

összehasonlításban a mezőny leszakadt részén van. A szerzők Ukrajna pozícióit az előbbi rangsorok alapján kialakított négy fejlettségi kategóriában elemzik, különös tekintettel az innovációk lehetőségeire, a tudástökére alapozva.

A tudás képessé teszi a gazdaság szereplőit, hogy eredményesen növeljék az innováció szintjét a vállalkozásban, amely pedig a gazdaság minden szintjén befolyásolja az innovációs szintet. Nyilvánvaló tehát, hogy a tudás növekvő használata biztosítani tudja a szervezetek növekvő hatékonyságát és jövedelmezőségét.

A tanulmány következtetése, hogy az innováció társadalmi háttere lényeges, ez biztosítja a nemzetgazdaság modernizálását. Az ismeretek, a szellemi tőke és az innováció révén egyre több high-tech termék hozható létre magas hozzáadott értékkel. Igazolt a kölcsönhatás az innovációs termékek értékesítése és a hozzáadott érték között, ezzel a szervezetek profitja is növelhető. Ahol elavultak a termelési technológiák, ott a termelők fokozatos piaci pozícióvesztéssel számolhatnak.

Az ukrán innovációk lassulása összefügg azzal, hogy 2014–2015-ben csökkent a GDP (2014: 177,8 Mrd USD, 2015: 130,7 Mrd USD), a politikai stabilitás nemzetközi rangsorában pedig a korábbi 77. helyről a 110. helyre esett vissza az ország. A beruházás kedvező feltételeinek rangsorában Ukrajna korábban a 127. volt, jelenleg csak a 136.

**TÁRGYSZAVAK:** szellemi tőke, tudásgazdaság, tudás, innovatív fejlesztési potenciál, innovatív fejlődési modell, versenyelőny, Ukrajna

CSEH GABRIELLA

## Általános kérdések

91/2016

### A fenntartható fejlődés mutatói: a 2030-as Agenda utolsó hiányzó eleme

Heidi Worley: Sustainable Development Indicators: the last missing piece of the 2030 Agenda  
*Population Reference Bureau*, 2016. március  
URL: <http://www.prb.org/Publications/Articles/2016/sustainable-development-indicators.aspx>

A Fenntartható Fejlődési Célok (FFC) múlt évi elfogadása fontos mérföldkőnek számít az ENSZ tagállamai számára az emberek életének javítását és egy

egészséges bolygó megőrzését célzó törekvések lendületbe hozása szempontjából. Stefan Schweinfest, az ENSZ Statisztikai Főosztályának vezetője szerint a 2030-as fenntartható fejlődési agenda felépítésének „utolsó hiányzó eleme” a fejlődési mutatók olyan átfogó keretének elfogadása, amely az említett célok 2030-ig történő eléréséhez nyújt iránymutatást az országok számára.

A 17 cél és 169 rész cél programját 2015 szeptemberében fogadták el az ENSZ-tagállamok állam- és kormányfőinek csúcstalálkozóján, hatályuk pedig 2016. január 1-től 2030 végéig terjed. Az új célok a részben megvalósult, korábbi nyolc Millenniumi Fejlődési Cél és az ezeket részletező 21 rész cél helyébe léptek, amelyek fókuszában egyebek között a szegénység leküzdése és az egészségügy állt. Az FFC ezeken túllépve a fenntartható fejlődés gazdasági és környezeti vonatkozásaira is kiterjed.

Az ENSZ Statisztikai Bizottsága, valamint a fenntartható fejlődési célokra vonatkozó mutatók intézményközi és szakértői csoportja kidolgozta az indikátorok keretrendszerét, amelynek segítségével mérhető lesz a globális célok irányába történő haladás mértéke, továbbá mind a megvalósulások, mind a hiányosságok értékelése. Az ENSZ Statisztikai Bizottságának márciusban tartott 47. ülészakán a szakértők áttekintik a javasolt keretrendszert, és a tervek szerint júniusban felterjesztik elfogadásra az ENSZ Gazdasági és Szociális Tanácsához és a Közgyűléshez. A munkák további részét országos szinten kell végezni, ahol a kormányoknak a sajátos körülményeiknek és adatgyűjtési lehetőségeiknek megfelelő indikátorokat kell kidolgozniuk. E mutatók a célok elérése és a fejlődés monitorozása szempontjából segíthetnek az adott ország prioritásainak megtalálásában, ám csak annyira, amennyire az alapul szolgáló adatgyűjtések minősége engedi. Számos szakértő szerint ugyanis sok ország nincs erre felkészülve. Az egészségügyet illetően például a WHO erőfeszítéseket tett az országok egészségügyi statisztikai adatgyűjtéseinek fejlesztésére, de ez csak lassan haladt, mert a kormányok vagy a donor szervezetek a népmozgalmi statisztikát és az ilyen adatgyűjtéseket nem tekintették prioritásnak, illetve a fejlesztési infrastruktúra részének. Ez most már változóban van, de kérdéses, hogy elég gyorsan-e az FFC 2030 szempontjából szükséges adatgyűjtések megvalósításához.

