiségű adat objektívebb, pontosabb, jobb minőségű, semlegesebb, kontextus nélkül is értelmesebb és használhatóbb tudást hoz létre, mint a tudáselőállítás hagyományos tudományos módszerei. Azáltal pedig, hogy a Big Data kiiktatja az emberi tényezőt és az elméletet az elemzési keretből, az emberiség ésszerűbb, hatékonyabb, olcsóbb döntéseket tud meghozni, amelyek közjószágként hozzájárulnak például az állami alrendszerek jobb megszervezésétől kezdve, a társadalmi egyenlőtlenségek csökkentésén át a korábban remélni sem mert hatékonyságú gyógyszerek kifejlesztéséig. Alighanem ezekkel állításokkal a *Replika* olvasói is találkozhattak már számos társadalomtudományi, adattudományi, marketinges és digitális bölcsészeti konferencián, és többen el is csábulhattak. Vannak ugyanis a Big Datának olyan aspektusai, amelyek a társadalomtudományok számára módszertanilag és a teória szempontjából is különösen vonzóvá tehetik. A 19. század vége óta küzd a társadalomtudomány azzal a problémával, hogy nem tud egyszerre minden embert vizsgálni, ahogy a fizikus teszi az atomokkal, vagy a biológus a molekulákkal. Még a legnagyobb minták is a társadalmak elenyésző százalékát teszik ki, és ezekből a mintákból kellene általános, törvényerejű következtetéseket levonnia a társadalomtudománynak, sőt, a természettudományokhoz hasonlóan predikciókat megfogalmaznia a társadalom csoportjainak viselkedéseire, döntéseire vonatkozóan. Quetelet és Galton, és nyomukban a 19. század megannyi statisztikusa és szociológusa, ha néha gyökeresen eltérő pozíciókból is, de meg voltak győződve arról, hogy a társadalmakat matematikailag leírható, a természettudományokhoz hasonló törvényszerűségek hozzák létre, mozgatják és szervezik. A Rész és az Egész közötti űrt különféle statisztikai-matematikai módszerekkel igyekeztek kitölteni, és ezek a módszerek sokszor megfelelőnek, sokszor elégtelennek, máskor félrevezetőnek bizonyultak. A Big Data mérete, egyidejűsége, skálázhatósága, relációssága látszólag kiiktatja ezt a szakadékot a Rész és az Egész között azzal, hogy egyszerre akár több százezer/millió ember több száz paraméterét is elemezhetővé teszi, amelyekből olyan új, finom információkhoz, összefüggésekhez is hozzájuthat a társadalomkutatás, amelyekről korábban csak álmodni mert, vagy még azt sem, mert nem tudta, hogy lehet róluk. Boyd és Crawford azonban úgy érvelnek, hogy még a több millió közösségimédia-felhasználó adatait elemző Big Data-kutatások is csupán egy méretükben jelentős, ám partikuláris mintán történnek meg, amelyek számos olyan módszertani problémát vetnek fel, amelyekre a kutatók ritkán reflektálnak, mégis a következtetéseiket úgy állítják be, mintha általában az emberekre vonatkoznának (Boyd és Crawford 2012: 670–672). Duncan J. Watts ugyanakkor ezt az átalakulást a csillagászat és a kémia kétszáz évvel ezelőtti forradalmához hasonlította, és azt állította, hogy ésszerűtlen amellet károskodni, hogy az emberi és társadalmi viselkedés a Big Data módszereivel való vizsgálata és megértése illegitim tudományos módszer lenne. Szerinte a róla való lemondás nem csak óriási kihagyott lehetőség, hanem a társadalommal szemben etikátlan is volna, ráadásul a Big Data elemzési keretei eleve alkalmazzák azokat a protokollokat, amelyek minimalizálják a károkozás valószínűségét (Watts 2014).

Kate Crawford az ehhez hasonló állításokról viszont úgy vélekedett<sup>11</sup>, hogy „a képzelőerő és a módszertan kudarcra azt állítani, hogy csak úgy lehet jó adatalapú tudományt művelni, ha ehhez emberek millióin kell kísérletezni a beleegyezésük nélkül” (Crawford 2014). Boyd és Crawford meggyőzően érvelnek amellett, hogy ezek az ígéretetek és állítások a Big Data magánmitológiájának és sales-stratégiájának a részei, így egy részükkel szemben már csak történeti tapasztalataink miatt is érdemes szkeptikusnak lennünk. Így a számok és adatok abszolút objektivitása és pontossága abból a téves felvetésből indul ki, hogy a nyers (*raw*) adat semleges. Régi tévedés ez, már a 17. században Bacon, Boyle és az indukcionisták a részecskékben (értsd itt: nyers adatokban) maguk is az elmélet-mentes, érdekektől meg nem fertőzött tiszta tudást látták (Poovey 1998: xviii), vagyis csak egy régi ígéret bújít új köntösbe. Ahogy Geoffrey Bowker rámutatott, a nyers adat egy oximoron, mert minden adat eleve kicsit „megfőzött” (2005). Bowker Claude Lévi-Strauss *A nyers és a főtt* 1964-es kötetére<sup>12</sup> utaló megjegyzése arra vonatkozik, hogy az adat nem „van”, hanem egy sor ideológiai és módszertani megfontolás, érdeklődés, technológiai feldolgozás során keletkezik. Emellett még a legtisztábbnak gondolt adatbázist is sokszor kell tisztítani, egységesíteni, következetessé tenni, tehát megannyi emberi döntés meghozatalával az elemzésre felkészíteni. Noha a Big Data azt mondja, hogy az adatok a póre *tényeket* mutatják be a maguk neutralitásában, nehéz elképzelni, hogy ezekhez az eredményekhez ne akarnánk jelentéseket, vagy kauzalitásokat társítani, főként úgy, hogy a pokolba vezető út gyakran a korreláció és a kauzalitás összetévesztésével van kikövezeve. Bernard Jansen és munkatársai egészen odáig mennek, hogy azt állítják, hogy a számérték (*count*) és a mérés (*measure*) közötti konceptuális zavar miatt a Big Databan az embereket leíró számok többnyire hamisak, ráadásul a Big Data elemzése során a kutatók és az algoritmusok a régi, jól ismert logikai-statisztikai tévedéseket és érvelési hibákat követik el. Ilyen tévedések többek között a szelekciós/mintavételi torzítások, a dimenzionalitás (túl nagy elemszámú adatbázis) átka, a crowdtruth probléma vagy Simpson paradoxona (Jansen és munkatársai 2022).<sup>13</sup>

