



A hivatalos statisztika problémái az adatrobbanás korában

Németh Zsolt

A hivatalos statisztika válsága az adatforradalomban

Absztrakt: Az adatforradalom minden korábban megtapasztalt változásnál gyorsabb és mélyebb átalakulást okoz a társadalomban és a gazdaságban. A hivatalos statisztika jelene és jövője szempontjából kulcsfontosságú kérdés, hogy a fundamentális változások közepette mennyiben tud megfelelni küldetésének és kötelezettségének: hogy hitelesen, magas színvonalú statisztikai adatokkal mutassa be a társadalom, a gazdaság és a környezet állapotát és folyamatait. A hivatalos statisztika számára a hosszú történelmi perióduson át természetes működési közeget biztosító racionális jogállam válságjelei, a nagy technológiai cégek adatgyűjtő és adatelemző tevékenysége, a közösségi média és benne a post-truth valóság, az összeesküvés-elméletek, a fake news, az identitáspolitikák felemelkedése, a politikai és kommunikációs környezet átalakulása korábban ismeretlen kihívásokat teremtett. Az adatforradalom új társadalmi, gazdasági valóságát hordozó, exponenciálisan bővülő adatvagyon szinte kizárólag a technológiai cégek magántulajdonában áll. Márpedig a cikk érvelése szerint ezek társadalmi tények, amikhez azonban a hivatalos statisztikának nincs hozzáférése. Ennek hiányában nem képes klasszifikációit, osztályozásait, fogalomrendszerét megújítani, alkalmassá tenni az új társadalmi valóság megragadására, így a valóságról alkotott képe egyre homályosabb, elmosódottabb, és mindennek az árát hitelvesztéssel fizeti meg.

Kulcsszavak: adatforradalom, hivatalos statisztika, válság, társadalmi tények, felügyelő kapitalizmus

„It is a duty of statisticians to explore the link between statistics, science and society and to lead intellectual reflections on the possible risk of reliance on data-centrism.”
(Radermacher 2018)

„Van, ami rock and roll, és van, ami nem.”
(URH)

Bevezetés

Egyetértés van az elemzők között abban, hogy korunk legfőbb jellemzője a társadalom, a gazdaság és a mindennapi élet jóformán minden részét átható, a történelemben korábban soha nem tapasztalt sebességű változás, átalakulás. Éppen a folyamatok gyorsaságából, abszolút jelenidejűségéből fakad az, hogy egyelőre nincs még konszenzusos elnevezése sem annak, ami a szemünk előtt történik. Másfél, két évtizeddel ezelőtt információs társadalomról beszéltünk, nincs egy évtizede annak, hogy az adatforradalom és a Big Data került az érdeklődés középpontjába, néhány éve a mesterséges intelligencia (AI) és az algoritmusos társadalom fogalmkörében igyekeznek értelmezni a változásokat. A társadalomtudományokban széleskörűen elterjedt a meglehetősen semleges digitalizáció címke, ami a megkérdőjelezhetetlen fontosságú digitális technológiai alapokra, a tárolókapacitások, a feldolgozási sebesség dinamikus növekedésére és a hálózatok szerepére helyezi a hangsúlyt. Az egyik legújabban terjedő fogalom a dataizmus (Harari 2017: 316). A formálódó új renddel szembeni kritikai attitűdöt jelzik a felügyelő kapitalizmus (*surveillance capitalism*; Zuboff 2016, 2019), illetve a kommunikatív kapitalizmus (Dean 2018: 34) kifejezések. Az, hogy ki milyen megnevezés mellett kötelezi el magát, függ részben attól, milyen szaktudomány eszköz- és fogalomrendszerével közelíti meg, illetve attól, hogy a fölöttébb komplex jelenség-halmaz melyik metszetére, területére fókuszál.

Ebben az írásban annak feltárására, megértésére teszünk kísérletet, hogy a folyamatban levő változások milyen hatást gyakorolnak a hivatalos statisztika¹ helyzetére, pozíciójára, azon vállalt küldetésének teljesítésére, hogy tevékenysége során gyűjtött adatok feldolgozásával, elemzések vagy adatállományok formájában történő közzétételével a felhasználók számára hiteles képet nyújt a társadalomról, a gazdaságról, a környezetről és ezek változásairól. Mivel a hivatalos statisztika adatokkal foglalkozik, tevékenysége szempontjából korunk

1 Az ENSZ Kögyűlésének a hivatalos statisztika alapelveiről 2014. jan. 29-én elfogadott határozata értelmében „A hivatalos statisztika a demokratikus társadalom információs rendszerének nélkülözhetetlen eleme, amely adatokat szolgáltat a kormányzat, a gazdaság és a nyilvánosság számára a gazdasági, demográfiai és társadalmi helyzetről, illetve a környezet állapotáról. Ebből a célból a hivatalos statisztikai szervezeteknek a gyakorlatban jól hasznosítható statisztikát kell előállítaniuk, és azt pártatlanul rendelkezésre bocsátaniuk, tiszteletben tartva az állampolgárok közérdeklő információkhoz való hozzáférési jogát”. (https://www.ksh.hu/docs/szolgáltatások/sajtoszoba/fundamental_principles_magyar.pdf) A hivatalos statisztika intézményrendszerének kulcsszereplője a nemzeti statisztikai intézmény, Magyarországon a KSH, de tagja lehet a 2016. évi CLV. törvény 4. § (2) értelmében az a szervezet, „amely közfeladatának részeként hoz nyilvánosságra hivatalos statisztikai adatokat, és amelynek szervezete és működése megfelel a Nemzeti Statisztika Gyakorlati Kódexében és az Európai Statisztikai Rendeletben foglaltaknak”. A tanulmányban hivatalos statisztikát előállítóként kezeljük a nagy nemzetközi szervezetek (ENSZ, OECD, Világbank) statisztikai részleget és mindenekelőtt az Eurostatot.

leírására a legrelevánsabbnak az adatforradalom² kifejezés tűnik, ezért a továbbiakban a statisztikát illetően ezt a kifejezést fogjuk használni.

A hivatalos statisztika beleütközik az adatforradalomba

Az új évezredben több évtizedes fejlődés eredményeképpen az ENSZ és más nagy nemzetközi szervezetek – mint a Világbank, az OECD, az EU – statisztikai részlegeinek égisze alatt létrejöttek a hivatalos statisztikának azok az adatbázisai, amelyek segítenek minden korábinál pontosabb képet alkotni a világunkban zajló társadalmi, gazdasági, környezeti folyamatokról. Globálisan informálnak a születéskor várható élettartam, az egy főre jutó GDP, a végső fogyasztási kiadások, az írástudók arányának – különösen a nők körében bekövetkezett – növekedéséről, a gyermekkorúak védőoltásának eredményeiről, a gyermekhalandóság csökkenéséről, a mélyszegénységben élők számának és arányának javulásáról. A klímaváltozásról és annak várható következményeiről adatok alapján vitatkozhatnak a szakértők, az aktivisták és a politikusok. A kiragadott példák talán jól mutatják, hogy minden korábinál közelebb kerültünk a ténytudatossághoz, a tényekre alapozott döntések kultúrájának terjedéséhez, köszönhetően a hivatalos statisztika eredményeinek.

Fontos előrelépések történtek a hivatalos statisztikai tevékenység sztenderdizálásában is. Az ENSZ Közgyűlése 2014-ben elfogadta a hivatalos statisztika alapelveit rögzítő dokumentumot,³ az Európai Statisztikai Rendszer tagországai pedig 2017-ben bevezették az Európai Statisztika Gyakorlati Kódexét.⁴

Ugyancsak a kétezres években következett be minőségi ugrás az információs technológiában. A tárolókapacitások és a feldolgozási sebesség dinamikus növekedéséhez a mindennapi életben egyre meghatározóbb szerephez jutó, gyakorlatilag folyamatos online jelenlétet biztosító mobil eszközök, a műholdak, kamerák, a legkülönfélébb érzékelők terjedése, az internet beépülése az üzleti és a magánéletbe biztosítja a mára már szinte felfoghatatlan méretű adatfolyamot. Ennek az adatforgalomnak a pillanatról pillanatra táguló adatuniverzuma nem csupán átszövi, hanem alapjaiban forgatja fel a társadalom és a gazdaság hagyományos szerkezetét.

Nézzük meg röviden, miként formálta az adatforradalom a gazdaságot, illetve a technológiai robbanásban vezető szerepet játszó, vagy annak lehetőségeire jó tempóban reagáló cégek miként lendítették tovább az adatforradalmat. A világ legnagyobb, legerősebb vállalkozásairól számos lista készül. Bármelyiket is böngésszük ezek közül, a piaci tőkeérték alapján készült 2018-as és 2019-es adatok szerint az első négy helyen konzisztensen a négy technológiai óriás, az Apple, az Amazon, a Microsoft és az Alphabet (Google)⁵ osztozik, őket

2 Az adatforradalom számos lehetséges definíciója közül az ENSZ tematikus kiadványában megfogalmazottat tekintjük irányadónak. Eszerint „az adatforradalom robbanás az adatok mennyiségében, a sebességben, ahogyan az adatokat előállítják, az adat-előállítók számában, az adatok terjesztésében és azon dolgok körében, amelyekről rendelkezésre állnak adatok az új technológiákból fakadóan, mint a mobiltelefonok, a »dolgok internete« és egyéb források, úgymint kvalitatív adatok, az állampolgárok által generált adatok, az érzékelők adatai; továbbá a növekvő adatigény a társadalom minden részéből” (A World That Counts 2014: 6).

3 https://www.ksh.hu/docs/szolgáltatások/sajtoszoba/fundamental_principles_magyar.pdf

4 <https://ec.europa.eu/eurostat/documents/4031688/9394119/KS-02-18-142-HU-N.pdf/5a4e298a-40e4-472c-aadf-5b831841a634>

5 GAFA vagy GAFAM néven is emlegetik őket: Google, Apple, Facebook, Amazon, de ideérthetjük a Microsoftot is.

követi a Berkshire Hathaway, több mint 320 milliárd USD-vel lemaradva, majd újabb technológiai cégek következnek, a Facebook, az Alibaba és a Tencent (PwC 2019: 19). A gazdaság, és különösen a technológiai vállalkozások piacának mozgékonyágát jelzi, hogy a 2019 első félévében közzétett listákon a vezető nyolcason belül a sorrend változik ugyan,⁶ de az őket a 9. helyen álló Jonson & Johnsontól elválasztó távolság is mintegy 60 milliárd USD. A PwC 2018-as összeállításának top 10 vállalata közül 2009-ben még egyik sem játszott ebben a kategóriában, az akkori listavezetők többsége pedig mostanra a top 100-ba sem fért be. Az elemzésben a gazdasági szektorok összehasonlításából az is kiderül (PwC 2018: 20), hogy a top 100-on belül a technológiai vállalkozások piaci tőkeértéke a legmagasabb, 4806 milliárd USD, és egy év alatt is 27%-os növekedést mutatott. A hagyományos ipari termelő vállalatok piaci tőkeértéke a technológiai cégek negyedét sem éri el, és mindössze 2%-kal bővült. Mindez arra utal, hogy a világgazdaságban valóban fundamentális átalakulás zajlott le, a gazdaság motorját már nem a hagyományos ipari termelés, hanem az adat mozgatja, alátámasztva a már szinte közhelyszerű megállapítást, hogy az adat az új olaj.⁷

A globális gazdaság makrofolyamatai és a technológiai cégek sikere mögött a világháló működése és emberek milliárdjainak a mindennapi tevékenysége, rutinja áll. A napi internethasználok száma 4,5 milliárd fölötti, akik több mint 6,6 milliárd Google-keresést indítottak, közel 6,5 millió blogbejegyzést készítettek, jóval több mint 700 millió Twitter-posztot tettek ki, az elküldött e-mailek száma 240 milliárd fölötti.⁸ Az aktív Facebook-felhasználók egy adott pillanatban mintegy 2,5 milliárdan vannak.⁹ Az internetes keresés, egy e-mail megírása, egy Facebook-poszt vagy Tweet-üzenet elkészítése mind időt vesz igénybe, akár a munka részeként, akár szabadidőben végzik. Mindez digitális lábnyomot hagy maga után, az adatok rögzítődnek a technológia cégek szerverein. Ha az idézett számokhoz hozzárendeljük azt az időt, amennyit ráfordítanak, óhatatlanul az következik, hogy a technológia és a gazdaság átalakulásával párhuzamosan alapvető változás megy végbe a társadalomban is. Megváltozik az idő strukturálásának módja, átrendeződnek az interperszonális kapcsolatok, a véleményalkotás módja és forrásai, sőt változik az emberek identitása – összefoglalóan azt mondhatjuk, korábban nem tapasztalt mértékű és sebességű társadalmi, kulturális átalakulás szemtanúi, résztvevői vagyunk.

Témánk szempontjából a legfontosabb kérdés az, hogy a hivatalos statisztika intézményrendszer, eszközei, fogalomrendszere alkalmas-e ezeknek a folyamatoknak a megragadására, mérésére, a társadalmi valóságról hiteles és jó minőségű statisztikai információk előállítására. A probléma gyökerét abban látjuk, hogy az intézményesülő, a nemzetközileg alkalmazott módszertani szabványokat sikeresen terjesztő hivatalos statisztika döntően hagyományos, mondhatni „analóg” adatforrásokra, kérdőíves adatgyűjtésekre és adminisztratív

6 <https://www.statista.com/statistics/263264/top-companies-in-the-world-by-market-value/>, <https://fxssi.com/top-10-most-valuable-companies-in-the-world> (letöltés ideje: 2019. október 28.).

7 <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data> De megfogalmazódnak ezzel ellentétes álláspontok is, lásd Martinez A.G. 2019. A vita akörül zajlik, hogy reálisak-e azok az elképzelések, amelyek azt vetik fel, hogy hasonlóan az Alaskan Permanent Fund eljárásához, amely éves kifizetést ad az alaszakai lakosoknak a kőolajbányászat jövedelmeiből, a Google-nak és a Facebooknak is fizetniük kellene felhasználóiknak a személyes adataik felhasználásából keletkező profitból. A szerző, aki korábban a Facebook monetizálásán dolgozó team tagja volt, természetesen elveti ezt a lehetőséget, érvelésében tagadja, hogy az adat összevethető lenne az olajjal.

8 <https://www.worldometers.info/hu/oldalról> 2020. május 4-én 22.50-kor letöltött adatok alapján.

9 <https://www.internetlivestats.com/> (letöltés ideje: 2020. május 4-én 22.50.).

adatokra támaszkodik. Attól, hogy ezeket a feldolgozás során digitalizálja vagy elektronikus formában gyűjti be, még megőrzi analóg karakterüket, ami alatt azt értjük, hogy nem közvetlenül az adatforradalom digitális valóságából származnak. Ez utóbbihoz egyelőre nincs, vagy legfeljebb sporadikusan van hozzáférése. Fontos leszögezni, hogy a hivatalos statisztika ebben a formájában is hasznos és nélkülözhetetlen. Amennyiben azonban elismerjük – és nem tehetünk másként –, hogy az adatforradalom pillanatról pillanatra bővülő társadalmi és gazdasági valósága létezik, akkor az előtte álló legeminensebb kihívás annak megválaszolása, hogy ezt miként vonhatja hatókörébe.

Mielőtt a válasszal kísérleteznénk, a fejezet elején hivatkoztam, kétségtelenül pozitív fejlemények mellett vessünk egy pillantást arra is, mi található a mérleg másik serpenyőjében, mert közel sem egyértelmű, hogy a világ alapvetően jó úton van-e, csak türelemmel ki kell várni a végét.

Az adatforradalom árnyoldala

Egyrészt a hivatalos statisztikák, mint például a foglalkoztatottság és a munkanélküliség, az árindex, a jövedelmi adatok vagy éppen a GDP, nem evidens módon, sőt tulajdonképpen azt mondhatjuk, egyre kevésbé tükröződnek vissza a személyes tapasztalatokban. Ennek csak az egyik oka, hogy viszonylag kevesen tanulmányozzák a hivatalos statisztikai adatokat, és a tömegtájékoztatókban, illetve a közösségi médiában ezekkel ellentétes információk terjednek (lásd erről: Rosling H., Rosling Rönnlund és Rosling O. 2018), másrészt – és ez utóbbi megállapításból következően – megrendült a hivatalos statisztikákba vetett bizalom. Erre később még részletesebben visszatérünk.