Daniel Rosenberg az adat (*data*), a tény (*fact*) és a bizonyíték (*evidence*) angol szavak etimológiai vizsgálatakor arra hívja fel a figyelmet, hogy a latin *datum* múlt idejű szó töve, a *dare*, vagy az adni, így az adat jelentését tekintve az, ami eleve adott, magától értetődő érv. Ezzel szemben a *fact* a latin *facere*, vagyis csinálni igéből származik, amely valami megtörténtre, létezőre vonatkozik. Végezetül az *evidence* a latin *videre*, vagyis látni szóból származtatható, így a bizonyíték az, amelyet a látásunkkal is megtapasztalunk, aminek szemtanúi vagyunk. Az adat így Rosenberg értelmezésében ellentétes mind a tény, mind pedig a bizonyíték jelentésével, amennyiben a tény ontológiai fogalom, a bizonyíték ismeretelméleti, az adat pedig retorikai. Rosenberg úgy érvel, hogy „ha egy tény hamisnak

---

11 A Watts, Crawford és mások részvételével zajló vita azután robbant ki, hogy kiszivárgott, hogy a Facebook megbízásából a Cornell Egyetem hatszázezer felhasználó adatait felhasználva azt vizsgálta, hogy miképpen lehet befolyásolni az érzelmeiket és a döntéseiket, és ezekre a manipulációs technikákra hogyan reagálnak.

12 A *Le Cru et le cuit* Lévi-Strauss négy kötetből álló *Mithologiques*-sorozatának első kötete. Az utolsó 1971-ben jelent meg.

13 Ezen problémák egy részét érinti Kmetty Zoltán is, aki arra is felhívja a figyelmet, hogy a Big Data-paradigma a szociológiai kutatások részterületeit nem egyenlő mértékben fogja érinteni (Kmetty 2018).

bizonyul, megszűnik ténynek lenni. Az adat, ha hamis is, ettől még adat marad” – ezzel rávilágítva az adat és igazság között feszülő ellentmondásra (Rosenberg 2013: 18).

Boyd és Crawford észrevételeinek egy része magára az elemzési lehetőségekre és ígéretre vonatkozik. Szintén számos érvet hoznak fel annak bizonyítására, hogy a Big Data, és itt most szűken a jó nagy méretű adatot értjük alatta, egyáltalán nem biztosít pontosabb, jobb minőségű elemzéseket, mint a kisebb mennyiségű elemet tartalmazó minták. Ezen túl az olyan, a társadalomtudományos vagy marketingkommunikációs tudást elméletben gyarapítandó módszerek, mint a közösségimédia-oldalokon zajló beszélgetések, interakciók lescrapelése (kis programok segítségével történő letöltése) és analízise egy sor, utólagos beavatkozás (botok, inaktív tagok azonosítása és kiszűrése, anonimizálás stb.) és egyéb gyakorlatias, ám emberi megfontolás után válnak elemezhetővé és nyernek értelmet.

Végezetül a szerzők mégiscsak azokon a vizsgálópontokon a legerősebbek, amelyeken a Big Data felhasználásának etikai dilemmáit tárgyalják. A Big Data-kultúra ugyanis akkor tud üdvtörténetté válni, ha hozzáférése van az adatokhoz, ám ez rögtön több kérdést is felvet. Önmagában a hozzáférés még nem jelenti azt, hogy az adat felhasználása etikus, jogszerű, nem sért személyhez fűződő jogokat. Másfelől azonban az adatokhoz való hozzáférés aszimmetrikus. Például a közösségimédia-oldalak személyhez fűződő adataihoz az ott dolgozók korlátlanul, más kutatók csak korlátozottan és pénzért férnek hozzá, amelyet egyre kevésbé tudnak megfizetni, szemben a kormányokkal és a hirdetőkkal. Mondhatnánk, hogy a kutatók még mindig jobban jártak, hiszen a felhasználók, civilek sok esetben még ennyire sem férnek hozzá a Big Datához, így – ahogy a szerzők is érvelnek – a digitális szakadék a „Big Data gazdagok” és „Big Data szegények” között tovább nő. Mindez pedig úgy megy végbe, hogy az adatalapú technológiai cégek semmilyen elszámolási és bejelentési kötelezettséggel nem tartoznak arról, hogy ezeket az adatokat mire használják, kiknek és milyen célból értékesítik, valamint az összes szabályozásukra vonatkozó kísérletet megakadályozzák.

### **Kísérlet(ek) egy diszciplína megteremtésére: hatalom és intézményrendszer**

Boyd és Crawford vitaindító cikke utáni szövegburjánzás egy része váltakozó amplitúdójú aggodalmak összefoglalása. Viktor Mayer-Schönberger és Kenneth Cukier (2018[2013]: 166–203) a kockázatok közül a privát szféra megszűnését, valamint az élet minden apró rezdülését is számszerűsítő „adatdiktatúrát” emelték ki fő kockázatként, és a széleskörű társadalmi kontroll szükségessége mellett az „adatbárók” szabályozását tartották szükségesnek. A szintén geográfus Rob Kitchin ezeket már nem kockáztatnak, hanem komoly aggodalomra okot adó megtörtént események sorozataként írta le (2014b). A már említett Craig Dalton és Jim Thatcher térképész-geográfusok által írt esszé, a 2014-es *What Does A Critical Data Studies Look Like, And Why Do We Care?* javarészt Boyd és Crawford cikkének főbb megállapításait ismétli meg, és a hely, a tér, az idő és az ember feltérképezésével foglalkozó térinformatika tapasztalatait beépítve gördíti tovább a problémát. Elsőre talán meglepőnek tűnik, hogy éppen két térképész az, akik nevet adnak a Big Data-eufória ellen gyűlő szisztematikus kritikai vizsgálatoknak, ám a lokációhoz köthető természeti

javak és emberi tevékenységek rögzítése, követése és láthatóvá tétele önmagában rengeteg problémát vet fel. Ugyan Dalton és Thatcher cikke elsősorban egy geográfusoknak szóló felhívás a történelmileg mindig új maskarába bújó Big Data normalizálódó hegemóniája elleni kritikai részvételre, az általuk felvetett kérdések minden tudományágban érvényesek lehetnek. Ezek a kérdések a következők: Milyen történelmi feltételek vezettek oda, hogy a Big Datáról egyáltalán kritikai kérdések merülhessenek fel? Ki kontrollálja és irányítja a Big Data előállítását és elemzését, és milyen motivációk, kényszerek vezérlik a munkájukat? Kik a Big Data alanyai, tárgyai, és ők miféle tudásokat állítanak elő? Hogyan alkalmazzák a Big Data-t a helyek, terek, tájképek előállítása során? Milyen egyéb tudások előállításában segíthet bennünket a Big Data?

Miközben ezekre a kérdésekre a mai napig is csak részleges válaszaink vannak, Dalton és Thatcher az esszéjük címében szereplő kérdésre, miszerint hogy is nézne ki a kritikai adatkutatás, Rob Kitchin a nyilván Deleuze és Guattari assemblage-elméletére rímelő adat-asszemblázs (*data assemblage*) fogalmával válaszolt, amely ernyő alá gyűjti azt az apparátust, amely kijelöli a kritikai adatkutatás vizsgálódási kereteit. Kitchin úgy fogalmaz, hogy az adat-asszemblázs egy komplex szocio-technológiai rendszer, amely számos, egymással összekapcsolódó eszközből és elemből áll, melyek elsődleges feladata az adat előállítása. Ez az asszemblázs nem csak az adatinfrastruktúrából áll, mint például egy nagy rendszer, adattár vagy archívum, hanem magában foglalja az összes olyan technológiai, politikai, társadalmi és gazdasági eszköztárat, amely az asszemblázsok természetét és működését meghatározzák. Ezek az apparátusok és elemek sokrétű kapcsolatok esetleges és összetett hálózatán keresztül kölcsönhatásba lépnek egymással és alakítják is egymást (Kitchin 2014b: 24). Magyarán: a kritikai adatkutatásnak nem csupán magát az adatot, mint tárgyat kell önmagában vizsgálnia, hanem a teljes, az adat létrejöttében szerepet játszó aktorhálózatot és intézményrendszert is.

2016-ban Andrew Iliadis és Federica Russo (2016) a kritikai adatkutatás mint diszciplína létrejöttét már történetiségében mutatták be. Érdemes a cikküket a megmozgott anyagmennyiség felől is nézni, és láthatóvá válik, hogy Dalton és Thatcher óhaja nyitott fülekre talált, „ott volt mindenki száján, csak ki kellett mondani” (Dalton, Taylor és Thatcher 2016: 1). Ez két dolgot mindenképpen jelenthet. Az egyik, hogy röpke két év alatt nagyon sok tudományterület képviselőit sikerült megszólítaniuk és nagyjából ugyanazok a kérdések aggasztották őket is. A másik pedig az, hogy a már létező adatkutatás/adattudomány elé biggyesztett kritikai jelző, ahogy a tudománytörténetben nem először, sokakat invitált arra, hogy valami felkapott dologhoz hozzászóljanak. Mindenesetre Iliadis és Russo cikke valóságos kincseshánya lehet azoknak, akik el akarnak igazodni a kritikai adatkutatás termékeny birodalmában.