Nagy nemzetközi szervezetek által készített elemzések, jelentések is rámutatnak az adatforradalmat kísérő számos negatív társadalmi és gazdasági következményre. A Világbank 2016-os jelentése (World Bank 2016) is több kritikus észrevételt fogalmaz meg. Mindenekelőtt rögzítik, hogy a világ népességének nagyobb fele nem rendelkezik internetcsatlakozással, és jórészt érintetlen maradt a digitális forradalomtól.¹⁰ Az egyéni sikertörténetek ellenére

a technológia hatása a globális termelékenységre,¹¹ a szegények és a középosztály lehetőségeinek bővülésére és a számonkérhető kormányzás terjedésére mind ez ideig elmaradt a várakozásoktól. ... A digitális technológia megváltoztatja a munka világát, de a munkaerőpiacok polarizáltabbá váltak és nő az egyenlőtlenség – különösen a gazdagabb országokban, de egyre inkább a fejlődő országokban is. És miközben nő a demokráciák száma, a szabad és tisztességes választások aránya csökken. Ezek a trendek folytatódnak, nem a digitális technológia miatt, hanem annak ellenére (World Bank 2016: 2).

Az idézett rész fontos jelenségekre mutat rá. A II. világháború után a nyugati világ jóléti társadalmában kialakult egy erős és jómódú középosztály, és ez biztosította a stabil politikai berendezkedés alapjait is. A racionális jogállam több évszázados fejlődés után az 1980-as évek végére érte el legérettebb formáját, a Szovjetunió és a kommunista rezsimek összeomlása

10 2020 szeptemberében az internethasználók száma meghaladta a 4,9 milliárd főt, ez a világ népességének 63,2%-a. <https://www.internetworldstats.com/stats.htm> Letöltés ideje: 2020. november 18.

11 A ledolgozott munkaórákra vetített termelékenység mediánja növekedésének öt éves átlaga 87 ország adatai alapján 1973 és 2015 között kevesebb mint negyedére csökkent (World Bank 2016: 3).

teret adott olyan vízióknak, hogy bizonyos értelemben bekövetkezett a történelem vége (Fukuyama 1994). Azonban a globalizáció folyamata éppen ebben az időben lépett szintet (Sassen 2001), és ehhez társult az új évezredben a digitális forradalom, melyek alapvetően rendezték át a társadalmakat és a gazdaságot, beleértve a munkaerőpiac polarizálódását, a nyugati társadalmak középosztályának a felbomlását. A globalizáció és a digitális forradalom azonban lényegében érintetlenül hagyták az államok szerveződési módját, intézményeit, az állami bürokráciát csakúgy, mint a politikai intézmények formai kereteit. Az utóbbiaknál két új jelenséget is megfigyelhetünk. Egyrészt parlamenti erővé, sőt kormányzati szereplővé válnak a centrumtól nagy távolságban elhelyezkedő politikai csoportok, pártok, másrészt egyre nagyobb számban azonosíthatók az úgynevezett hibrid rezsimek, amelyek fenntartják ugyan a parlamentáris demokrácia egyes elemeit, így a rendszeres időközökben lebonyolított választásokat, de olyan módon manipulálják a szabályokat, hogy fenntartható legyen az államot foglyul ejtő, autoriter módon kormányzó párt és a mögötte álló érdekcsoportok hatalma. Az adatforradalom világa szerfőlött komplex, összefüggéseivel, hatásaival, rövid és hosszú távú következményeivel éppen csak ismerkedünk. Nagy bizonyossággal mondhatjuk azonban, hogy elsöpörte azt a társadalmi és politikai konszenzust, a nyelvet, fogalmakat, a közbeszédnek azt a módját, ahogyan felbomlása előtt a jóléti társadalmak a világot, annak problémáit értelmezték. És ez nem hagyja érintetlenül a racionális jogállam egyik legalapvetőbb intézményét,¹² és nyelvének, beszédmódjának integráns részét, a hivatalos statisztikát sem. A hivatalos statisztika osztozik az állam, a bürokrácia intézményeinek helyzetében, amennyiben éppen bürokratikus jellege, a hosszú idő alatt kialakult fogalomrendszerének kötöttségei miatt alkalmazkodása a példátlanul rapid változásokhoz finoman szólva sem problémamentes. Ezeket a korlátokat azért is érdemes figyelembe venni, mert a Világbank jelentése a digitális forradalom érzékelt hátulütőinek kezelésében – egyébként helyénvaló módon és érveléssel – éppen az államok cselekvésében lát lehetőséget. Az internetalapú gazdaság kedvez a természetes monopóliumoknak, a versenyalapú üzleti környezet hiánya még koncentráltabb piacokat eredményezhet, ahol a magasabban képzettek, a jobb kapcsolatokkal rendelkezők és ügyesebbek szerzik meg a haszon és az előnyök javát (World Bank 2016: 3). Egyúttal tartósan fennáll a digitális szakadék minden országban a nemek, földrajzi területek, korcsoportok és jövedelmi csoportok között. A kockázatok és a káros hatások csökkentése érdekében a jelentés szerint az államoknak erősíteniük kell az úgynevezett „analog kiegészítőket”, amelyek közül háromra helyeznek kiemelt hangsúlyt: a szabályozásra, az emberek készségeinek, jártasságának növelésére, hogy jobban élhessenek a digitális lehetőségek előnyeivel, és az elszámoltatható intézményekre, hogy az állam megfeleljen a polgárok szükségleteinek és igényeinek. A digitális technológia pedig megerősítheti az analog kiegészítőket, felgyorsítva ezáltal a fejlődés ütemét (World Bank 2016: 5). Közben a jelentés érzékenyen és pontosan mutat rá a technológiai forradalom kockázataira és káros következményeire, a megoldási javaslata idealizmussal terheltnek tűnik, és ellentmond az államok működéséről tett megállapításainak is. Mert nehezen képzelhető el, hogy egy politikai és gazdasági érdekcsoportok által foglyul ejtett állam az erőforrásait olyan célokra használná, melyek aláássák saját hatalmát. Sőt maga is kiemeli azt a kockázati tényezőt, hogy az államok és a gazdasági társaságok a digitális technológiát az állampolgárok ellenőrzésére

12 A későbbiekben még részletesebben foglalkozunk a racionális jogállam és a hivatalos statisztika kapcsolatával, alátámasztva az utóbbi fontosságáról tett állítást.

használják, nem pedig arra, hogy befolyáshoz juttassák őket. A technológia – és itt elsősorban az internetre gondoljunk – csak ott ad nagyobb beleszólási lehetőséget a polgároknak, ahol az államban már létezik az erre való nyitottság, fogékonyság. Ellenkező esetben az internet inkább növeli a kormányzatok kapacitását az állampolgárok ellenőrzésére, kontrolljára (World Bank 2016: 24–25). Annyit mindenképpen rögzíthetünk, hogy az adatforradalom egyik legfőbb kihívása a technológiai változások és az állami, társadalmi intézmények működésmódjának, reakciókészletének ehhez kevésbé illeszkedő viszonyában nevesíthető. És ez alól nem kivétel a hivatalos statisztika intézményrendszere sem, aminek hagyományos küldetéséből fakadóan informálnia kellene a döntéshozókat, a minden rendű és rangú felhasználókat a társadalomban és gazdaságban zajló folyamatokról. Ez a kijelentés tűnhet akár elhamarkodottnak, megalapozatlannak, sőt önellentmondásnak is, hiszen eddig is adatokra, statisztikákra támaszkodva igyekeztünk megragadni a jellemző trendeket. A kérdés azonban az, hogy hol keressük, miben véljük azonosíthatónak az adatforradalom legalapvetőbb következményeit, problémáit.

Heves kritikák minden oldalról

Az előző nagy, a társadalmat minden eresztékében megrázó átalakulás az ipari forradalom volt, ennek a történelmi korszaknak a helyébe lép, lépett az adatforradalom. Ahogyan az ipari forradalom kezdetét a vadkapitalizmus jellemezte, azt mondhatjuk, az adatforradalom jelenlegi szakasza is vad időszakában van, nevezhetjük akár *vad dataizmusnak*, vagy némi analógiával az ipari kapitalizmus kezdetére utalva *az eredeti adatfelhalmozás korának*, és mint fentebb állítottuk, a társadalom hagyományos intézményei, strukturái kevésbé képesek követni a változásokat és adekvát módon reagálni azokra. Az adat a legfőbb érték és a legfőbb erőforrás. Tevékenységeink többségének immár marad digitális lenyomata, ami a nyersanyag az adatalapú gazdaság lejárólüzemei számára. 2019-ben a világ népességének 66,6%-a rendelkezik – a csecsemőket is beleszámítva – mobilszelőkkel, ami 5,15 milliárd eszközt jelent,¹³ és ezek akkor is hagyhatnak digitális nyomot, ha épp nem használják őket, ehhez jönnek még az internetre kapcsolt számítógépek, a műholdak által érzékelt adatok, a térfigyelő kamerák, a bankkártyák, az okos mérők, okos órák, és a sort még hosszan lehet folytatni. Ilyen típusú adatokkal fizetünk a látszólag ingyenes szolgáltatásokért, mint a Google-kereső, vagy a közösségi média platformjai, ahol a felhasználók mindennapi életük apró rezdüléseit, nemritkán intim részleteit osztják meg írásban és képben. Ezek az adatok gyakorlatilag visszavonhatatlanul tárolódnak a szolgáltatók rendszereiben. Becslések szerint napjainkban az adatok értéke már meghaladja az ipari termelőkapacitások értékét. Az ezek fölötti diszponálásért kiélezett verseny, küzdelem folyik az államok és a transznacionális közösségi médiát üzemeltető és tartalomszolgáltató vállalkozások, technológiai cégek között.

Kulcskérdés, hogy milyen információkat hordoznak az adatok. A 2000-es évek elején a Google a befektetők nyomására a felhasználók adatainak elemzésével igyekezett hatékonyabbá tenni a hirdetéseket, hogy a felhasználók magatartásáról szerzett tudással minél célzottabb, az igényeiknek, szokásaiknak, preferenciáiknak leginkább megfelelő hirdetéseket

¹³ <https://www.bankmycell.com/blog/how-many-phones-are-in-the-world> (letöltés ideje: 2019. november 28. 15.20).

juttassanak el hozzájuk. Ebből alakult ki a digitális gazdaságnak az az iparága, ami a felhasználók magatartásáról gyűjtött és elemzett adatok alapján igyekezett minél rövidebb távra, a lehető legnagyobb pontossággal megjósolni az egyének magatartását, a legkisebbre szűkíteni az intervallumot a vállalkozások döntése és a felhasználók viselkedése között.¹⁴ Napjainkra már túl vagyunk egy újabb fordulaton, a cél immár nem csupán a magatartás előrejelzése, hanem inkább a befolyásolása, sőt irányítása. És ez az, ami az adatforradalom természetével, társadalmi hatásaival kapcsolatban a legélesebb kritikákat váltja ki.

A formálódó szép új világra minden oldalról zúduló kritikák egyik fókuszpontjában a nyugati típusú társadalomfejlődés kulcsfontosságú fogalmának, az egyéni szabadságnak, a magánélethez fűződő jogoknak a kérdése áll. Az elemzők között ki az államban, ki a cégekben látja a nagyobb fenyegetést, de van olyan is, aki azt vizionálja, hogy új korba lépünk, ahol már a személyiség sem több, mint algoritmusok összessége (Harari 2017: 275). Jack M. Balkin klasszikus libertáriánus nézőpontból vizsgálja az egyik legalapvetőbb demokratikus jog, a szólásszabadság állami korlátozásának eseteit az Egyesült Államokban (Balkin 2014). Külön pikantériája az elemzésnek, hogy valamennyi példája a mindmáig mintademokráciának tartott Egyesült Államokból származik. Az esetek alapján klasszifikálja a szólásszabadság korlátozásának régi (old school) és új (new school) technikáit. Álláspontja szerint a szólásszabadság demokratikus modellje, valamint a tudásnak és a véleményeknek a nyilvános szférája, ami legitimálja a demokráciát, széles körű intézményrendszeren nyugszik, mint a telefon-társaságok, könyvtárak, könyvesboltok, iskolák, egyetemek, széles sávú szolgáltatások stb., melyek alapvető jelentőségűek (Balkin 2014: 2301). Azonban a digitális infrastruktúrának ugyanazon jellemzői, amelyek demokratizálták a véleménynyilvánítást, egyúttal a leghatásosabb és legcsábítóbb célpontjai a vélemények regulálásának és ellenőrzésének. Nem lát semmi meglepőt abban, hogy a vállalkozások és a kormányzatok igyekeznek úgy formálni a digitális infrastruktúrát, hogy az a lehető legjobban növelje a kontrollt és a felügyeletet. A hosszú távú trend a szólásszabadság, valamint a véleményregulálás és ellenőrzés, a megfigyelés infrastruktúrájának az összekapcsolódása (Balkin 2014: 2305, 2310). Bár Balkin cikkében sokkal élesebben fogalmaz, ezekre a veszélyekre a Világbank jelentése is felhívja a figyelmet, nemzetközi példákra hivatkozva, amikor a kormányzatok blokkolják az internetet vagy szűrik a számukra nemkívánatos tartalmakat.¹⁵ Álláspontjuk szerint az internet szűrése és a cenzúra számos negatív hatással bír, így például miközben költséges és máshonnan von el forrásokat, az emberek nem tudnak nyíltan kicserélni és megvitatni gondolatokat, ami előfeltétele az innovatív és produktív társadalomnak (World Bank 2016: 26). Balkin meglátása szerint napjainkban, miközben tovább élnek a szólásszabadság korlátozásának régi módszerei, gondoljunk itt a televízió, a rádió, a nyomtatott sajtó kormányzati szócsővé degradálására, azok kiegészülnek az új technikákkal, mint a digitális hálózatok és a kapcsolódó szolgáltatások, az internetes keresők, a fizetési rendszerek, a hirdetőik felett gyakorolt ellenőrzés. Az új technikáknak három, gyakran együtt érvényesülő jellemzőjét emeli ki. Az első a közvetett cenzúra, amikor az állam a digitális infrastruktúra tulajdonosára gyakorol nyomást, hogy blokkolja, szűri vagy cenzúrázza a számára nemkívánatos tartalmakat. A második a public-private együttműködés, illetve a magánfelek bevonása a vélemények

14 Az előrejelzési iparban dolgozókat nevezi McGuinnis preditocracynak (McGuinnis 2012: 67).

15 2016-ban a Google 6951 megkeresést kapott kormányoktól, amelyben keresési tartalmak eltávolítását kérték, a legnagyobb számban Törökországból, az Egyesült Államokból és Braziliából.

regulázásába és felügyeletébe, amikor a magánszereplők közreműködnek azoknak a beszélőknek, oldalaknak az azonosításában, amelyeket az állam figyelni, szabályozni szeretne, vagy ki akar zárni. A kormányzat vegyesen használja az ösztönzés és az elrettentés módszerét, akár törvények alóli mentességet kínálva annak, aki készségesen közreműködik a felügyeletre és ellenőrzésre irányuló törekvéseiben. A magánfelek a jogi bizonytalanságok elkerülése vagy a normális üzletmenet feltételeinek biztosítása érdekében akár direkt kormányzati nyomás nélkül, kezdeményezően mutatkoznak segítőkésznek. A harmadik a digitális megelőző korlátozás kormányzatok által kidolgozott új formája, ami a magán-infrastruktúra tulajdonosait célozza, és széles körben elterjedt az Egyesült Államokban. A véleménynyilvánítás szabadságát korlátozó szájzártechnikák a szerző szerint teljesen szokványos és bürokratizált elemei a digitális felügyeletnek, rutinszerűek, és ráadásul láthatatlanok. Ezek alkalmazására példa a WikiLeaks-sztori, amikor kormányzati felhívásra az Amazon azonnal eltávolította a szervereiről a WikiLeaks-et, a domain szolgáltatók megszüntették a szervereik elérhetőségét, a PayPal, a Visa és a MasterCard felfüggesztette az átutalások teljesítését, az Apple pedig eltávolította az App Store-ból azokat az applikációkat, amelyek lehetővé tették a felhasználóknak a WikiLeaks elérését (Balkin 2014: 2328).

Egy későbbi írásában Balkin a szabad véleménynyilvánítást fenyegető veszélyeket vizsgálva a hangsúlyt az államról az algoritmos társadalomra helyezi (Balkin: 2017). Ez utóbbit úgy jellemzi, hogy a nagy közösségimédia-platformok beékelődnek a hagyományos nemzetállamok és az individuumok közé, és algoritmusokkal és mesterséges intelligenciával irányítják az embereket (Balkin 2017: 1). Az algoritmos társadalom célja a gyakorlati mindentudás, a képesség arra, hogy a lehető legtöbbet lehessen tudni arról, ki mit, mikor és hol tesz, valamint a képesség mindezek előrejelzésére. A gyakorlati mindentudás lehetőségéből következik egy fontos politikai kérdés: ki birtokolja és irányítsa ezt a mindentudást? (Balkin 2017: 6–7). Napjainkban kiélezett harc folyik a nagy nemzetközi digitális hálózatok és közösségimédia-platformok tulajdonosai, valamint a nemzetállamok között, akik szeretnék ellenőrizni és irányításuk alá vonni ezeket a túl nagy hatalomra szert tett szereplőket. Az egyének véleménynyilvánítási szabadságát tehát több irányból éri fenyegetés. Ilyen a Big Data problémája, ami az egyénekről gyűjtött óriási mennyiségű adattal megnyitja a felügyelet, az ellenőrzés, a diszkrimináció és a manipuláció lehetőségeit úgy az állam, mint a magántársaságok részéről¹⁶ (Balkin 2017: 3). Az egyének véleménynyilvánítási szabadságának jövőjét Balkin szerint a kormányzatok és a magántársaságok közötti harc kimenetele dönti el. Bármennyire is kritikus és szkeptikus az államok szerepével kapcsolatban, úgy látja, az államoknak ellensúlyt kellene képezniük a magánkézben levő ellenőrzési és felügyeleti technológiák fejlesztésével szemben. A törvényalkotásnak mint eszköznek az alkalmazása sem könnyű azonban, ha a technológiai óriások úgy identifikálják magukat, mint a Google: „az online világot nem igazán köti az állami jog ... ez a világ legnagyobb kormányzat nélküli tere”¹⁷.

Vannak, akik sokkal pozitívabban látják a véleménynyilvánítás lehetőségét és szerepét az online világban. McGinnis amellett érvel, hogy az internet, az „új média” megtöri a hagyományos média, a sajtó, a televízió és a rádió monopóliumát azzal, hogy minden korábbinál több ember számára teszi lehetővé véleménye kifejtését, terjesztését. Az előnyöket nem is

16 A 2016-os amerikai elnökválasztási kampányban Donald Trumpnak dolgozó Cambridge Analytica – a nyilvánosságra került információk szerint – a „This Is Your Digital Life” Facebook-applikáció 270 ezer felhasználóján keresztül több mint 87 millió Facebook-felhasználóról egyenként 4-5 ezer személyes adatot szerzett meg.

17 Eric Schmidt és Jared Cohen írják ezt *The New Digital Age* című könyvük első oldalán, idézi: Zuboff (2016: 3).

csupán a különböző nézőpontok megjelenésében látja, hanem különösen a szakértői blogok megjelenésében és működésében, ahol a hozzáértők nyilvános vitáiból, eszmecseréiből rendkívül kifinomult vélemények jöhetnek létre például közpolitikai kérdések kapcsán, amit az érdeklődők szabadon követhetnek. Az ilyen eszmecserék alkalmasabbá tehetik a társadalmat (és a kormányzatot) a közügyek tényszerűbb kezelésére, és ha bekerülnek a hagyományos médiába, még szélesebb közönséghez juthatnak el, ezenkívül támogatják a tényeken alapuló újságírást (McGinnis 2012: 64–66). Egy ponton hozzászól azonban, hogy a szabad beszéd korlátozása az interneten alássa a politikai rendszer és a közpolitikák döntéshozóinak elszámoltathatóságát. És ezzel tulajdonképpen visszakanyarodtunk Balkin aggodalmához és fenntartásaihoz.

Shoshana Zuboff alapvetően baloldali gondolkörre alapozva ostorozza az adatforradalom következményeit, amit következetesen felügyelő kapitalizmusnak nevez. Meglátása szerint a kapitalizmus a szemünk előtt ment át több átalakuláson. Régen a profit a termékekből és a szolgáltatásokból keletkezett, aztán spekulációból, most pedig az egyének adatain alapuló felügyeletből. Ez utóbbi mutáció magyarázhatja, miért maradt el mindedig, hogy a digitalizáció döntő hatással legyen a gazdasági növekedésre, ugyanis a benne rejlő lehetőségek nagy részét a profit alapvetően parazita formájává alakították (Zuboff 2016: 5). A kapitalizmus korábbi formáival szemben a digitális világban a felhasználók nem eladók, nem vásárlók, és nem termékek, hanem a szabad nyersanyag forrásai, akik táplálnak egy új típusú termelési folyamatot (Zuboff 2016: 5). Zuboff – másokkal együtt – a legaggasztóbb fejleménynek a nagy IT-szolgáltatóknál, különösen a közösségimédia-platformok üzemeltetőinél felhalmozott magatartási adatokat tartja, melyekkel kereskednek, és a felhasználók viselkedésének befolyásolására használják őket. Mindemellett kialakul egy alapvető egyensúlytalanság: a magánélethez fűződő jogok erodálódását a cégek üzleti titkainak fokozott védelme kíséri, és a magatartás fölötti kontroll gyakorlása és annak technikai teljesen láthatatlanok az érintettek számára. Alapesetben döntési jog, hogy valaki valamit nyilvánosságra hoz vagy titokban tart. A felügyelő kapitalizmus újraosztja ezeket a jogokat és a felügyelő rezsim intézményeinél koncentrálja őket, teljesen új dimenzióját nyitva meg ezzel a társadalmi egyenlőtlenségeknek, egyúttal roncsolva a demokráciát (Zuboff 2019: 263; Zuboff 2016: 7).

Természetesen vannak, akik sokkal több pozitívumot látnak az adatforradalom társadalmi hatásaiban, következményeiben. Alex Pentland egy új társadalomtudomány alkalmazásának szükségességéről beszél, amit Auguste Comte nyomán társadalmi fizikának (Social Physics) nevez. Ennek a feladata lenne, hogy éljen napjaink példátlan lehetőségével, és az adatok biztonságos használatával segítse az alkalmazkodást az adatok új világához, ahhoz a világhoz, amely tisztességebb, hatékonyabb, befogadóbb, és amely nagyobb lehetőségeket nyújt, mint bármi korábban (Pentland 2018: 1). Érdekes és izgalmas azonban, ahogyan ezt az utópiát összeköti azzal a gondolattal, ami szerint a nyugati társadalmak legnagyobb tévedése: a ragaszkodás ahhoz az ideához, ami szerint „racionális individuumok” lennének. Ennek mindkét elemét elveti; nem csupán racionálisak nem vagyunk, de még kevésbé individuumok, sokkal inkább egy szociális faj tagjai, ezért a kutatásoknak nem az egyénekre, hanem az egyének közötti kapcsolatokra kell fókuszálniuk, mégpedig a mindennapi élet legapróbb rezdüléseit is rögzítő adatokkal, beleértve a digitális kommunikáció minden típusát annak tartalmával együtt. Hasonló álláspontot fogalmaz meg Dessewffy és Láng is:

az új információs korszakban a társadalom a hálózatban folyó áramlások köré épül: tőke, információ, technológia, kódok és szimbólumok ezekben a hálózatokban áramlanak. A hálózati áramlás nemcsak az egyik lehetséges működési módja a társadalmi rendszernek, hanem a legfontosabb szerkezeti forma, amely meghatározza a gazdasági, politikai és szimbolikus szféra működését (Dessewffy és Láng 2015: 160).

Európai nagyvárosokban folytatott empirikus kutatásuk során Pentland és munkatársai azt állapították meg, hogy a társadalmi kapcsolatok elemzése *háromszáz százalékkal jobban jelezte előre az emberek magatartását a hagyományos társadalmi-demográfiai változóknál, mint amilyen a nem, a kor, az iskolai végzettség, a jövedelem* (Pentland 2018: 9). És ezzel ismét az emberek magatartásának előrejelzésénél, majd befolyásolásánál tartunk, noha Pentland azt reméli, ezzel elindítható egy okosabb, harmonikusabb társadalom építése. „Az individuum halála” című (Pentland: 2014) rövid írásának végén elismeri, hogy a magánélet védelme nélkül a vállalatok és a kormányok hatalma a magatartás manipulálására gyakorlatilag határtalan. Felvet ugyan egy technikai javaslatot, ami szerint az egyének a banki tranzakciókhoz hasonlóan hagyhatnák jóvá az adataik használatát, de ezt – jegyezzük meg – azóta sem alkalmazták sehol, és azzal tér ki a jelentéktelennek talán nem mondható probléma elől, hogy az egy másik történet.

Néhány tolakodó, nehéz kérdés

Óhatatlanul felmerül a kérdés, hogy vajon mi köze mindehhez a hivatalos statisztikának? Úgy vélem, sokkal több, mint ami esetleg az első pillanatban látszik.

Ha továbbra is fenntartjuk – és szerintem helyes, ha így teszünk –, hogy a hivatalos statisztika küldetése a tudósítás a társadalom, a gazdaság, a környezet állapotáról és változásairól, akkor megfelel-e vállalt feladatainak, ha nem reflektál a kapitalizmus természetének gyökeres átalakulására, a szabad véleménynyilvánítás korlátozására, az új szájzártechnikák terjedésére, az egyének magatartásának egyre hatékonyabb és nagyobb léptékű befolyásolására, sőt irányítására, a demokrácia intézményeinek, hosszabb időn át stabilnak mutatózó működési módjának megroppanására? Jogosnak tűnő ellenvetés lehet, hogy ezek a jelenségek kívül esnek a hivatalos statisztika hagyományos tárgyköreire, számára legfeljebb külső peremfeltételként jelentkeznek, és sokkal inkább a politikatudomány, a jogtudomány, a politikai szociológia érdeklődési körébe esnek. Egyébként is, a hivatalos statisztika az állami bürokráciának mint nagy társadalmi alrendszernek a része, ezért teljesen irreális elvárás, hogy a rendszer működési deficitjére vonatkozó kritikák itt fogalmazódjanak meg. Különösen az autoriter és hibrid rezsimok esetében jól látszik, mennyire rövid pórázon tartják a statisztikai hivatalukat.¹⁸ Az igazi probléma azonban nem az, hogy a statisztika közvetlenül reflektál-e

18 A Kínai Nemzeti Statisztikai Hivatal igazgatóját letartóztatták egy héttel azután, hogy sajtótájékoztatón a politikusoknak nem tetsző GDP-adatot jelentett be. Igaz, formailag korrupcióval vádolták. <https://geopoliticalfutures.com/china-admits-its-statistics-are-wrong/>

De érdekes összefüggéseket látnak az elemzők az Orosz Állami Statisztikai Szolgálat új vezetőjének 2016-os kinevezése és a statisztikai adatok módosítása között is. <https://www.obserwatorfinansowy.pl/in-english/macroeconomics/russian-trouble-with-statistics/>

a peremfeltételekben végbemenő változásokra, hanem hogy az adatai elemzésekor tekintettel van-e ezekre, illetve az adataiból lehet-e következtetni rájuk. Más megfogalmazásban és konkrétabban az vethető fel, hogy amennyiben helytállóak azok a talán disztópikusnak gondolt feltevések, amelyek szerint a nagy technológiai cégek az algoritmusokat és a mesterséges intelligenciát az egyének magatartásának befolyásolására, irányítására használják, akkor mit mér a statisztika? A statisztika végső soron az emberi magatartásokat – döntéseket, választásokat –, illetve ezek következményeit méri, legyen szó a mindennapi élet apróságairól (hamburger vagy hot dog) vagy cégek sokkal komplexebb döntési folyamatának eredményéről, mint például egy új termék bevezetése. Az egyszerűség kedvéért maradjunk az egyének mindennapi cselekvéseinél. Az individuum a társadalom szövetébe mélyen beágyazva éli mindennapi életét, ami egy bizonyos aspektusból döntések, választások folyama. Ezeket befolyásolják egyéni tapasztalatai, felhalmozott tudása, ismeretei (a múlt), aspirációi (a jövőre vonatkozó elképzelések) mellett az egyebek között a családtagokból, barátokból, ismerősökből, munkatársakból álló társadalmi háló tagjai, a Facebook-csoportban forgó vélemények, értékek, elvek, jól-rosszul megragadott érdekek, aktuális hangulata, fizikai és mentális állapota, pillanatnyi hatások, és a sor még tetszőlegesen folytatható. Az adott helyzetben választott magatartásnak van konzisztenciája, a magatartások sorozatának mintázata körvonalazza a személyiséget, és ez az, amire rámozdultak a Google, a Facebook, az Amazon és sokan mások. A társadalomtudományok alapvetése pedig abban áll, hogy a hasonló társadalmi-demográfiai jellemzőkkel leírható csoportok bizonyos helyzetekben mérhető valószínűséggel hasonló magatartást tanúsítanak. A családi szocializáció, az oktatási rendszer, a kultúra, az egyházak, a jogrendszer nyitották és zárták az adott szituációban és adott társadalmi helyzetben kívánatos, javasolt, helyes, elvárt magatartások körét. Ezt a részben az egész társadalomra érvényes, részben rétegspecifikus magatartási kánont fellazította a tömegmédiá színpadra lépése, de ennek is megvan a maga differenciáltsága, hiszen nem ugyanazok olvassák mondjuk a *The Guardian* vagy a *Frankfurter Allgemeine Zeitung*-ot, mint a *Daily Start* és a *Blicket*. Ugyanez a tagoltság megfigyelhető a televíziócsatornák és -műsorok nézői között is, miközben jól érzékelhető egy eltolódás a tömegmédiában a hírekről, eseményekről, tényekről való tájékoztatástól a kész vélemények tolmácsolása felé, vagyis nem „terhelik” a fogyasztót az önálló véleményalkotás bonyodalmaival, hanem azt igyekeznek megmondani, mit célszerű, ildomos, *nota bene* helyes gondolni a világ eseményeiről, dolgairól. Ezekre a folyamatokra ráerősített a reklámpiar berobbanása. A tömegmédiá legújabb generációs reklámjai már nem csupán arról igyekeznek meggyőzni a nézőt, hallgatót, hogy milyen kávét, teát, mosószert, fogkrémet, sört, telefont, autót stb. vásároljon, hogy a vaj vagy a margarin az egészségesebb, hanem életérzést tolmácsolnak. Akkor vagy trendi, ha ezt követed, így viselkedsz, így választasz – xy celeb ajánlásával.