A szerzők már a szöveg elején leszögezik: az adatokat a hatalom egyik megnyilvánulási formájának tekintik, amely az adatechnológiára épülő vállalatoknak óriási pénzt termel, amely képes az ember érzésein túl a kultúrát is manipulálni, és amely egyben az objektivitás látszatának fegyverét adja a kutatók kezébe. Számos bizonyítékot sorolnak arra, hogy az adatot sokszor az érintettek tudomása nélkül használják fel, hogy az adatot gyakran eltérítik az iparági proficélokot szolgáló kutatói agendák, és hogy nincs sem megfelelő önszabályozás, iparági védelem, sem protokoll a hackerek és a titkosszolgálati megfigyelések ellen. „Miközben az adat mindaz, amit fentebb írtunk, az adat hiánya ugyanígy

szembetűnő – az adat hiánya ugyanúgy a hatalom jele, azé a hatalomé, amelyik nem akarja, hogy lássuk, amelyek rejtőzve akar maradni” – írják. A megannyi ismert probléma ellenére azonban úgy látják, hogy a Big Data-forradalmon nincs fék, hiszen néhány év alatt „a társadalom- és természettudományok (és tegyük hozzá gyorsan, a bölcsészettudományok – BA) területén az adatintenzív, gyors eredményt ígérő pozitivisták módszerei szinte teljesen kiszorították a hosszabb távú, posztpozitivisták, elméletibb megközelítéseket” (Iliadis és Russo 2016: 1). Ebben a társadalomelmélettel foglalkozók bizonyára maguk is szerezhettek keserű tapasztalatokat az elmúlt egy évtizedben a külföldi folyóiratokhoz benyújtott elméleti cikkeik visszautasításakor.

## A kritikai adatkutatás alkalmazhatósága és egyes fogalmai

A kritikai adatkutatással is kicsit az a helyzet, mint a kritikai kultúrakutatással (*cultural studies*). Noha megfigyelései, interdiszciplináris módszerei, sőt, eszményei nagyon sok tudományterületbe beépültek, viszonylag kevés kutató kevés munkája viseli a kritikai adatkutatás kítűzőjét magán. Ez már Iliadis és Russo cikke megjelenésekor is világos volt, a szélrózsa minden irányából igyekeztek összegyűjteni azokat a régebbi és újabb tudományos és népszerű műveket egyaránt, amelyekre rá lehetett aggatni, hogy a kritikai adatkutatás szellemében íródtak.

Ezeknek, és az azóta született munkáknak egy része továbbra is a kritikai adatkutatás elméleti kereteit és vizsgálati módszereit polírozza, más része elvégzi a kritikai piszkosmunkát, megint mások új irányok kijelölésén fáradoznak (Hepp, Jarke és Kramp 2022; Jarke és Bates 2024). Az alkalmazhatóság szempontjából a második kategória az izgalmasabb, mivel ezen kutatók többsége túlmegy az elméleti vívódásokon és konkrét kutatást végeznek, nem egy esetben aktivizmust is megfogalmaznak. Ilyenek a *data justice* témakörébe tartozó tanulmányok, amelyek a Big Data és az algoritmusok társadalmi egyenlőtlenségek újratermelésében, az elnyomó struktúrák fenntartásában, vagy akár a rasszizmus és a genocídiumokban játszott szerepét mutatják be és elemzik (O’Neill 2016). Részből a *data justice* és a reprezentáció kérdéskörét érintik a Big Data feminista kritikái (D’Ignazio és Klein 2019; D’Ignazio 2024), valamint a digitális szakadékok csökkentését célzó adatműveltségi (*data literacy*) kutatások (D’Ignazio 2017). A *data justice* reprezentációpolitikai leágazásait jelentik a *counterdata* ellenkulturális praxisába illeszthető aktivizmusok. A *counterdata* azokat a közösségi, „nem hivatalos”, részvételi adatgyűjtési gyakorlatokat foglalja magába, amelyek az állami-hivatali vagy éppen a vállalati adatgyűjtésekkel, valamint az általuk interpretálható valósággal szemben születnek meg, és az adatgyűjtés alapjául szolgáló kérdések felvetésének jogát az állampolgárok közösségének szolgáltatja vissza (Olojo 2024). Szintén a reprezentációra, illetve annak hiányára reflektál a *missing data* fogalma, amely azokat az adatköröket foglalja magába, „amelyek nincsenek, de létezniük kellene” (Onuoha 2016). Az adatok hiányossága/elitközlése mögött sokszor történeti-politikai, vagy üzletstratégiai okok húzódnak meg, hiányuk megállapítása, valamint a pótlásuk sürgetése szintén a kritikai adatkutatás egyik fontos területe (Jungs de Almeida, Klein és D’Ignazio 2024). A *data colonialism* (Couldry és Mejias 2019, 2024) azt a komplex viszonyrendszert hivatott átfogni, amely az emberi élet

minden aspektusát az adatosítás révén a kapitalizmus szolgálatába állítja. A privát szféra abúza, a megfigyelés és a kontroll, valamint az autokráciák, adatcégek és a hirdetőik ösz-szefonódásának kérdésével foglalkoznak a data és a surveillance összevonásából képzett *dataveillance* tanulmányok és monográfiák (Dijck 2014; Zuboff 2014, 2019). A kritikai adatkutatás megállapításai továbbá jól ráilleszthetők az adatok vizuális reprezentációinak kritikai megközelítéseire is (Engelbrechtsen és Kennedy szerk. 2020)<sup>14</sup>, valamint az adatvezérelt, *smart city* várostervezés- és irányítás problémáira (Mattern 2021).