Azt látjuk tehát, hogy az emberi társadalmaknak kialakulásuktól létfeltétele, hogy tagjait az éppen érvényesnek tartott normák követésére szorítsa, kezdetben a közösség kontrolljával, nyomásával, aztán intézményekkel, köztük a politikai hatalom és uralom intézményeivel, a vallási előírásokkal. Az eszközök egyre kifinomultabbá váltak, a nyers erőszak és a represszió helyébe a jogrendszer lépett, megnőtt az oktatás és a kultúra szerepe, míg eljutottunk a magatartás befolyásolását rendkívül szofisztikáltan és lágyan végző tömegmédiáig és reklámparig. A modern hivatalos statisztika mintegy 150 évre visszatekintő történetében az előállított információk minőségével, megbízhatóságával kapcsolatban soha nem merült fel fundamentális kétely amiatt, hogy a mért adatokban manifesztálódó magatartások

milyen befolyásoló hatásoknak tudhatók be.¹⁹ Sőt a nagy nemzetközi szervezetek, az ENSZ, a Világbank, az OECD, az EU erőfeszítéseinek köszönhetően mostanra kiépültek azok a statisztikai rendszerek, amelyek összehasonlítható, módszertanilag megfelelően harmonizált adatbázisaikkal lehetővé teszik a világban végbemenő folyamatok, változások tényeken alapuló követését. Soha nem tapasztalt bőségben rendelkezünk adatokkal a demográfiai folyamatokról, a gazdaságok állapotáról, az oktatás, az egészségügy, az élmezés helyzetéről, az ivóvízkészletekről, a kereskedelemről, és még hosszan sorolhatnánk. Mi a probléma? Van-e racionális értelme a hivatalos statisztika válságáról dörmögni azért, mert az adatforradalom következtében a minden rezdülésünket rögzítő technológia adatfolyamának elemzésével a technológiai óriások befolyásolni, netán irányítani akarják a magatartásunkat? Nincs ebben semmi új, mondhatnánk, ami meg új, az előrelépés, hiszen régen a normákat nem követőket megkövezték vagy máglyára küldték, most meg mindössze a felhasználói élmény fokozódásának kellemességéből zárják ki magukat. A statisztika attól még statisztika maradhat. Ezt a nézőpontot nem féltetlenül osztja mindenki, és a kérdés eldöntetlen marad mindaddig, amíg a hivatalos statisztikának és a kutatóknak nincs hozzáférésük a GAFA és más technológiai cégek adatvagyonához, hogy elemzésekkel lehessen eldönteni, milyen impulzusok, interakciók vezérlik a statisztika által mért magatartásokat, és ezek alapján lehessen megújítani a hivatalos statisztika fogalomrendszerét, osztályozásait, klasszifikációit.

Válságjelek

William Davies cikke (Davies 2017) meglehetősen nyugtalanító képet fest a statisztika helyzetéről. Nyitó megállapítása szerint a statisztika képessége arra, hogy pontosan megjelenítse a világot, hanyatlak.

A hivatalos statisztika tekintélyének csökkenésében, az adatai hitelességébe vetett bizalom megingásában több tényező együttes hatása is szerepet játszott/játszik. Az első ilyen a racionális jogállam, a nyugati típusú liberális demokrácia válsága. A történetük az elmúlt bő négy száz évben szorosan összefonódott, a XX. század második felétől pedig a kormányzati és közpolitikai célok kitűzésének és eredményességének vagy eredménytelenségének megítélésében a statisztikai mutatók játsszák a talán legfontosabb szerepet, gondoljunk a GDP-re, az államháztartás és a költségvetés helyzetére, a foglalkoztatottság és a munkanélküliség mérőszámaira, a demográfiai, oktatási, népegészségügyi jellemzőkre. Azt mondhatjuk, a hivatalos statisztika nagyjából az állam szolgálatlánya szerepkörébe került. A közelmúltban felerősödött populista politikai mozgalmak elitellenes frázisaiban a hivatalos statisztika vált az egyik célponttá. Szólamaik szerint a statisztikára alapozott politika elitista, nem demokratikus, és hanyag az embereknek a közösségükbe és nemzetükbe fektetett érzelmi erőfeszítéseivel szemben. A statisztika csak egy újabb eszköz, amivel London, Washington DC és Brüsszel privilegizáltjai megkísérlik ráerőltetni a világnézetüket mindenki másra. A számokat manipulálnak tartják, és elutasítják a kvantitatív tényekre építő elitizmust. A statisztika bonyolult módszertanon alapuló, adatokkal, tényekkel operáló nyelvezete szemmel láthatóan valóban irritáló azok számára, akik előnyben részesítik a ködös, mitikus narratívát a nemzet, a politikai

19 A jobbra politikai nyomásra visszavezethető „botlások” nem érintik jelen írás gondolatmenetét és távol esnek annak tárgyától.

közösség múltjáról, állapotáról és perspektíváiról (Davies 2017). Egy másik nézőpontból – és a szerző ezt osztja – a statisztika éppen nem elitista. Lehetőséget teremt arra, hogy az emberek, újságírók, politikusok véleményt cseréljenek a társadalomról mint egészről, mégpedig nem anekdoták, érzetek és előítéletek, hanem megalapozott ismeretek alapján (Davies 2017). Mindenesetre a populizmus térnyerése komoly kihívás a racionális jogállam és vele együtt a hivatalos statisztika számára, az utóbbi hitelességének aláásása, módszereinek, nyelvezetének, világlátásának kikezdése éppen az egyik harci terep.

A racionális jogállam válsága mögött azonosítható két további olyan társadalmi jelenség, amelyek közvetlenül is erodálják a hivatalos statisztikát és működési kereteit, az egyik a „post-truth”²⁰ valóság felemelkedése a közösségi média világában, a másik az identitáspolitikák eluralkodása. Az emberek érzelmeit célzó, a tényeken nyugvó racionalitást elitista machinációként megvető, az összeesküvés-elméleteket favorizáló, ködös mítoszokat terjesztő post-truth valóság egyszerű, könnyen befogadható és népszerű válaszokat kínál a gyorsan változó világ bizonytalanságai között nehezen tájékozódó közönségének. Számukra a statisztika félrevezetés, manipuláció, hazugság. És olyan nagyszerű érzés ebben egyetérteni, összekacsintva, hogy „de bennünket aztán nem fognak átverni”. Ha belegondolunk, az ilyenfajta kritikákban azért van némi igazság, még ha nem is abban a formában, hogy a statisztikusok szándékosan vezetnének félre bárkit is. Az állammal összefonódó története során a statisztika figyelme – a legutóbbi időkhöz – elsősorban a társadalomra mint *egészre* irányult, annál is inkább, mert tárgya szerint társadalmi tömegjelenségekkel foglalkozik. A globalizáció egyik hatásaként a társadalmak, és különösen igaz ez a térbeli társadalomszerkezetre, egyre mozaikosabbá váltak. Egymás szomszédságában található prosperáló és leszakadó városrészek, tömbök, ráadásul hosszabb időn át ugyanazon közösségeknél jelenhetnek meg egyik oldalon a hasznok, a másikon a költségek. Kevésbé lesz fogékony a foglalkoztatási adatok javulására és az átlagbérek emelkedésére az, aki másfél éve munkanélküli, napi megélhetési gondokkal küzd, és ugyanezt látja a szomszédjainál, a rokonságában és a sarki kocsmatörzsvendégeinél.

Napjainkra nem csupán a térbeli társadalomszerkezet vált mozaikossá, hanem feloldódtak, széttöredeztek azok a fogalmak is, amikkel az emberek önmagukat jellemzik, azonosítják.

A 60-as évektől kibontakozó identitáspolitika folyamatosan feszegette a statisztikai klasszifikáció kialakult rendszerét. A statisztikai adat csak akkor hihető, ha az emberek elfogadják a felkínált demográfiai kategóriák limitált körét, amit a szakértők alakítanak ki, nem a válaszadók. De ahol az identitás politikai ügyé válik, az emberek igénylik, hogy a saját fogalmaik szerint definiálják őket, mondjuk a gender, a szexuális irányultság, a faj, az osztály vonatkozásában. [...] A statisztikai klasszifikációk és definíciók hagyományos formáit szétfeszítik a változékony identitások, attitűdök és gazdasági pályák. A demográfiai, társadalmi, gazdasági változások egyszerű, elfogadott indikátorokban való megjelenítésére tett erőfeszítések veszítenek a legitimitációjukból (Davies 2017).

Ma az, hogy valaki valamelyik focicsapat drukkeré, egy popsztár rajongója, egy autómárka vagy teleföntípus híve, valamelyik influenszer követője, egy televíziós sorozat tántoríthatatlan

²⁰ Az *Oxford Dictionary* meghatározása szerint a kifejezésben az jelenik meg, amikor a közvélekedés formálásában az objektív tények kisebb befolyással bírnak, mint az érzelmek vonzereje vagy a személyes meggyőződések. <https://en.oxforddictionaries.com/definition/post-truth>

nézője, esetleg egy politikai szubkultúra szószólója, erős identitásképző tényező, ami meghatározó befolyással bír mindennapi életének szerkezetére, az időfelhasználására, viselkedésére, vásárlási szokásaira, véleményalkotására, interperszonális kapcsolataira, legyenek azok face to face alapon működő vagy az interneten szerveződött csoportok. Ezekkel a színes, jól átélhető identitásokkal szemben a statisztika klasszikus szürke sémái kiüresedtek, mert ugyan kinek a számára vonzó, hogy középfokú végzettségű, középkorú, betanított szellemi munkát végző alsó középosztálybeli nőként vagy férfiként lássa önmagát otthon a tükörben, miközben meccsre készülve kedvenc csapatának a harci színeit mázolja az arcára? A statisztikának természetesen az identitáspolitika felemelkedésében, az identitások elburjánzásában, a társadalmi tagoltság növekedésében sincs semmi felelőssége. Mulasztásként az róható a terhére, hogy klasszifikációiban, skáláiban nem követte elég érzékenyen ezeket a változásokat, aminek az árát ma hitelvesztésben fizeti meg. Ugyan jó ideje tesz komoly erőfeszítéseket arra, hogy egyre kisebb területi egységről egyre több adatot állítson elő, a klasszifikációval kapcsolatos kritika ettől még nyomasztóan igaz. Erre hívja fel a figyelmet Radermacher is: „az ex ante/top-down klasszifikációk nem fedik le az egyének identitását” (Radermacher 2019: 106).

Ha a racionális jogállam megrendülése, a post-truth valóság és az identitáspolitikák terjedése nem jelentene elegendő kihívást a hivatalos statisztika számára, az adatforradalom közvetlen következményeként szembesülnie kell a nagy transznacionális cégek támadásával és az általuk kikényszerített versenyhelyeztetel. A Facebook és a Google adatelemzői teljesen másféle „statisztikákat” állítanak elő, mint azok a szakértők, akiket most éppen széles körben elítélnek. Az új elemzőket az anonimitás és a titokzatosság sokkal nagyobb politikai befolyáshoz juttathatja, mint korábban bármely társadalomtudóst. A cégek számára az adatforrások és az alkalmazott módszerek titkossága olyan versenyelőny, amit önszántukból nem fognak feladni, ráadásul nincsenek olyan társadalmi mechanizmusok, amelyek megóvnák az embereket attól, hogy az így keletkezett információkra ösztönös reakciókkal, érzelmi előítéletességgel reagáljanak. Sőt az olyan társaságok, mint a Cambridge Analytica, kifejezetten igyekeztek felszínre hozni ezeket az érzéseket. Egyáltalán nem látszik, ezekből az adatfeldolgozásokból mikor lesznek ahhoz fogható hasznot hajtó információk a társadalom tagjainak közös ügyeiről folytatott diskurzushoz, vagy mondjuk a köz- és szakpolitikák nyilvános megvitatásához, értékeléséhez, amelyeket hiteles adatok formájában a statisztika nyújt (Davies 2017). Érdeemes rámutatni a GAFA-cégek és a statisztika elemzési módszereinek alapvető módszertani különbségére is. Az előbbieket a szinte felmérhetetlen mennyiségben és sebességgel keletkező mikroadatokra algoritmusokat eresztenek, mesterséges intelligenciát alkalmaznak, anélkül, hogy előzetesen kutatási tervekkel, hipotézisekkel, szakirodalmi áttekintéssel babrálnának, aztán igyekeznek értelmezni az eredményeket. Ezzel az induktív módszerrel szemben a statisztika deduktív megközelítést alkalmaz, vizsgálatait során a hosszabb időn át kialakított és alkalmazott fogalmi, klasszifikációs kereteket tölti ki új adatokkal (Kuonen 2018). Ez kétségtelenül jól szolgálja az időbeli összehasonlíthatóság, a társadalomban hosszabb idő alatt végbemenő folyamatok megragadásának szempontjait, ugyanakkor kétségeket támaszt eredményei relevanciájával kapcsolatban.

A fentiekből azt a következtetést vonhatjuk le, hogy a hivatalos statisztika számára több száz éven keresztül természetes működési közeget jelentő racionális jogállam legitimitásának megroppanása, a post-truth valóság felemelkedése, az identitáspolitikák eluralkodása

és a mindmáig domináns deduktív módszer együttes hatására a hivatalos statisztika által a társadalomról, annak folyamatairól, változásairól alkotott kép életlenné, homályossá vált, a hitelessége pedig csorbát szenvedett.

Walter Radermachernak, az Eurostat korábbi főigazgatójának „Hivatalos statisztika a Big Data korában, lehetőségek és fenyegetések” címen 2018-ban közzétett írása (Radermacher 2018a) számos ponton egybecseng Davies gondolataival. Fontosnak tartja a statisztika szerepének elemzését történeti kontextusban, a liberális kormányzati modellel való összefonódásra tekintettel, és jelzi a post-truth valóság terjedésének és az algoritmusok, a mesterséges intelligencia ellenőrizetlen használatának veszélyeit is. Cikkében a hangsúlyt nem a válságjelek listázására helyezi, sokkal inkább gondolkodási irányokat vet fel, javaslatokat tesz elvégzendő feladatokra. Ugyanakkor széles történeti, tudománytörténeti és kulturális összefüggésekbe ágyazva érvel mélyen elkötelezetten a hivatalos statisztika fontosságáért és jövőbeli szerepe mellett.

A hivatalos statisztikának alapvető szerepet kell játszania a modern társadalmakban, vezérelni a közpolitikákat, támogatni az üzleti döntéseket és hozzájárulni, hogy az állampolgárok elbírálhassák az elért előrelépéseket és összehasonlíthassák magukat a szomszédjaikkal (Radermacher 2018a: 227).

Rámutat a veszélyekre, kockázatokra is:

... a statisztikai tudás csodálatos erejének azonban vannak veszedelmei. Megismerő eszközből, amely emancipáló és részvételi, azonban igazi technokratikus tirannussá válhat, amely, különböző mértékben, menedzsmentelvek (tényekre alapozott döntéshozatal), főáramú gondolkodásmód vagy ideológia mögé rejtőzik” (Radermacher 2018a: 227).

Nem fejt ki, de itt gondolhatunk a kínai digitális diktatúra kiépülésére csakúgy, mint a technológiai cégek magatartást manipuláló gyakorlatára. Ezzel szemben nagyon impresszíven fogalmazza meg, miben látja a hivatalos statisztika hivatását, küldetését, és mit nem tart annak.

A hivatalos statisztika, és aminek maradnia is kell, a társadalomról rendelkezésre álló tudás megosztásának útja, a racionalitás egy eszköze, a döntéshozatal és a hatékonyság erősítésének egy módja. Bizalmat kell ösztönöznie, nem gyanakvást. Meggyőznie kell, nem nyomasztani. Segítenie kell, nem leigázni. Felszabadítania kell, nem alávetni. Feltárnia kell, nem félrevezetni” (Radermacher 2018a: 227).

Kérdés, hogy a diagnosztizált problémák és a statisztika jelen- és jövőbeli szerepére vonatkozó vizíó közötti feszültség miként oldható fel. Mint láttuk, a statisztika hitelvesztése és marginalizálódása mögött kimutathatók belső okok, mint a klasszifikáció és a deduktív módszer, de ezek egyrészt önmagukban nem okozói a válságnak, másrészt jelentős és összehangolt erőfeszítésekkel akár orvosolhatók a rendszeren belül. A külső okok – az adatforradalom következtében a tényeken alapuló igazság feltöredezése szubjektív valóságokra alapjaiban rengeti meg a valóság értelmezésének azt a hagyományát, amelyben a hivatalos statisztika eddigi története nem csupán mélyen gyökerezett, hanem annak egyik zászlóshajója volt – összehasonlíthatatlanul nagyobb szerepet játszanak, és ráadásul kívül esnek a hivatalos statisztika hatókörén. Egyúttal annak a racionális jogállamnak az autoritását és legitimitását kezdték ki, ami történetileg a hivatalos statisztika természetes működési közege.