### A kritikai adatkutatás kritikái és ellentmondásai

A kritikai adatkutatás sokat idézett klasszikus szövegeinek jellemző műfaja a vitairat. Ez a tény egyfelől elhomályosítja a kritikai adatkutatáshoz kapcsolódó valódi kutatások érdemeit, másfelől azonban muníciót adhat azok számára, akik szerint a kritikai adatkutatás az állításait nem tudja adatokkal alátámasztani. Ezzel szemben az adattudomány (*data science*) és a Big Data-alapú társadalom- és bölcsészettudományi kutatások azzal, hogy időközben elsajátítani látszanak a természettudományos kutatások módszertani apparátusát, eltávolodnak a *soft science*-ek sokszor vádként megfogalmazott módszertani renyhességétől, ideologikusságától, interpretatív jellegétől. Éppen ezért a Big Data-ra épülő kutatások sokszor úgy tehetnek, mintha immunitást élveznének a kritikai adatkutatás megállapításaival szemben.

Ugyan a kritikai adatkutatás örvén született írások egy kis része valóban ijesztő disztópiákat fest le a jelenről és a jövőről, azért a többségük mégiscsak arról szól, hogy a Big Data óriási lehetőség, de egyben felelősséggel is jár. Ugyanakkor ez a feltartott mutatóujj is sokakat irritálhat. Ahogy Craig Dalton, Linnet Taylor és Jim Thatcher (2016) is megjegyzi, a „kritikai” előtag odabiggyesztésével „fennáll annak a veszélye, hogy számos, a kutatásaikat eleve kritikusan, etikusan végző kutatót megsért, és feleslegesen választja szét a kritikai elméletet a kutatásaikba beillesztő tudósokat azoktól, akik szigorú empirikus kutatásokban vesznek részt.” Így a kritikai elmélet politikussága és aktivizmusa révén a tudomány művelői számára terhes lehet még akkor is, ha maguk is tisztában vannak az adat politikai és hatalmi kérdéseivel, és eleve kötik őket a tudományetikai normák. Éppen ezért a kritikai adatkutatásnak nyitottnak kell maradnia az ellenkritikára, különben maga is dogmává merevedik.

Természetesen voltak korábban is, és ma is vannak, akik azt mondják, hogy a kritikai adatkutatás által felvetett problémák valójában álproblémák. Részben azért, mert ha reálpolitikailag nézzük, akkor a Big Data minden aggodalmaskodás és botrány ellenére végigsöpört az államszervezéstől kezdve a tudományokon át egészen az emberek mindennapi életéig, és túl lehetett élni. Másfelől az új forradalom, a Big Data-ra épülő mesterséges intelligenciáé ugyanezeket a kérdéseket hozza fel ismét, így most már erre kellene koncentrálni. Megint mások azt mondják, hogy a kritikai adatkutatás valójában különféle pesszimizmusok (technopesszimizmus, számpesszimizmus) 21. századi reinkarnációja, és ugyanúgy, ahogy a korábbi disztópiák sem váltak valóra, úgy a Big Data-apokalipszisre

<sup>14</sup> Lásd továbbá a kötetben Gregor Anikó és Katona Eszter (2024) cikkét.

is kicsi az esély. Számos ellenvetés arra hívja fel a figyelmet, hogy a kritikai adatkutatás meglátásai méltánytalanul egyoldalúak, miközben a Big Data által előidézett összestrádalmi előnyökkel összehasonlítva a kockázatok eltörpülnek.

A kritikai adatkutatás problémafelvetéseinek legnagyobb ellenzői azonban maguk a technológiai cégek, és konzervatív-libertáriánus kommentátoraik. A jól ismert érvek szerint minden szabályozásra, átláthatóságra, elszámoltathatóságra és demokratizálódásra vonatkozó kísérlet súlyos beavatkozás a vállalkozások szabadságába és a szabad piacba. Az olyan adatvállalkozók és befektetők, mint Elon Musk, Peter Thiel vagy David Sacks az- zal vádolják azokat, akik például a vállalataik által gyűjtött felhasználói adatok nyílt kutató- tási célokra való felszabadításáért kampányolnak, hogy azokat csak a szélsőbaloldali *social justice* elméleteik igazolására használnák fel.

A nyílt adat (*open data*), és ezeken belül is a nyílt kormányzati és nyílt tudományos adatok helyzete ellentmondásos. Egyfelől ezeknek az adatoknak a mennyisége nő, minősége folyamatosan javul (*Global Data Barometer 2022; Open Data Maturity 2023*), és ez javarészt a közérdekű és tudományos adatok átláthatóságáért és demokratizálódásáért elindított civil mozgalmaknak és transznacionális szabályozásoknak köszönhető (mint például az Európai Unió tudományfinanszírozásában a Responsible Research and Innovation (RRI) keretrendszere). Másfelől azonban jelentős regionális egyenlőtlenségek tapasztalhatók, amely a szegényebb országok és az autoriter rezsimek állampolgárait jelentős hátrányba hozza. Érdeemes továbbá figyelemmel kísérni azt az ellentétes folyamatot is, amely akár a közpénzből finanszírozott kormányzati és tudományos adatokat, akár a felhasználók által gyűjtött, generált adatokat kereskedelmi áruvá konvertálja, így ezekből az adatokból végeredményben megint csak azok fognak profitálni, akik meg tudják őket venni (Lindman és Kuk 2015). Ez a folyamat sem minden régiót és országot érint egyformán. Mások viszont arra hívják fel a figyelmet, hogy a nyílt adat, nyitott tudomány, vagy a kormányzati elszámoltathatóság jelszavait a profitérdekelt vállalatok trójai falóként használják fel arra, hogy az így elérhetővé vált kutatási adatok módszertanát, eredményeit megkérdőjelezzék, manipulálják, a tudományos konszenzussal szembeni ellennarratívát állítsanak fel, miközben rájuk nem vonatkoznak a nyílt hozzáférésre vonatkozó törvények (Levy és Johns 2016).