Radermacher fontosnak tartja, hogy a hivatalos statisztika megőrizze alapvető értékét, a nyilvános, ellenőrizhető módszerekkel előállított statisztikai információk magas minőségét, miközben érzékenyen reagál a változó felhasználói igényekre és tartja magát a szigorú szakmai etikai kódexhez (Radermacher 2018a: 227). Változtatni kell egyebek között a kommunikáción is, az információkat a nagyon különböző felkészültséggel és statisztikai műveltséggel rendelkező csoportokhoz is el kell juttatni. Végül rendkívül ambiciózus feladatot jelöl ki a statisztikus közösség számára, amit értelmezhetünk úgy is, mint a külsőnek nevezett okokra adható reakció első és elengedhetetlen lépését:

a statisztikusok kötelessége a statisztika, a tudomány és a társadalom közötti kapcsolat feltárása, és vezető szerepet kell betölteniük az adatközpontúságba vetett túlzott bizalom kockázataira vonatkozó intellektuális reflexióban” (Radermacher 2018a: 230).

Útkeresés és reakciók a hivatalos statisztika intézményeiben

A jelzett folyamatok értelemszerűen nem hagyják érintetlenül a hivatalos statisztikai szolgálatokat sem. A 2010-es években a nagy nemzetközi szervezetek statisztikai részlegei sorra jelentettek meg dokumentumokat az adatforradalom, azon belül is leginkább a Big Data hatásairól, következményeiről, a benne rejlő lehetőségekről. Ilyen az ENSZ főtitkára által létrehozott Independent Expert Advisory Group on a Data Revolution for Sustainable Development (IEAG), a Data Revolution Report vagy a World Bank kiadványa (World Bank 2014). Az OECD honlapján kiemelten foglalkozik a Big Data témakörével, számos tanulmányt publikálva a gyakorlati alkalmazás tapasztalatairól (például Reimsbach-Kounatze 2015). A World Bank honlapján is számos elemzés található a Big Data alkalmazási területeiről, és helyet adnak egy Big Data blognak is.²¹

A számunkra legközvetlenebb hatással bíró intézmény, az Eurostat stratégiai dokumentumaiban számos érzékeny diagnózis olvasható korunk gyorsan változó viszonyainak a hivatalos statisztikára gyakorolt hatásáról. Az ESS²² Vision 2020 című, 2014-ben kiadott anyag ennek jó példája. Ebben négy fő kihívást azonosítanak.

1. Az adatforradalom: az új adatforrások óriási lehetőséget kínálnak a hivatalos statisztika időszerűségének és relevanciájának javítására csakúgy, mint a válaszadói terhek csökkentésére. Másrészt az új adat-előállítók részéről nő a verseny nyomása, ami megváltoztathatja a hivatalos statisztika szerepét. Kulcskérdés, hogy mi a jövője a megbízható és magas minőségű információinfrastruktúrának egy ilyen megváltozott környezetben.

2. Új típusú mérés szükségessége: új és komplex valóság alakult ki, amit meg kell ragadni a hivatalos statisztikának – nem engedve át a terepet az új adatelemzők manipulációinak.

3. A statisztika költsége: a döntéshozók kevésbé vannak tudatában, hogy a magas minőségű statisztika előállításához befektetések szükségesek. Folyamatosan csökkenő költségvetési források mellett kellene előállítani releváns és magas minőségű hivatalos statisztikákat.

4. Európa jövője: bizonytalan, milyen irányba halad az EU, a jelenlegi állapotot a gyengébb vagy erősebb integráció szakasza fogja követni.

²¹ <http://www.worldbank.org/en/search?q=big+data>

²² ESS – European Statistical System, magyarul ESR – Európai Statisztikai Rendszer.

További fontos gondolatok a dokumentumban az együttműködés és a kooperáció szükségességének hangsúlyozása úgy a nemzeti hivatalok között, mint a felhasználók minden csoportjával, kiemelve a kutatókat, elemzőket, döntéshozókat, újságírókat, illetve a statisztikai műveltség, jártasság terjesztése.

A hivatalos statisztika énképének legfontosabb eleme tehát a magas minőségű, megalapozott és hozzáférhető módszertannal készített statisztikák előállítása, amelyek reflektálnak a felhasználói igényekre. Ezt olyan értéknek tekintik, amelyet minden körülmények között meg kell őrizni, annál is inkább, mert ezen nyugszik tevékenységének hitelessége. Nincs ok arra, hogy ezt az alapállást megkérdőjelezzük. Ráadásul a közelmúltban keletkezett egy olyan pozitív tapasztalat, ami szerint a hivatalos statisztika nemzetközi intézményrendszere a legutóbbi, 2008-as válságot követően, építve a Stiglitz–Sen–Fitoussi-jelentés (Stiglitz, Sen és Fitoussi 2009) ajánlásaira, képes volt megújulni és új mérések, új indikátorok bevezetésével a makrogazdasági mutatók mellett a korábbiaknál sokkal részletesebb, érzékenyebb képet adni a különböző társadalmi csoportok élethelyzetéről, a jólétről, a szegénységről, a társadalmi egyenlőtlenségekről. Utólag is azt mondhatjuk, hogy a hivatalos statisztika reakciója viszonylag gyors volt és szakszerű, miközben jobbra a bevált eszközeit alkalmazva fejlesztett és a működő keretek közé integrált a felmerült igényeket kielégítő méréseket és outputokat. (Komoly attitűdbeli, módszertani megújulásnak tekinthető, hogy a lakossági adatgyűjtésekben minden korábbinál nagyobb teret kapott a válaszadók szubjektív vélekedésének vizsgálata.) Ez volt ugyanakkor az az utolsó felvillanás, amikor a statisztika az észlelt kihívásokra adott válaszokban proaktív, kezdeményező szerepet töltött be. A sikert még ki sem élvezhette, rögtön újra kihívott állapotba került az adatforradalom berobbanásával, és azóta tulajdonképpen fut az események után. A kihívottságba történő, jelenleg tartósan tűnő beragadtág leginkább az alábbi korlátokból fakad:

- *Szerkezeti korlát:* a hivatalos statisztika intézményeinek bürokratikus jellege. A bürokratikus jelleg adottság, a statisztikai szolgálat mindenütt része a kormányzati bürokráciának. Méretüknél fogva is nehezen mozdulnak, nem rugalmasak, minden irányváltás átfutási ideje hosszú. Az alapfeladatok ellátása lekötik az erőforrásaikat. Mind ez ideig nem kényszerültek azoknak a receptoroknak a kifejlesztésére, amelyek lehetővé tennék a releváns környezeti változások gyors leereagálását. Mentségül jegyezzük meg, eddig nem is szembesültek az adatforradalomhoz fogható kihívással. Ugyan az említett dokumentumokban is felmerül az igény a flexibilisebb szervezeti modell kialakítására, vagyis a korlát érzékelt, azonban ez egyelőre egyrészt a kívánság szintjén van, másrészt a megvalósítás beleütközik a további korlátokba.
- *Erőforráskorlát:* értve ezalatt pénzügyi, költségvetési és a humán erőforrásokban meglévő korlátokat is. A kormányzatok – különösen az 1990-es évektől – a menetrendszerűen érkező megszorításokat mindig ráterhelték a statisztikai szolgálatra is, amely kényszerűen alkalmazkodott a körülményekhez, leginkább arra koncentrálva, hogy az alapfeladataiban érvényt tudjon szerezni a maga elé támasztott magas minőségi követelményeknek. Az adatforradalom a gazdasági és társadalmi környezetben radikálisan átrendezte az erőviszonyokat. A statisztikai hivataloknak ma az informatikai és infrastrukturális fejlesztésekben a legtökeerősebb, legprofitábilisabb és nem utolsósorban legagresszívabb cégekkel kell versenyezni, mint a Facebook, a Google, az Amazon, a mobileszközök gyártói és az alkalmazások fejlesztői. De

ugyanaz a helyzet a humán erőforrásokkal is: a közsféra fizetései és karrierpályái valószínűleg sehol a világon nem versenyképesek. Nem hiszem, hogy van ma olyan statisztikai hivatal, amelyik a munkaerőpiacon adattudósokat tudna elhalászni a Facebook vagy a Walmart előtt. Az erőforráskorlát olyan adottság, amivel a jelek szerint továbbra is együtt kell élni. A feloldás a humán erőforrások területén partnerségen alapuló együttműködésekkel képzelhető el, amelyekben elsősorban kutatókat, külső szakértőket kellene bevonni fejlesztési programokba.

- *Percepciók korlát:* a hivatalos statisztika foglya a saját fogalmainak, tapasztalatainak, elköteleződéseinek, abból indul ki, hogy a munka, amit végez, az vitán felül értékes és nélkülözhetetlen a világ számára. Legfeljebb a világ nem tud róla. A problémát érzékeli, mégpedig sajátos dimenzióban, amikor a statisztikai kultúra, a statisztikai műveltség terjesztésének fontosságáról beszél. Vagyis amit csinál, amikor előállítja a magas minőségi követelményeinknek megfelelő statisztikai anyagait, az rendben van, a fogadókészség deficitjét pedig a felhasználók megfelelő jártasságának hiányossága okozza. A post-truth valóságba átcúszott vagy ott sajátos öntudatokra ébredt társadalmi csoportokra tényleg tud hatást gyakorolni a kínálatában szereplő statisztikákkal, csak éppen meg kell őket művelni egy kicsit statisztikailag? A Trump-jelenség, a Brexit, a brazíliai választás ezzel kapcsolatban elég sok kétséget ébreszt. (A hazai példákat most hanyagolom.)

A percepciók korlátát legszembetűnőbb megnyilvánulási formája, hogy az adatforradalom rendkívül komplex, fundamentális változásokat hozó következményeit a hivatalos statisztika hajlamos a Big Data jelenségre szűkíteni. Ez bizonyos értelemben természetes, hiszen hivatása szerint adatokkal foglalkozik, ennyiben a terep ráadásul ismerős. Megjelentek új adatforrások, a feladat ezek beépítése a hivatalos statisztika rendszerébe. A kezdetben fellángoló lelkesedést hűtötték ugyan a gyakorlati alkalmazásban mutatkozó nehézségek, de mára kialakult egy racionális, professzionális és fölöttébb innovatív hozzáállás. Az adatforrások expanzióban vannak, elképesztő a napi bővülésük is, a feldolgozókapacitások és technológiák velük haladnak, a távlatok beláthatatlanok, és mindez még költséghatékonyak is bizonyulhat. A Big Datáról azonban már jól látszik, hogy ez csak egy mennyiségi vetülete az adatforradalomnak, és szakmai csőllátás erre fókuszálni, ha a hivatalos statisztikát szeretnénk pozicionálni, megtalálni a jövőbeli helyét, szerepét. Mivel a hivatalos statisztika missziója továbbra is az, hogy a világunkról, annak társadalmi és gazdasági, környezeti állapotáról és változásairól adjon hiteles képet, nem eshet abba a csapdába, hogy a Big Data eltakarja előle a valóságot, bármekkora behemót is. Márpedig a dolog pillanatnyilag úgy áll, mintha a legfőbb probléma a Big Data statisztikává szelídítése lenne.

Az Európai Statisztikai Rendszer intézményei kettős nyomás alatt dolgoznak. Egyrészt a rendszeres adatgyűjtésekre támaszkodva elő kell állítani azokat a kulcsfontosságú mutatókat, indikátorokat, amelyek nélkülözhetetlenek a kormányzatok számára: negyedévente GDP-t, havonta a foglalkoztatottságra és munkanélküliségre vonatkozó adatokat és a fogyasztói árindexet, a háztartások jövedelmének és fogyasztásának indikátorait, az államháztartási adatokat, a mezőgazdasági információkat, stb. Ez olajozottan, megbízhatóan és megfelelő

minőségbiztosítási rendszerbe ágyazottan működik. Az Eurostat egy-egy témára fókuszálva sorra jelentet meg színvonalas, magas minőségű elemzéseket.²³ Saját hagyományos feladatkörén belül mérhető eredményeket ért el abban, hogy minél szélesebb felhasználói kör igényeinek megfelelő, magas színvonalú statisztikai információkat állítson elő. Ugyanakkor adekvát válaszokat kellene adnia a társadalmi, politikai, kommunikációs környezetében végbement változásokra, és ebben talán távolabb van a céltól, mint az ESS Vision 2020 megfogalmazása idején. Hiába vannak eredmények mondjuk a Big Data statisztikai célú felhasználásában, a post-truth valóságban és a fake news világában, a közvélemény, de még a politikai döntéshozók ingerküszöbét sem éri el a mégoly kitűnő statisztikai elemzések –feloldódnak az internet világának mindent átható zajában, zöreijében.

Az elmúlt mintegy két évben több Eurostatos dokumentumban mutatkoznak jelei annak, hogy oldódnak a percepciók korlát kötöttségei. Egy az European Statistical System Committee által 2017-ben elfogadott helyzetjelentésben több fontos megállapítás szerepel. Kiemelik, hogy az interneten és a közösségi médiában egyre gyorsabban egyre több statisztika jut el a felhasználókhoz, akiknek a többsége nem rendelkezik statisztikai jártassággal, és nem tudják megkülönböztetni a megbízható hivatalos statisztikai információt a hamis adatoktól.

A felhasználók bizalmát veszélyeztetik a terjedő visszaélések a statisztikával (mint a félreértelmezés, az előre meghatározott üzenethez illesztett szelektív bemutatás, stb.) és a hivatalos statisztikán kívül generált megbízhatatlan adatok burjánzása. Minél több médiafigyelem irányul az ilyen manipulatív üzenetekre és hamis hírekre, annál kevesebb ember képes különbséget tenni az adatok minőségében és felismerni a megbízhatóakat. Következésképpen bizalmatlanná válhatnak minden további, a sajtóban és a közösségi médiában felbukkanó számokkal kapcsolatban, még ha azok hivatalos forrásból származnak is. Ez az úgynevezett „post-truth valóság” aláássa azt a szaktekintélyt, amit a hivatalos statisztika mint hiteles ténylapot magának mondhatott. ... Ez a post-truth környezet által előállított tendencia leginkább a mindennapi embereket érinti, akik nem rendelkeznek azzal a felkészültséggel, amivel megítélhetnék a számok és forrásaik megbízhatóságát. Ezzel szemben a professzionális felhasználók (így a közpolitikai döntéshozók, tudósok, kutatók és a média) bizalma magas maradt, bár ők sem mindig képesek/akarják megérteni a számok üzenetét és a statisztikai szervek által alkalmazott adat-előállítási módszerek értelmét (ESS Committee 2017: 4).

Helyesen mutatnak rá arra a kulcsfontosságú problémára is, hogy a közösségi médiában, a mobilhálózatokban, az okos eszközök használatával és más automatikus adatgenerálók által keletkező adatok gyakorlatilag 100%-ban a magánszektor tulajdonában vannak, és a statisztikai szervezeteknek nem áll módjukban sem az, hogy kísérletezzenek ezekkel, sem az, hogy beépítsék őket a szokásos statisztikai folyamatokba (ESS Committee 2017: 4–5). Erre a tényre egyébként felhívta a figyelmet a Világbank jelentése (WB 2016: 28) és több más elemző is.

Az ESS VIN munkacsoportja az Eurostat 2021–2027-es Európai Statisztikai Programjához fogalmazott meg átfogó innovációs javaslatokat (ESS VIN: 2019). Kiemelkedő jelentőséget tulajdonítanak a jobb kommunikációnak úgy a felhasználókkal, mint a hivatalos statisztika terjesztőivel: kutatókkal, oktatókkal, újságírókkal és civil szervezetekkel. Felve-tik a felhasználói csoportok jellemzőinek figyelembevételét a statisztikák előállítása során,

23 <https://ec.europa.eu/eurostat/publications/statistical-working-papers>

mégpedig ún. perszónák létrehozásával. Ez a megközelítés és technológia tulajdonképpen erősen hasonlít az internet célzott, személyre szabott hirdetéseihez, azzal a nem csekély különbséggel, hogy nem manipulálni akar, hanem tájékoztatni. Fontos új eleme lehet a hivatalos statisztikai szolgálatok tevékenységének az ún. „tényellenőrzés” (fact-checking). Azon túl, hogy a statisztikai szolgálatoknak reagálniuk kell a statisztikák súlyos félreértelmezésére, megtévesztő felhasználására, ez lehet egy olyan szolgáltatás, amit más, tényeket ellenőrző szervezeteknek nyújtanak, továbbá az ESS-en kívül keletkező hasznos, megbízható adatok alkalmazásában támogatást nyújtanak a döntéshozóknak (ESS VIN 2019: 4).

Az egyetemekkel, kutatóintézetekkel, a felhasználókkal és a döntéshozókkal kialakítandó új típusú kapcsolatok kiépítése és működtetése visszatérő eleme úgy az Eurostat, mint számos nemzeti intézmény stratégiai dokumentumaink. A konkrét lépésekre példa az Eurostat által 2018-ban közzétett tanulmánykötet (Eurostat Outlook Report 2018), amelyben széles kört lefedő elemzők, kutatók, felhasználók írásai szerepelnek a hivatalos statisztika aktuális, helyenként égető problémáiról. A kötet érzékletes helyzetképet ad arról, hol tart a hivatalos statisztika az adatforradalom kihívásaira adható válaszokban. A lényegre törő írásokról összefoglalóan azt mondhatjuk, hogy többnyire egy-egy szakterületre fókuszálva, más esetekben elvi kérdéseket boncolgatva adnak diagnózist. A hangsúly tehát a diagnózison van, ami nagyon fontos fázisa a problémamegoldásnak, de még nem maga a megoldás. A szerzők egyebek között a nemzetközi vándorlás, a globalizáció, a munkaerőpiac, a fenntartható fejlődés, a digitalizáció témakörében rámutatnak a jelenleg előállított statisztikák korlátaira, és javaslatokat tesznek módszertani fejlesztésekre, új statisztikák kidolgozására, új adatforrások integrálására. A jó hír tehát az, hogy a kötet demonstrálja: a felhasználói, elemzői oldal és az Eurostat szakértői között folyik a párbeszéd, a közös gondolkodás. Az még nyitott kérdés, hogy ennek eredményei mikor mutatkoznak új vagy megújuló hivatalos statisztikai információkban.

Corine Eyraud a fentebbi fejtegetésekkel azonos következtetésre jut a hivatalos statisztika helyzetével kapcsolatban (Eyraud 2018). Korunk post-truth és intellektualizmusellenes vonásainak kibontakozásával párhuzamosan nő az emberek gyanakvása, a fenntartás a statisztikával szemben.

Nem biztos, hogy ez a szkepticizmus új és terjedő jelenség az emberek között, ami azonban elég bizonyos, hogy a bizalmatlanságot a szakértelemmel szemben világszerte egyre inkább a politikusok gerjesztették, és ez egyre nagyobb médiafigyelmet kap. Beláthatjuk, hogy politikusok vagy menedzserek (a köz- és a magánszférából egyaránt) rendszeresen használják a statisztikát az emberek félrevezetésére azért, hogy tényekre alapozottnak beállított politikai vagy gazdasági döntéseket igazoljanak, vagy oly nehezen érthetővé teszik őket, hogy a nem szakértő emberek ne legyenek képesek megkérdőjelezni a választott irányokat és döntéseket. Így a statisztika az uralmi rendszer részévé vált. Az első tennivaló az állampolgárok és a statisztika közötti szakadék áthidalása, ami gátat vethet annak, hogy ilyen módon és ilyen célokra használják. Ez azonban messze túlnyúlik azon a körön, amire a hivatalos statisztikának önmagában hatása van (Eyraud 2018: 103).

Márpedig – mintegy rímelve erre – Radermacher azt írja, ha egyszer elveszik a statisztika iránti bizalom, akár évtizedekig is eltarthat annak újjáépítése. A hiteles és releváns statisztikák jövőbeli előállítása érdekében a statisztikai intézményeknek át kell tekinteniük és revízió alá kell venniük a jelenlegi statisztikai irányítást a jövőbeli fejlesztések érdekében. Az is világos,

hogy a bizalom kérdése messze túlmutat a statisztikába vetett bizalmon. Végző soron a megbízható statisztika elengedhetetlen ahhoz, hogy magának a társadalomnak a működése megbízható legyen (Radermacher 2018b: 110).

Remélhetőleg elegendő oldalról és elegendő alkalommal jártuk körül a hivatalos statisztika helyzetét és annak minden kihívását ahhoz, hogy megkíséreljük összefoglalni, amit észleltünk.

- a) Megfontolandó érvek szólnak amellett, hogy az adatforradalommal összefüggésben a digitalizált világ minden rezdülését rögzítő adatfolyamra és a feldolgozásukat szolgáló algoritmusokra és mesterséges intelligenciára építve a kapitalizmus új szakaszába, a *felügyelő kapitalizmus* korszakába lépett. Ugyanazok az eszközök, rendszerek, amelyek a szólás és a szabad véleménynyilvánítás szabadságát biztosítják, egyúttal eszközei a szólásszabadság korlátozásának, a szájártechnikák alkalmazásának.²⁴ Miközben a kormányzatok és a technológiai óriások ebben készek együttműködni, egyúttal kiélezett versenyben állnak egymással a korlátozó, manipulatív, a véleményeket és magatartásokat irányító hataloméért. A szólásszabadság mellett sérülnek a demokrácia más alapintézményei is, mint a választások szabadsága, a döntések számonkérhetősége. Ez utóbbi eszközéül kellene szolgálniuk a mindenki számára egyenlően hozzáférhető hivatalos statisztikai információknak, amik elengedhetetlen, egyelőre mással nem pótolható eszközei a köz ügyeiről folytatott racionális és nyílt diskurzusnak. Jogos aggályok merülnek fel azonban azzal kapcsolatban, hogy a felügyelő kapitalizmusban mennyi tér marad az ilyen típusú eszmecseréknek. A trendek a legkevésbé sem kedveznek annak a politikai kultúrának, annak a döntéshozatali gyakorlatnak és közbeszédnek, amiben a hivatalos statisztika integráns szerepet töltött be.
- b) A néhány jelszót, ügvet, vagy éppen post-truth hívószót zászlóra tűző *populista politikai mozgalmak* a régi establishmentet mint elitista, nagyképű uralmat támadják, amelyik érzéketlen „a zemberek” igazi problémáira, akikkel elveszítette a kapcsolatot, a nyelvüket sem beszéli. Fő célpontjaik közé tartoznak a szakértők, beleértve a statisztikusokat is, mint a félrevezető praktikák részeseit, így a racionális jogállam megroggyanása magával rántja a hivatalos statisztikát is, mint működésének egyik emblémáját. Tekintélye, hitelessége csökken a közvélekedésben, párhuzamosan annak az érzetnek a terjedésével, hogy a statisztika nem tükrözi vissza a mindennapi élet tapasztalatait az árakról, a munkanélküliségről, a keresetekről stb. A szakértő felhasználói kör ugyan továbbra is bízik a hivatalos statisztikákban és professzionális felhasználó, azonban ők ugyanúgy a támadások célpontjai, ezért nem segíthetik a statisztika hitelességének visszaállítását. Ugyanakkor a hatalomra kerülő populisták irányzatok, az államot foglyul ejtő gazdasági és politikai érdekkörök, a vezérelvű hibrid rezsimek elsősorban a hatalmuk megőrzésére törekednek, nem az egész társadalmat átfogó, nyilvános, számonkérhető és vitatható programok megvalósításában érdekeltek; számukra a statisztika, a benne felhalmozott tudás nem képvisel értéket, legfeljebb a propaganda egyik eszközeként tekintenek rá.

²⁴ Ennek az írásnak nem tárgya a szólásszabadság terjedelmének és határainak vizsgálata. Fontos rögzíteni azonban, hogy nem a terroristák, bűnözők, pedofilhálózatok stb. szólásának és szervezkedésének a szabadságáról van szó.

Fellazul tehát az a szövetség, egymásra utaltság, ami évszázadokon át a statisztika természetes működési közegét jelentette.

- c) Az internet és a közösségi média határtalan világában burjánzanak a *post-truth* tartalmak és összeesküvés-elméletek. Ezekhez jellemzően alacsony a belépési küszöb, nem kívánnak érdemleges felkészültséget és erőfeszítést, viszont egyszerű, könnyen érthető és könnyű azonosulást kínáló magyarázattal szolgálnak bármire, formát adva érzelmeknek, indulatoknak, és elutasítanak minden eltérő nézetet. Éles ellentétben állnak a statisztika logikus fogalomrendszerével, nomenklatúráival, transzparens, noha kétségtelenül nem mindig triviális módon érthető módszertanával, és egyúttal folyamatos kihívást jelentenek neki. A *post-truth* valóság digitális tartalmainak áradatában a statisztikai információ nagyon távol esik a fősodortól, inkább jelentéktelen hordaléknak számít. Ebben a (szubjektív) valóságban a szerepe még marginálisnak sem mondható, azonosíthatatlan.
- d) Az *identitáspolitikák* eluralkodása megkérdőjelezi a hivatalos statisztika hagyományos társadalmi-demográfiai változóinak relevanciáját.²⁵ Ez felerősíti a *b)* pontban jelzett tüneteket, táptalaja az elitizmusra hivatkozó kritikáknak és egyik forrása a hitelvesztésnek.
- e) A *globális technológiai cégek*, a GAFÁ-óriások és tartalomszolgáltatók több szempontból is egyenlőtlen versenyre kényszerítik a hivatalos statisztikát. Egyfelől aránytalanul nagyobbak az algoritmusok és a mesterséges intelligencia fejlesztésére rendelkezésre álló forrásaik, amiket mindenféle kontroll és transzparencia nélkül alkalmaznak. Az adatok birtokosai elemző, feldolgozó tevékenységükkel oda jutottak, hogy képesek az emberek magatartását közvetlenül befolyásolni. Másfelől, ha lehet, még ennél is nagyobb probléma, hogy az adatforradalom, a digitalizáció során keletkező gyakorlatilag valamennyi adattal magántulajdonukként rendelkeznek, ezekhez a hivatalos statisztikának nincs hozzáférése. Márpedig az adatforradalom társadalmi változásaival konform klasszifikációkat, osztályozásokat, nomenklatúrákat a hozzáférés hiányában nehéz lenne elkészíteni. A jogi feltételeket állami szabályozás alakíthatja ki, ami egyúttal a vad dataizmus és az eredeti adatfelhalmozás korának konszolidációját is hozhatná. Kérdés azonban, hogy az *a)*, *b)* és *c)* pontban írtak alapján a kormányzatok érdekükben állónak tekintenek-e a jogi beavatkozást és az ezzel járó konfliktust, miként az is, hogy adott esetben a felhasználóknak, az internet és a közösségi média népének az identitásában a kormányzat vagy a szolgáltató iránti lojalitás az erősebb, különös tekintettel arra, hogy a véleményeket manipuláló és a magatartást irányító eszközök terén a technológiai cégek lépéselőnyben vannak.

Az előző pontokban említett jelenségek, tényezők sok szálon összefüggő, erős, zárt rendszerre álltak össze. Akárhogyan forgatjuk, rendezzük őket, kölcsönösen következnek egymásból és erősítik egymást. Egyenként és összességükben is kedvezőtlen hatással vannak a hivatalos statisztika elismertségére, hitelességére, az adatforradalom előtt betöltött szerepének

²⁵ Érdemes lenne megfontolni, hogy például a politikai közvélemény-kutatások pontatlanságai, tévedései mennyiben vezethetők vissza arra, hogy a mintavételnél még mindig a hagyományos társadalmi-demográfiai változókra támaszkodnak.

fenntarthatóságára. Alapjaiban rendeződtek, rendeződnek át azok a társadalmi és politikai viszonyok, amelyekben a hivatalos statisztika a felhalmozódott tudással és szakértelemmel a közjó egyik letéteményese volt. Rögzítsük, hogy a statisztikának ebben nincs felelőssége, a változásoknak ennek ellenére kárvallottja. Felelőssége legfeljebb annyiban van, hogy késve reagál, jó ideig beleragadt a Big Data statisztikává szelídítésének feladatába, miközben klaszszifikációinak, osztályozásainak, módszereinek megújításával mindmáig adós. Ma már ezek is napirendre kerültek mint problémák, bár a konkrét megoldások még váratnak magukra. A stratégiai dokumentumokban és szakértői anyagokban az európai hivatalos statisztika képviselői eljutottak a helyes diagnózisig, valamennyi, a hivatalos statisztika elismertségének, tekintélyének, szavahihetőségének megőrzésére, helyreállítására tett javaslatokról elmondható, hogy helyes. Ezek mind szükségesek, azonban minden bizonnyal nem elégséges feltételei a kitűzött célok elérésének.

A skolasztikusoktól az adatforradalomig. És merre tovább?

Ha valakinek statisztikai információra van szüksége, szeretné tudni, mennyi a háztartások zöldségfogyasztása, vagy mennyit költenek villamos energiára, mennyi búza termett tavaly, miként alakult az infláció, az államháztartás helyzete, vagy éppen a külkereskedelem egyenlege, ma már néhány kattintással megtalálhatja az interneten. Ez hasonlóképpen természetes, mint felkattintani a villanykapcsolót. Kattintás – statisztikai adat. Kattintás – felkapcsolódik a lámpa. Mindkettő magától értetődő mozdulat. Az emberek döntő többsége számára ugyanakkor reflektálatlan abban az értelemben, hogy amiként nem gondolkodnak el a villanyvezetékben a kapcsoló által kiváltott fizikai folyamatokon, ugyanúgy nem merülnek bele azokba a módszertani, történeti, kulturális, tudománytörténeti, matematikai, szociológiai, közgazdasági, informatikai stb. kérdésekbe, amelyek eredményeképpen megtalálják (nem találják meg) a keresett információt. Ez a reflektálatlan bizalom alapesetben része a rendszernek, és működik mindaddig, amíg valami nem szaggatja meg a bizalmi hálót, illetve amíg nem történnek változások az összekapcsolódó, egymásra épülő elemek rendszerében.

A statisztikai adatokkal kapcsolatos viták, az, hogy helytálló információkat közölnek-e a valóságról, két fő nézőpontból merülhetnek fel. Az egyik magának a mérésnek a pontossága, helyessége. Ha a mérés tárgya egy jól meghatározható dolog, mint egy adott területen a búzatermés mennyisége vagy a sertésállomány nagysága, sem a búza, sem a sertés fogalma nem igényel további tisztázást, legfeljebb az alkalmazott módszer, a számbavétel pontossága vehet fel kérdéseket. Noha elképzelhetők tévedések, akár a csalás szándéka is, a problémák magának a mérésnek a körén nem terjednek túl. A másik, sokkal bonyolultabb eset az, amikor annak az objektumnak a léte és fogalma, relevanciája, a valóságot leképező ereje kérdőjeleződik meg, ami konvencióként beépült a statisztika és a társadalomtudományok eszköztárába, fogalomrendszerébe (Desrosières 1998: 1). A 2008–2009-es válságot követően viták indultak a GDP-nek mint széleskörűen használt makrogazdasági mutatónak a korlátairól, revíziójának szükségességéről (Stiglitz, Sen és Fitoussi 2009). Magától értetődő az igény, hogy a mérések pontosításával, a konszenzusosan használt fogalmaknak, a statisztika nyelvének a folyamatos finomításával, megújításával, ha kell, a régi fogalmak újakra cserélésével a statisztika megőrizze a képességét a körülötte változó világ folyamatainak megragadására, tükrözésére. Végül is így tud megfelelni lényegi feladatának, a valóság tömör,

számszerű jellemzésének. Amíg malacokat kell számolni, addig nem tornyosulnak a látóhatáron megoldhatatlan akadályok. A szilárd talaj akkor illan el a lábunk alól, ha elkezdjük firtatni a jellemezni akart valóság mibenlétét. A kérdést ugyan passzolhatnánk a filozófusoknak, de ettől a definíciónk fele még a levegőben lógna.