### Replika 93/2015

Érdekes majdnem közel tíz év távlatából újra elolvasni a Dessewffy Tibor és Láng László által szerkesztett Big Data és szociológia *Replika*-blokk cikkeit és reflexióit (2015). A kritikai adatkutatás korai szempontjai hol lábjegyzetekben kerülnek elő, hol érvekként (Németh 2015, Székely 2015), hol pedig morális pánikként (Z. Karvalics 2015). Utóbbi azért is érdekes, mert a Z. Karvalics László az általa kipellengérezett „what if” világot (2015: 191) akkor a „kritikaival” együtt a fikciós és ezoterikus irodalom polcára tette.

Utóbbiak ideológiának vagy vallásnak tekintik a „szép új adatvilág” profétáinak lelkesedését, akikkel szemben örömmel tetszelegnek az orwelli forgatókönyvek feltáráinak, egyfajta kollektív lelkiismereti kapuőröknek a kisebbségi szerepében. Ha nem kezdené már az akadémiai közbeszédet is, önmagát kritikainak tartva, meghatározni ez

a szemléleti kiindulópont és gondolkodói attitűd, s ha nem indulna ki egyre több publicista abból a feltételezésből, hogy a médiafigyelmet az apokaliptikus adatjövőrre irányítva lehet felkelteni, akkor még gondolhatnánk azt, hogy időben vagyunk. De sajnos elkéstünk (Z. Karvalics 2015: 191).

A kritikai adatkutatást valaha gyarapítók, illetve a társadalomtudományokban a kritikai elmélet szempontjait is érvényesítők szinte majdnem mindegyike maga is Big Data-felhasználó kutató. Másfelől azonban a kritikai adatkutatás által felvetett problémák nem lehetőségként, esetleges kockázatként és elhanyagolható mellékhatásként merültek fel, hanem akkor is zajló, és azóta csak erősödő realitásként. Az elmúlt évtizedben a Big Data kormányzati és vállalati felhasználása körüli jól ismert botrányok és visszaélések, melyek között megtalálható a választói akarat illegitim befolyásolása, a totális megfigyelőrendszerek kiépítése, a tartalommegosztás algoritmikus eltérítése, a kutatási adatok manipulációja mind arra hívják fel a figyelmet, hogy a kritikai adatkutatás által felvetett problémák érvényesek voltak már akkor is. Ugyan némely platformok, adatvállalatok és kutatók elkezdtek beépíteni a gyakorlataikba a kritikai adatkutatás egyes szempontjait, a szemünk láttára történő MI-forradalom továbbra is sok munkát fog adni a kritikai elméletnek.

### Hivatkozott irodalom

- Anderson, Chris (2008): The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. *Wired*, Interneten: <https://www.wired.com/2008/06/pb-theory/> (letöltve: 2024.11.02.).
- Barry, David M. (2011): The Computational Turn: Thinking About the Digital Humanities. *Culture Machine* (12): 1–22.
- Beniger, James R. (2004[1985]): *Az irányítás forradalma. Az információs társadalom technológiai és gazdasági forrásai*. Budapest: Atlantisz.
- Bowker, Geoffrey C. (2005): *Memory Practices in the Sciences*. Cambridge, MA: MIT Press.
- boyd, danah és Kate Crawford (2012a): Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon. *Information, Communication, & Society* 15(5): 662–679. DOI: <https://doi.org/10.1080/1369118X.2012.678878>
- boyd, Danah és Kate Crawford (2012b): Az adatrengeteg kínos kérdései: Vitaindító egy kulturális, műszaki és tudományos jelenségről. *Információs Társadalom* 12(2): 7–23. DOI: <https://dx.doi.org/10.22503/infarts.XII.2012.2.1>
- Brooks, David (2013): The Philosophy of Data. *The New York Times*. Interneten: <https://www.nytimes.com/2013/02/05/opinion/brooks-the-philosophy-of-data.html> (letöltve: 2024.11.02.).
- Buck-Morss, Susan (1995): Envisioning Capital: Political Economy on Display. *Critical Inquiry* 21(2) (Winter, 1995): 434–467. DOI: <https://doi.org/10.1086/448759>
- Couldry, Nick és Ulises A. Mejías (2019): *The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism*. Stanford, CA: Stanford University Press.
- Couldry, Nick és Ulises A. Mejías (2024): *Data Grab. The New Colonialism of Big Tech and How to Fight Back*. Chicago: Chicago University Press.
- Crawford, Kate (2014): The Test We Can – and Should – Run on Facebook. *The Atlantic*. Interneten: <https://www.theatlantic.com/technology/archive/2014/07/the-test-we-canand-shouldrun-on-facebook/373819/> (letöltve: 2024.11.02.).
- Cryle, Peter és Elisabeth Stephens (2017): *Normality. A Critical Genealogy*. Chicago: Chicago University Press
- Dalton, Craig és Jim Thatcher (2014): What Does A Critical Data Studies Look Like, And Why Do We Care? *Society & Space*, 2014. május 12. Interneten: <https://www.societyandspace.org/articles/what-does-a-critical-data-studies-look-like-and-why-do-we-care> (letöltve: 2024.11.02.).
- Dalton, Craig, Linnet Taylor és Jim Thatcher (2016). Critical Data Studies: a dialog on time and space. *Big Data & Society* 3(1) 2016 június. DOI: <https://doi.org/10.1177/2053951716648346>
- Desrosières, Alain (1998): *The Politics of Large Numbers. A History of Statistical Reasoning*. Cambridge, MA: Harvard University Press.