Desrosières nagyszerű munkájában (Desrosières 1998) a statisztika, a statisztikai kultúra történetét a tudományos kultúrába, azt pedig az általános kultúrába ágyazva vizsgálja. Sok egyéb mellett felidézi, hogy a statisztikai gondolkodás fejlődése során a matematika, a csillagászat, az orvosi megfigyelések, a racionális közigazgatás felől érkező igények beépítését, felhasználását a statisztikai eljárásokban rendszerint viták kísérték, és ezeknek a vitáknak a gyújtópontjában visszatérő hullámokban a statisztikai adat és a valóság viszonya állt. A probléma gyökere mélyen visszanyúlik az európai bölcsélet múltjába. A skolasztikus filozófiában élénk vita folyt az ún. realista és nominalista irányzatok között arról, hogy mi tekinthető valóságosnak. A realisták álláspontja szerint „az általános éppúgy ténylegesen létezik, mint az egyes”, ezzel szemben a nominalisták nem fogadták el az általánost létezőnek, „a nominalista ... az egyetemesnek soha nem, csakis az egyedinek tulajdonít valódi létezését” (Panofsky 1986: 9). Ezt a nézetet képviselte a jeles nominalista gondolkodó, William Occam is. Occam és a nominalisták álláspontja nem csak tudománytörténeti szempontból érdekes, és nem is azért, mert ez nyitott utat a modern természettudományos gondolkodásnak, hanem jelentőséggel bírhat az adatforradalomban a statisztika előtt álló feladatok feltárásában is. Időközben a realista gondolatkör a középkori ontológiai és idealisztikus változattól indulva egy anyagszerűbb és empirikusabb irányba mozdult. Ez a fejlemény eszközöket igényelt a dolgok összekapcsolásához egy felsőbb szintű realitásba, és ilyennek bizonyult a 18. századi naturalisták terminológiája (Linné) és Quetelet statisztikai átlaga a 19. században. A realizmusnak ez az eszközökkel jól ellátott új fajtája, feljavítva az individualista nominalizmussal – ami elengedhetetlen volt az egyes elemek rögzítéséhez és kódolásához –, létrehozta azt a húzóerőt, ami jellemző a tudás statisztikai formájára (Desrosières 1998: 69–70). „Ez a húzóerő lényeges a statisztikai munkában végbemenő mágikus átalakulásban: ez az átmenet a valóság egyik szintjéről a másikra magával hozza a váltást egyik nyelvről a másikra” (Desrosières 1998: 70).

A valóság szintjei közötti mozgás, a váltás a közvetlenül megfigyelhető és megragadható objektumról a származtatott, absztrakt fogalmi szintre – a mágikus transzformáció – mégsem vált egy csapásra a statisztikai munka bevett elemévé. Ma már talán nehéz elképzelni, de még Quetelet átlagát is – a számítás egyszerűsége ellenére – viták övezték. És nem azért, mert az ellenlábások felkészületlenek vagy vaskalaposak lettek volna. Kétségeik inkább abból fakadtak, hogy az absztrakcióval nyer-e annyit a statisztika felhasználója, mint amennyi bizonytalanság keletkezik az információvesztés miatt. Moreau de Jonnés, aki 1833 és 1851 között vezette a Statistique Générale de la France intézményét, az átlagot valami olyan fikciónak tekintette, ami talán megfelel egyes embereknek, de soha nem fogják tudni helyettesíteni a valódi számokat. A statisztika könyvelési típusú, adminisztratív elve tagadta az aggregátumok valóságosságát (Desrosières 1998: 85). Egy holland főtisztviselő, Keverberg báró 1827-ben azért ellenezte Quetelet javaslatát, hogy becsléssel állapítsák meg Hollandia népességét, mert kétségbe vonta, hogy egy néhány helység adata alapján számított koefficiens, mint a születési ráta, kiterjeszhető lenne az egész országra. Amellett, hogy megnevezetlenül bár, de itt a mintavétel problémái merültek fel, Keverberg érvelése szerint a népességről és az azt alkotó elemekről egyetlen módon nyerhető egzakt tudás, egy valódi és teljes

cenzussal, egy regiszter létrehozásával, amely tartalmazza valamennyi lakos nevét, korát, foglalkozását (idézi: Desrosiérés 1998: 88). Keverberg felfogása maradéktalanul megfelel a kortárs statisztika feladatköréről alkotott képnek, miszerint nagy mennyiségű egzakt adattal kell támogatnia az állami döntéshozatalt. Miközben úgy tűnik, értetlenül állt Quetelet módszertani újítása előtt, a regiszterre vonatkozó felvetése maga is modern gondolat volt. Napjaink hivatalos statisztikájában a mintavétel, a matematikai módszerek alkalmazása és a regiszterek egymást kiegészítve, támogatva nélkülözhetetlen részei a módszertani apparátusnak. Ez a történet rámutat arra is, hogy a hivatalos statisztika fejlődése sem lineáris, miként általában a tudományé sem. A statisztika, a tudomány és a társadalom egymást mélyen átható változásai során mellékvágányra sodródott eszmék, megválaszoltnak tűnő kérdések újra előkerülnek, és hozzájárulhatnak új problémák megoldásához.

A valóság mibenléte nem csak a statisztika számára jelentett több hullámban visszatérő kérdésfelvetést, hanem más társadalomtudományok, így a szociológiának is. A szociológia számára annyiban egyszerűbb a helyzet, hogy nem magának a társadalomnak a létét kell igazolnia, azt adottnak veszi, érdeklődési körén kívül esik, hogy az Csuang Ce pillangójának az álma vagy sem – a mibenléte, milyensége, működési módja foglalkoztatja. „A mindennapi élet valóságát elfogadjuk mint valóságot. Egyszerű létén túl nincs szüksége további bizonyításra. Egyszerűen itt van, mint nyilvánvaló, kényszerítő tény. Tudom, hogy valóságos” (Berger és Luckmann 1998: 43). További pontosítást jelent még az a meghatározás, hogy „a »valóságot« mint a jelenségek olyan minőségét határozzuk meg, amelyek akaratunktól függetlenül léteznek (nemlétüket nem tudjuk elérni), a »tudást« pedig mint azt a bizonyosságot, hogy a jelenségek valóságosak, és meghatározható tulajdonságokkal rendelkeznek” (Berger és Luckmann 1998: 11). Az ilyen módon felfogott, tételezett társadalmi valóság elemzéséhez, megértéséhez szükség van közvetlenül hozzáférhető, megfigyelhető, rögzíthető operacionalizált alkotóelemekre. Legáltalánosabban ilyenként fordulhatunk a társadalmi tényekhez. Durkheim klasszikus megfogalmazásában „társadalmi tény minden olyan – állandósult vagy nem állandósult – cselekvésmód, amely képes kényszerítő erővel hatni az egyénre, vagy pedig ami egy adott társadalomra általánosan jellemző, és egyéni megnyilatkozásaitól független, önálló léttel rendelkezik” (Durkheim 1978: 36). A társadalmi tények ebből következően tehát cselekvésmódok, amikre a nevezetes durkheimi maxima vonatkozik: „a társadalmi tényeket dolgoknak kell tekintenünk” (Durkheim 1978: 37). Desrosiérés rámutat, hogy a definíciónak kétféle olvasata is lehetséges, az egyik szerint állítás a valóságról (a társadalmi tények dolgok), a másik szerint módszertani választás (társadalmi tényeket úgy kell kezelünk, mint ha dolgok lennének) (Desrosiérés 1998: 3). Végül úgy foglal állást, hogy „a történelem megérteti velünk, miként váltak a társadalmi tények dolgokká, és ennek megfelelően miként váltak dolgokká mindenki számára, aki statisztikai eszközöket használ” (Desrosiérés 1998: 3). A későbbiekben Durkheim rögzít még néhány, statisztikusok számára különösen megszívlelendő jellemzőt a társadalmi tényekről.

A társadalmi jelenségek mindazonáltal dolgok, s dologként is kell kezelünk őket. [...] (E)z az egyetlen adat (datum), amellyel a szociológia rendelkezik. Dolog ugyanis mindaz, ami adva van, ami megfigyelésre kínálja vagy inkább kényszeríti magát. Dologként kezelni a jelenségeket azt jelenti, hogy olyan adatot látunk bennük, amely a tudomány kiindulópontját alkotja. A társadalmi jelenségek vitathatatlanul rendelkeznek ezzel a tulajdonsággal (Durkheim 1978: 49).

Az adatforradalomban minden adat – és az adat minden. Ha megfordítjuk a fenti durkheimi logikát, arra jutunk, hogy az adat dolog, a dolgok pedig társadalmi tények. A fenti, a skolasztikától a szociológiai elméleteknek a társadalmi valóság mibenlétét rögzítő áttekintésére azért volt szükségünk, hogy megvizsgálhassuk az adatforradalom társadalmi tényeinek és a hivatalos statisztikának a viszonyát. Az internethasználók, a Google-keresések, az e-mail-üzenetek, a Twitter- és a blogbejegyzések, a Facebook-posztok száma jóval meghaladja a napi 200 milliárdot. A naponta keletkező adatmennyiség nehezen megbecsülhető, de ennél nagyságrendekkel nagyobb lehet, mivel ezek java része nem elemi, hanem többszörösen származtatott adat, hiszen például egy komment, egy e-mail vagy blogbejegyzés is adatokat vagy adatokra hivatkozó adatokat tartalmaz. Az adatfolyam mögött pedig emberi magatartások vannak, ráfordított idővel, és ebben a hálózatban információk, impulzusok, eszmék, nézetek, vélemények áramlanak, hatást és ellenhatást váltanak ki, formálva és dinamikus működésben tartva az adatforradalom társadalmát. A Thomas-elv szerint, ha az ember egy szituációt valósággal definiál, az valóságos a következményeiben.²⁶ Vagyis a társadalomkutatók dolgát az teszi izgalmassá és módfelett komplikálttá, hogy a valósággal együtt változik az egyéneknek a valóságról alkotott percepciója is, ami ugyancsak részévé válik a változó valóságnak. De elmondhatjuk azt is, hogy „a digitalizált adatok felszabadítják gondolkodásunkat a korábban kényszerűségből használt leegyszerűsítő, homogenizáló és holisztikus fogalmak használata alól” (Dessewffy és Láng 2015: 168). Napjainkban ráadásul mindez elképesztő gyorsasággal történik, eltérő mértékben érintve az individuuumok különböző (és fluktuáló) csoportjait. Ez az adatforradalom társadalmi valósága.

A működő nemzetközi statisztikai rendszer minden kétséget kizáróan globálisan is értékes, fontos információkat állít elő hagyományos eszközeivel, az adatgyűjtésekkel és az adminisztratív források felhasználásával. Ezeket minden túlzás nélkül mondhatjuk nélkülözhetetlenek és alapvetőnek. Az előző fejezetben foglaltuk össze azokat a masszív, zárt rendszerre összeálló tényezőket, amelyek megakadályozzák a hivatalos statisztikát abban, hogy aktív, hiteles és mértékadó szereplője és formálója legyen a világban zajló változások értelmezésének. Az adatforradalom korában mélyen egymásba fonódva él a régi, „analóg” és a digitális világ. Nem arról van szó, hogy a népesség egy része a digitális szakadék egyik oldalán van, a másik része pedig a túloldalon – a határok sokkal elmosódottabbak, nem lehet őket földrajzi régiók, generációk, társadalmi helyzet alapján egyértelműen kijelölni. Minden bizonynyal egy menő influenszernek, vagy magát főhivatású Instagram személyiségnek tekintő embernek is megvannak a tradicionális társadalmi kapcsolatai, a magányosan élő idős falusi asszony pedig mobiltelefonjával akaratlanul is része, forrása az adatfolyamnak. Az adatforradalom társadalmi valóságáról szóló fentebbi fejtegetéssel viszont remélhetően sikerül meggyőzően érvelni amellel, hogy az Internet of Data is a valóságunk integráns része, sőt társadalmi, gazdasági hatásaiban, a mindennapi élet szerkezetének átalakításában játszott szerepe miatt, valamint azért, mert minden jel szerint a változásoknak ez a fő iránya, kevés kockázatot vállalva a nagyobb fontosságúnak is gondolhatjuk.

A hivatalos statisztika válságának legnyilvánvalóbb jele az, hogy nincs hozzáférése az adatforradalom új valóságához, így pedig nem tudja betölteni korábban már többször

26 <https://www.oxfordreference.com/view/10.1093/oi/authority.20110803104247382>

hivatkozott küldetését. A GAFA és más technológiai cégek adattudósai permanensen elemzik a digitális adatfolyamot, és immár eljutottak oda, hogy nem csupán képesek előrejelezni egyének és csoportok viselkedését, de befolyásolni is tudják azt. Mindeközben a hivatalos statisztika – és általában az empirikus társadalomkutatás, részben kényszerűségből – bele- ragadt a saját konvencióiba, fogalomrendszerébe. A kialakult klasszifikációit osztályokról, rétegekről, társadalmi csoportokról, háztartástípusokról továbbra is jól fel lehet tölteni ada- tokkal, kifinomult matematikai statisztikai módszerekkel előállnak eredmények is, de egyre inkább elvékonyodik az a szál, ami az ilyen módon képzett csoportok és társadalmi viselke- désük, magatartásuk között fennáll. Két alapvető probléma körvonalazódik tehát a hivatalos statisztikával kapcsolatban: az adatforradalomban nagy sebességgel multiplikálódó társadal- mi tények döntő hányadáról nem vehet tudomást, amivel pedig foglalkozik a módszereivel és konvencióival lefedett körben, azoknak egy jelentős része immár nem releváns.

Logikai úton tulajdonképpen triviálisan következhetne a megoldás, hogy nemzetközi jogalkotással a technológiai cégeket arra kellene kötelezni, hogy elemzési célokra jól de- finiált feltételek mellett biztosítsanak hozzáférést a digitális társadalom (beleértve ebbe a gazdaságot is) működése során keletkező adatokhoz. Itt azonban nyilvánvalóan sokkal bo- nyolultabb kérdéssről van szó, mint a GDPR esetében, ahol az EU a tagállamok területén ér- vényes adatvédelmi szabályokat léptetett életbe, de ezzel csupán a személyes adatok kezelése során érvényesítendő szabályokat írta elő, nem a cégek birtokában levő adatokhoz kívánt harmadik félnek hozzáférést biztosítani. Az adat érték. A nagy adatgazdáknak vastos primer gazdasági érdekei fűződnek az adatmonopóliumuk fenntartásához, miközben a legkevésbé sem érdekeltek abban, hogy az adatok elemzésével a társadalomról és a gazdaságról megsze- rezhető tudáson osztozzanak kormányzati szervekkel és kutatókkal.

Beláthatatlan távlatokat nyitna meg, ha a hivatalos statisztika és a kutatók számára elem- zési célra megnyílnának az adatforradalom társadalmának működését rejtő adatbázisok. Elsőként ezek felhasználásával kellene megújítani a statisztika nyelvét, fogalomrendszerét, kidolgozni új változókat, klasszifikációkat, osztályozásokat. Az új adatforrások módszer- tani fordulatot is igényelnek, az eddigi alapvetően deduktív megközelítésről az induktív- deduktív-induktív ciklusokba szerveződő üzemmódra kellene váltani, amire a nyitottság ki- olvasható az áttekintett dokumentumokból (lásd még: Radermacher 2019: 59). A hivatalos statisztika története során a munkájához ma már elválaszthatatlanul kötődő módszertani újításokat vett át kívülről és épített be az eszköztárába. Az alkalmazott matematika megter- mékenyítő hatásán túl például a könyvviteli technikákból eredeztethető a statisztikai tábla, az államigazgatásból származik a regiszter, a piac- és közvélemény-kutatóktól pedig a survey módszer. Most pedig elérkezettnek látszik az idő, hogy a GAFA-cégek adattudósainak ered- ményeit, az algoritmusok és az AI elemzési célú felhasználását honosítsa meg. A fejezet kezdetének történeti hivatkozásaira visszautalva egy újnominalista megközelítés, egyfajta occami szemlélet szükségessége merülhet fel, amennyiben *az elemi tényekből, az adatokból kell újra konstruálni a valóságról alkotott képünket.*

Az elmúlt években a hivatalos statisztikai szolgálatoknál egyre több helyen épültek ki azok a lehetőségek, ahol a kutatók ún. biztonságos környezetben, vagyis az adatszolgáltatók azonosíthatatlanságának, felfedhetetlenségének kompromisszumoktól mentes érvényesítése mellett mikroadat-állományokon dolgozhatnak. Ez a gyakorlatban igazolja azt, hogy lehet- séges a mikroadatok felhasználása olyan módon, hogy az egyéneknek és a szervezeteknek nem sérülnek az egyéni adataikhoz fűződő jogai. A kutatószobában és más biztonságos

módon történő elemzésekhez a kutatók használhatják a statisztikai adatgyűjtések során keletkező állományokat (a KSH például elvileg minden adatát rendelkezésre bocsátja kutatószobai felhasználásra), de további lehetőségeket nyit meg, hogy ezeket össze lehet kapcsolni adminisztratív forrásokból származó állományokkal is. Túlás nélkül mondhatjuk, hogy ezzel a kutatók előtt olyan távlatok nyíltak meg és olyan perspektíva, mint amikor a biológusok először pillanthatnak mikroszkópba. A mikroadatokon és az összekapcsolt állományokon keresztül a társadalom mélyében végbemenő folyamatokról ezekkel az eszközökkel olyan összefüggések tárulnak fel, nyerne empirikus igazolást, amikről korábban legfeljebb sejtéseket lehetett megfogalmazni (például Hajdu, Kertesi és Kézdi 2019; Hajdu, Kertesi, Kézdi és Szabó-Morvai 2020). A konkrét eredményeken túl a mikroadatok kutatási célú hozzáférhetővé tétele azt is igazolja, hogy a statisztikai hivatalok rendelkeznek azokkal a készségekkel, technológiával és know-how-val, amikkel a legszigorúbb adatvédelmi kötelek mellett is transzparensten és ellenőrizhető módon kezelnek és nyitnak meg a felhasználók előtt akár érzékeny információkat tartalmazó adatbázisokat is. Ez egy olyan erősségük, amivel az adatvagyonukkal átláthatatlan gyakorlatot folytató technológiai cégek előtt járnak, akiknek a ténykedése ehhez képest asztal alatti machinációnak minősíthető.

Az adatforradalom társadalmi valóságát jelentő online világ adatfolyamához való rendezett hozzáférés tényleg egy új társadalomtudomány kezdetét jelentheti, még ha nem is feltétlenül a Pentland által megfogalmazott értelemben. Fontos az a meglátás is, hogy „a számítógépes társadalomtudomány talán legnagyobb kihívása, hogy az értelmezési kapacitás lépést tud-e tartani a számítógépes módszerek outputjainak rohamos bővülésével” (Rudas és Péli 2018: 942). Amiként a fizika tudományának behatolása a kvantumvilágba forradalmasította tudásunkat az univerzum működéséről, ugyanúgy a társadalom elemi részecskéinek, az adatoknak az üzemszerű bekapcsolása a társadalomtudományi elemzésekbe új fejezetet nyithat a társadalom működésének, változásainak megértésében (például Vosoughi, Roy és Aral 2018). És ebben eminens szerepe lehet a hivatalos statisztikának, ami tapasztalataival, tudásával, szabályozottságával, transzparenssé működésével, intézményrendszerével egyfajta interface funkciót tölthet be a magánkézből levő adatok és a felhasználók között.

Rengeteg persze a bizonytalanság azzal kapcsolatban, hogy vannak-e egyáltalán reális esélyek, sőt mi több, komoly kormányzati szándékok a magántulajdonban álló adatvagyonok közcélú hozzáférését biztosító jogszabályok megalkotására. A dolgok jelenlegi állása szerint valószínűbbek tűnhetnek olyan forgatókönyvek, hogy az elslumosodó politikában befolyásos kormányzatok fű alatt inkább kiegészítsék a GAFA-cégekkel, és igyekezzenek hasznot húzni a manipulációs technikák alkalmazásából, nem pedig a társadalomról megszerezhető új tudásnak nyitnak utat. Ez a tudás pedig egyrészt alkalmas lehet az aktuális társadalmi problémák kezelését célzó közpolitikák megalapozására, ugyanakkor csúf vitákkal jár, elemzési eredményeket lobogtató mindenféle „gyanús” alakokkal, mely viták bizony magukban rejtik a politikai bukás kockázatát. Akkor már inkább valami jól csomagolt és kellemes felhasználói élményt nyújtó manipuláció?

A tét nagy. És nyilván nem csupán arról van szó, hogy néhány tízezer statisztikus munkájára szükség lesz-e a jövőben is. Több évszázados múltja azt igazolja, hogy a hivatalos statisztika helyzete érzékeny indikátora annak, miként viszonyulunk a társadalom működéséhez, a folyamatosan keletkező problémák kezeléséhez, a köz ügyeihez. Lehet-e tényekre alapozott célokat megfogalmazni egy társadalom vagy annak egyes csoportjai számára, és nyilvánosan hozzáférhető-e az információk ezek haladásáról, és lehet-e erről véleményeket közzétenni,

tárgyszerű vitákat folytatni? Valami ilyesmi, amit demokráciának és demokratikus nyilvánosságnak hívnak, és ennek integráns szereplője a hivatalos statisztika. Vagy olyan kor felé haladunk, amihez képest az irodalom és a filmművészet nagy antiutópiai, disztópikus víziói szélsőliberális elhajlásnak minősülnek?

A hivatalos statisztikának az adatforradalom megváltozott társadalmi viszonyai miatt előállt válságáról szóló, helyenként talán túl körülményes gondolatainkat azzal foglalhatjuk össze, hogy a megoldás kulcsa az adatforradalom társadalmi valóságát hordozó, magántulajdonban álló adatvagyonhoz való elemzési célú hozzáférés. Az elért eredményeit megőrizve, a tevékenységét folytatva a már hatókörébe vont társadalmi, gazdasági, környezeti folyamatok megfigyelésében, működését mielőbb ki kell terjeszteni az adatforradalom új valóságára is, azt mintegy művelés alá vonni. Ha ez megvalósul, olyan tudáshoz és muníciónhoz jut a statisztika és annak minden felhasználója, amivel megerősítheti, visszanyerheti pozícióit az adatforradalom világában, és ismét fajsúlyos szereplője lehet a változó valóság hiteles értelmezésének. Egyúttal esélye lenne az adekvát válaszra a fontosságát és szavahihetőségét a felügyelő kapitalizmus, a post-truth valóság, az identitáspolitikák és a populista politika oldaláról érő kihívásokra.

William Davies a következő figyelmeztetéssel zárta többször idézett írását (Davies 2017): „a harc, amit hosszú távon meg kell vívni, nem az elit vezérelte tényalapú politika és az érzelmek populista politikája között zajlik, hanem azok között, akik még elkötelezettek a közös tudás és a nyilvános érvelés mellett, és akik profitot termelnek ezeknek a dolgoknak a folyamatos bomlasztásából”.

Hivatkozott irodalom

- Balkin, Jack M. (2014): Old-School/New-School Speech Regulation. *Yale Law School Faculty Scholarship Series Paper* (127): 2296–2342. Interneten: http://digitalcommons.law.yale.edu/fss_papers/4877.
- Balkin, Jack M. (2017): Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation. *Yale Law School, Public Law Research Paper* (615). DOI: <https://doi.org/10.2139/ssrn.3038939>
- Berger, Peter L. és Thomas Luckmann (1998): *A valóság társadalmi felépítése*. Budapest: Józsefvég.
- Data Revolution Group (2014): *A World That Counts*. (Kutatási jelentés az ENSZ számára.) Interneten: <https://www.undatarevolution.org/wp-content/uploads/2014/12/A-World-That-Counts2.pdf>.
- Davies, William (2017): How Statistics Lost Their Power – and Why We Should Fear what Comes Next. *The Guardian* (január 19.). Interneten: <https://www.theguardian.com/politics/2017/jan/19/crisis-of-statistics-big-data-democracy>.
- Dean, Jodi (2018): Kommunikatív kapitalizmus és osztályharc. *Fordulat* (23): 33–52.
- Desrosières, Alain (1998): *The Politics of Large Numbers: A History of Statistical Reasoning*. Cambridge, MA: Harvard University Press.
- Dessewffy Tibor és Láng László (2015): Big Data és a társadalomtudományok véletlen találkozása a mütőasztalon. *Replika* (92–93): 157–170.
- Durkheim, Émile (1978): *A társadalmi tények magyarázatához*. Budapest: Közgazdasági és Jogi Könyvkiadó.
- ESS Committee (2017): *ESS priorities beyond 2020. Position paper of the European Statistical System Committee adopted on 16 November 2017*. Interneten: <https://ec.europa.eu/eurostat/documents/7330775/8463599/ESS+priorities+beyond+2020+final.pdf/42ea415b-0b40-418b-8c76-25a8e219365c>.
- ESS VIN (2019): *Proposed Cross Cutting Innovation Actions for the Multiannual Action Plan*. (2019. október 17.)
- Eurostat (2014): *Eurostat ESS Vision 2020*. Interneten: <https://ec.europa.eu/eurostat/documents/7330775/7339647/ESS+vision+2020+brochure/4baffca-9469-4372-b1ea-40784ca1db62>.

- Eurostat Outlook Report (2018): *Power from Statistics: Data, Information and Knowledge*. Luxembourg: Eurostat. Publications Office of the European Union. Interneten: <https://ec.europa.eu/eurostat/documents/7870049/9087487/KS-FT-18-005-EN-N.pdf/8c143a88-a67c-46ca-a24d-c34046fc4a49>.
- Eyraud, Corine (2018): Stakeholder Involvement in the Statistical Value Chain: Bridging the Gap between Citizens and Official Statistics. In *Eurostat Outlook Report, 2018 Edition: Power from Statistics: Data, Information and Knowledge*. Luxembourg: Eurostat. Publications Office of the European Union, 103–106.
- Fukuyama, Francis (1994): *A történelem vége és az utolsó ember*. Budapest: Európa.
- Hajdu Tamás, Kertesi Gábor, Kézdi Gábor és Szabó-Morvai Ágnes (2020): The Effects of Expanding a Neonatal Intensive Care System on Infant Mortality and Long-Term Health Impairments. *CERS-IE Working Papers 2020/20 May 2020*. Interneten: <https://www.mtakti.hu/wp-content/uploads/2020/05/CERSIEWP202020.pdf>.
- Hajdu Tamás, Kertesi Gábor és Kézdi Gábor (2019): Health Differences at Birth between Roma and Non-Roma Children in Hungary. Long-Run Trends and Decompositions. *Population and Development Review* 45(3): 631–657. DOI: <https://doi.org/10.1111/padr.12276>
- Harari, Yuval Noah (2017): *Homo Deus*. Budapest: Animus Kiadó.
- Kuonen, Diego (2018): *Production Processes Official Statistics & Data Innovation Processes Augmented by Trusted Smart Statistics: Friends or Foes?* (Konferencia-előadás: Big Data for European Statistics, Szófia, Bulgária, 2018. május 14–15.) Interneten: https://webgate.ec.europa.eu/fpfis/mwikis/essnetbigdata/images/c7/BDES_2018_Diego_Kuonen_Keynote_speech.pdf.
- Martinez, Antonio Garcia (2019): No, Data is Not the New Oil. *Wired* 2019. (február 26.). Interneten: <https://www.wired.com/story/no-data-is-not-the-new-oil/>.
- McGinnis, John O. (2012): A Politics of Knowledge. *National Affairs* (Winter 2012): 58–74. Interneten: <https://www.nationalaffairs.com/publications/detail/a-politics-of-knowledge>.
- Panofsky, Erwin (1986): *Gótikus építészet és skolasztikus gondolkodás*. Budapest: Corvina.
- Pentland, Alex (2014): The Death of Individuality. *New Scientist* (2963): 30–31.
- Pentland, Alex (2018): Data for a New Enlightenment? A Transcendent Decade. *Madrid BBVA* 85–105. Interneten: <https://www.bbvaopenmind.com/wp-content/uploads/2019/02/BBVA-OpenMind-Alex-Pentland-Data-for-a-New-Enlightenment.pdf>.
- PwC (2018): *Global Top 100 Companies by Market Capitalisation*. Interneten: <https://www.pwc.com/hu/hu/kiadvanyok/assets/pdf/global-top-100-companies-2018.pdf>.
- PwC (2019): *Global Top 100 Companies by Market Capitalisation*. Interneten: <https://www.pwc.com/gx/en/audit-services/publications/assets/global-top-100-companies-2019.pdf>
- Radermacher, Walter (2018a): Official Statistics in the Era of Big Data Opportunities and Threats. *International Journal of Data Science and Analytics* 6(3): 225–223. DOI: <https://doi.org/10.1007/s41060-018-0124-z>
- Radermacher, Walter (2018b): The Future Role of Official Statistics. In *Eurostat Outlook Report, 2018 Edition: Power from Statistics: Data, Information and Knowledge*. Luxembourg: Eurostat. Publications Office of the European Union, 107–113.
- Radermacher, Walter (2019): *Official Statistics 4.0. Facts for People in the 21. Century*. (Doktori értekezés.) Róma: University of Rome “La Sapienza” Department of Statistics. Interneten: https://iris.uniroma1.it/retrieve/handle/11573/1237923/1032618/Tesi_dottorato_Radermacher.pdf.
- Reimsbach-Kounatze, Christian (2015): The Proliferation of “Big Data” and Implications for Official Statistics and Statistical Agencies: A Preliminary Analysis. *OECD Digital Economy Papers* (245). DOI: <http://dx.doi.org/10.1787/5js7t9wqzvg8-en>.
- Rosling Hans, Anna Rosling Rönnlund és Ola Rosling (2018): *Tények*. Budapest: Libri.
- Rudas Tamás és Péli Gábor (2018): Számítógépes társadalomtudomány. Bevezető. *Magyar Tudomány* (179): 939–942. DOI: <https://doi.org/10.1556/2065.179.2018.7.1>
- Sassen, Saskia (2001): *The Global City*. Princeton: Princeton University Press.
- Stiglitz, Joseph, Amartya Sen és Jean-Paul Fitoussi (2009): *Report by the Commission on the Measurement of Economic Performance and Social Progress*. Interneten: <https://ec.europa.eu/eurostat/documents/118025/118123/Fitoussi+Commission+report>.
- The Economist (2017): The World's Most Valuable Resource is No Longer Oil, but Data. *The Economist* (május 6.). Interneten: <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>.
- Vosoughi, Soroush, Roy Deb és Aral Sinal (2018.): The Spread of True and False News Online. *Science* (6380): 1146–1151. DOI: <https://doi.org/10.1126/science.aap9559>

- World Bank (2016): *World Development Report 2016. Digital Dividends*. Washington, DC: World Bank. DOI: <http://doi.org/10.1596/978-1-4648-0671-1>
- World Bank (é. n.): *World Development Indicators*. Interneten: <https://databank.worldbank.org/reports.aspx?source=world-development-indicators>.
- Zuboff, Shoshana (2016): The Secrets of Surveillance Capitalism. *Frankfurter Allgemeine Zeitung* (2016. március 5.). Interneten: http://www.faz.net/aktuell/feuilleton/debatten/the-digital-debate/shoshana-zuboff-secrets-of-surveillance-capitalism-14103616.html?printPagedArticle=true#pageIndex_2.
- Zuboff, Shoshana (2019): Surveillance Capitalism: An Interview with Shoshana Zuboff. *Surveillance & Society* 17(1–2): 257–266. DOI: <https://doi.org/10.24908/ss.v17i1/2.13238>

Németh Zsolt

Szociológus, statisztikus, Központi Statisztikai Hivatal