- Dessewffy Tibor és Láng László szerk. (2015): Big data és szociológia. *Replika* (92–93): 155–230.
- D'Ignazio, Catherine (2024): *Counting Feminicide. Data Feminism in Action*. Cambridge, MA: MIT Press.
- D'Ignazio, Catherine (2017): Creative data literacy. Bridging the gap between the data-haves and data-have nots. *Information Design Journal* 23(1): 6–18. DOI: <https://doi.org/10.1075/idj.23.1.03dig>
- D'Ignazio, Catherine és Lauren F. Klein (2019): *Data Feminism*. Cambridge, MA: MIT Press.
- Dijck, José van (2014). Datafication, Dataism and Dataveillance: Big Data Between Scientific Paradigm and Ideology. *Surveillance & Society* 12(2): 197–208. DOI: <https://doi.org/10.24908/ss.v12i2.4776>
- Dusek Tamás (2024): A mérés és méréselmélet néhány alapkérdése. *Statisztikai Szemle* 102(2): 107–157. DOI: <https://doi.org/10.20311/stat2024.02.hu0107>
- Engelbrechtsen, Martin és Helen Kennedy szerk. (2020): *Data Visualization in Society*. Amsterdam: Amsterdam University Press. DOI: <https://doi.org/10.2307/j.ctvzgb8c7>
- Foucault, Michel (1990): *Felügyelet és büntetés. A börtön története*. Budapest: Gondolat.
- Global Data Barometer (2022): *First Edition Report – Global Data Barometer*. ILDA. DOI: <https://doi.org/10.5281/zenodo.6488349>
- Gould, Stephen Jay (1981[1996]): *The Mismeasure of Man*. New York: WW Norton
- Gregor Anikó és Katona Eszter (2024): Egy ábra és ami mögötte van: az adatvizualizáció politikája. *Replika* (134): 35–53. DOI: <https://doi.org/10.32564/134.3>
- Hepp, Andreas, Juliane Jarke és Leif Kramp szerk. (2022): *New Perspectives in Critical Data Studies. The Ambivalences of Data Power*. Cham: Palgrave Macmillan. DOI: <https://doi.org/10.1007/978-3-030-96180-0>
- Horkheimer, Max és Theodor W. Adorno (2020): *A felvilágosodás dialektikája. Filozófiai töredékek*. Budapest: Atlantisz.
- Iliadis, Andrew és Federica Russo (2016): Critical Data Studies: An Introduction. *Big Data & Society* 3(2). DOI: <https://doi.org/10.1177/2053951716674238>
- Jansen, Bernard J. et al. (2022): The Illusion of Data Validity: Why Numbers about People Are Likely Wrong. *Data and Information Management* 6(4) October 2022: 1–14. DOI: <https://doi.org/10.1016/j.dim.2022.100020>
- Jarke, Juliane és Jo Bates szerk. (2024): *Dialogues in Data Power. Shifting Response-abilities in a Datafied World*. Bristol: Bristol University Press.
- Jungs de Almeida, Alessandra, Lauren Klein és Catherine D'Ignazio (2024): Missing Data. In *Keywords of the Datafied State*. Jenna Burrell, Ranjit Singh és Patrick Davison (szerk.). Data & Society, 194–205. DOI: <http://dx.doi.org/10.2139/ssrn.4734250>
- Kaplan, Frédéric és Isabelle de Lenardo (2017): Big Data of the Past. *Frontiers in Digital Humanities* 4(12). DOI: <https://doi.org/10.3389/fdigh.2017.00012>
- Kehlmann, Daniel (2006): *A világ fölmérése*. Budapest: Magvető.
- Kitchin, Rob (2014a): Big Data, New Epistemologies and Paradigm Shifts. *Big Data & Society* 1(1): 1–12. DOI: <https://doi.org/10.1177/20539517146528481>
- Kitchin, Rob (2014b): *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*. London: Sage DOI: <https://doi.org/10.4135/9781473909472>
- Kitchin, Rob és Tracy Lauriault (2014): Towards Critical Data Studies: Charting and Unpacking Data Assemblages and Their Work. *The Programmable City Working Paper* 2.
- Kmetty Zoltán (2018): A szociológia helye a Big Data-paradigmában és a Big Data helye a szociológiában. *Magyar Tudomány* (179): 683–692. DOI: <https://doi.org/10.1556/2065.179.2018.5.11>
- Lévi-Strauss, Claude (1964[1983]): *The Raw and the Cooked. Mythologiques Vol. 1*. Chicago: Chicago University Press.
- Levy, Karen és David Merritt Johns (2016): When Open Data is a Trojan Horse: The Weaponization of Transparency in Science and Governance. *Big Data & Society* 3(1): 1–6. DOI: <https://doi.org/10.1177/2053951715621568>
- Lindmann, Juho és George Kuk (2015): From Open Access to Open Data Markets: Increasing the Subtractability of Open Data. In *2015 48th Hawaii International Conference on System Sciences (HICSS)*. IEEE Computer Society 2015, 1306–1313. DOI: <https://doi.org/10.1109/HICSS.2015.159>
- Lohr, Steve (2015): *Dataism. Inside the Big Data Revolution*. London: Oneworld.
- Mattern, Shannon (2021): *A City is Not a Computer: Other Urban Intelligences*. Princeton: Princeton Architectural Press.
- Mayer-Schönberger, Viktor és Kenneth Cukier (2018): Big Data. Forradalmi módszer, amely megváltoztatja a munkánkat, gondolkodásunkat és az egész életünket. Budapest: HVG Könyvek.
- Muller, Z. Jerry (2018): *The Tyranny of Metrics*. Princeton: Princeton University Press.
- Németh Renáta (2015): A számok tényleg magukért beszélnek? Hozzászólás Deseffly Tibor és Láng László írásához. *Replika* (92–93): 203–208.
- Olojo, Seyi (2024): Counterdata. In *Keywords of the Datafied State*. Jenna Burrell, Ranjit Singh és Patrick Davison

- (szerk.), *Data & Society*, 171–181.
- O’Neill, Cathy (2016): *Weapons of Math Destruction. How Big Data Increases Inequality and Threatens Democracy*. New York: Crown Books.
- Onuoha, Mimi (2016): *The Library of Missing Datasets V 1.0*. Interneten: <https://mimionuoha.com/the-library-of-missing-datasets> (letöltve: 2024.11.02.)
- Open Data Maturity (2023): *Report 2023*. Luxembourg: Publications Office of the European Union. DOI: 10.2830/384422
- Poovey, Mary (1998): *A History of the Modern Fact. Problems of Knowledge in the Sciences of Wealth and Society*. Chicago: Chicago University Press.
- Porter, Theodore M. (1986): *The Rise of Statistical Thinking, 1820–1900*. Princeton: Princeton University Press.
- Porter, Theodore M. (1995): *Trust in Numbers. The Pursuit of Objectivity in Science and Public Life*. Princeton: Princeton University Press.
- Prensky, Marc (2009): H. Sapiens Digital: From Digital Immigrants and Digital Natives to Digital Wisdom. *Innovate: Journal of Online Education* 5(3). Interneten: <https://nsuworks.nova.edu/cgi/viewcontent.cgi?article=1020&context=innovate> (letöltve: 2024.11.02.)
- Puschmann, Cornelius és Jean Burgess (2014): Metaphors of Big Data. *International Journal of Communication* (8): 1690–1709.
- Rosenberg, Daniel (2013): Data Before the Fact. In „*Raw Data*” is an Oxymoron. Lisa Gitelman (szerk.). Cambridge, MA: MIT Press, 15–40.
- Scott, James C. (1998): *Seeing Like a State. How Certain Schemes to Improve the Human Condition Have Failed*. New Haven, London: Yale University Press.
- Siddique, Ashees Kapur (2024): *The Archive of Empire: Knowledge, Conquest, and the Making of the Early Modern British World*. New Haven: Yale University Press.
- Székely Iván (2015): Az adatmentes zónák szükségessége és helyessége. Helytelen reakció Dessewffy Tibor és Láng László írására. *Replika* (92–93): 209–225.
- Vincent, James (2022): *Beyond Measure. The Hidden History of Measurement*. London: Faber & Faber.
- Vörös Miklós és Nagy Zsolt (1995): Kultúra és politika a mindennapi életben. Bevezetés a kritikai kultúrákutásba. *Replika* (17–18): 153–157.
- Watts, Duncan J. (2014): Stop Complaining About the Facebook Study. It’s a Golden Age for Research. *Guardian*. Interneten: <https://www.theguardian.com/commentisfree/2014/jul/07/facebook-study-science-experiment-research> (letöltve: 2024.11.02.)
- Ziman, John (1996): Postacademic Science. Constructing Knowledge with Networks and Norms. *Science & Technology Studies* 9(1): 67–80.
- Z. Karvalics László (2015): A Nagy-Adat-jelenség társadalomtudományi lehorgonyzásához. *Replika* (92–93): 189–201.
- Zuboff, Shoshana (2014): A Digital Declaration. Big Data as Surveillance Capitalism. *Frankfurter Allgemeine Zeitung* (FAZ.net). Interneten: <https://www.faz.net/aktuell/feuilleton/debatten/the-digital-debate/shoshana-zuboff-on-big-data-as-surveillance-capitalism-13152525.html> (letöltve: 2024.11.02.)
- Zuboff, Shoshana (2019): *The Age of Surveillance Capitalism. The Fight for Human Future at the New Frontier of Power*. London: Profile.

Bátorfy Attila:

---

médiakutató, információdízájner, az ELTE BTK Művészetelméleti és Médiakutatási Intézetének oktatója.