A mutatórendszer kidolgozása számos tanácskozás, konferencia és ötletbörze, azaz hosszabb konzultációs folyamat eredménye. Az előzetes, decemberi változat nagyobb számú (229) indikátorra tett javasla-

tot, amelyek többségéről azóta általános megegyezés született, a kisebb részét azonban tovább kellett vitatni és finomítani. Az ENSZ Statisztikai Bizottsága elé már a mutatók végleges listája került. A rendszer középpontjában a globális mutatók állnak, de az országok közreműködésével ezeket regionális, országos és országon belüli mutatókkal kell kiegészíteni. Tematikus mutatók kidolgozására is sor kerül.

Az említett szakértői munkacsoport útmutatást és támogatást nyújt az adatok részletezéséhez, lehetőség szerint nem, kor, lakóhely (város vagy vidék), valamint más releváns jellemzők szerint, hogy a mutatók elkülönítsék a sérülékeny népességcsoportokat és az FFC által meghatározott egyéb tényezőket. A jelentéseket évente kell elkészíteni, az adatgyűjtések pedig a nemzeti statisztikai hivatalok feladatát képezik. 2016 őszén az ENSZ világfórumot fog tartani a fenntartható fejlődés adatairól.

Az idén március 30. és április 1. között tartott mexikóvárosi utótálálkozóra a szakértői csoport a keretrendszer megvalósításának első lépéseként három szintet határozott meg az indikátorokkal kapcsolatban:

- Kidolgozott módszertan létezése és adatok széles körű rendelkezésre állása.
- A módszertan megvan, de az adatok hozzáférése problematikus.
- Még nem sikerült nemzetközileg is elfogadható módszertant kidolgozni.

Az ENSZ főtitkárának 2015 utáni fejlesztési tervekkel foglalkozó főtanácsadója elsősorban azért aggódik, mert jó alapadatok nélkül szinte lehetetlen a fejlődés monitorozása. Szerinte mintegy száz országban a legalapvetőbb adatok beszerzése is komoly nehézséget jelent a megfelelő népmozgalmi statisztikai rendszer híján. A Világbank becslése szerint például világméretben közel 230 millió öt éven aluli gyermek nem szerepel a nyilvántartásokban. Sőt, a Brookings Institution elemzése szerint sok ország nem tudja teljesíteni a nemzetközi normáknak megfelelő, tízévenkénti népszámlálásokat sem.

Indiában és más országokban viszont egyéb kihívásokkal kell szembenézni. Van adatgyűjtési rendszer, de problémát jelent a népesség nagysága. Egyes szakértők úgy vélik, hogy az FFC nagyrészt az indiai siker függvénye lesz, ahol a világnépességének csaknem 18%-a él. India már kifejezte kétségeit a mutatók monitorozásának képességével kapcsolatban, ugyanis az országnak már azért is küszködnie kellett, hogy az ötéves tervezési ciklusokhoz tartozó 25 alapvető mutatót monitorozni tudja. Egy

1,3 milliárd lakost számláló országban az adatgyűjtés hatalmas terhet jelent.

Vegyük például a nem fertőző betegségeket, összességükben a vezető halálokokat, amelyek egyúttal a legnagyobb terheket jelentik Indiában. Az FFC 3. pontja az egészségügy, a 3.4 jelű rész cél pedig: 2030-ig a megelőzés és kezelések révén harmadával kell csökkenteni a nem fertőző betegségek miatti korai halálozást, valamint javítani kell a mentális egészséget és jóllétet. E cél eléréséhez a javasolt mutatók:

- a szív- és érrendszer megbetegedéséből, rákbetegségből, cukorbetegségből vagy a krónikus légzőszervi betegségből eredő, 30 és 70 éves kor között bekövetkező halálozások valószínűsége;
- az öngyilkossági halálozási arány;
- a káros alkoholfogyasztás az ország szokásai alapján, egy naptári évben, tiszta alkohol liter/főre (15 évesek és idősebbek) átszámítva.

Indiának nincsenek megfelelő forrásai az ilyen mutatók mérésének monitorozásához. Nincsenek átfogó információs rendszerei a megbetegedési és halálozási adatok minden területről történő gyűjtéséhez, így csak modellszámításokat végeznek a nem fertőző betegségekkel összefüggő terhek megállapítására. India országos egészségügyi irányítási információs rendszere a kormányzati eszközökre korlátozódik. A nem fertőző betegségek prevalencióját és kezelését célzó monitorozás programjait nyolc év elteltével is alig tudták elkezdni. Az egyes szövetségi államok kezdeményezései nem elegendőek az országos szintű erőfeszítések sikeréhez. Nincsenek ismert információk a kábítószer- és alkoholfogyasztással foglalkozó intézményekről sem. Egyszóval a fejlődés monitorozása valóban jelentős kihívást fog jelenteni India számára. Hatalmas adatrendszereket kell létrehozni, és a különféle ágazatok közötti együttműködés is növekvő emberi és pénzügyi erőforrásokat fog igényelni.

Az ENSZ Statisztikai Bizottságának ülésén mindezeket a kihívásokat meg kell vitatni, ideértve a statisztikai rendszer fejlesztésének finanszírozási problémáit, valamint az FFC megvalósításánál hasznosítható lehetőségeket, amelyeket az adatok forradalma nyújt. Napjainkban különféle szervezetek és programok foglalkoznak a statisztikai kapacitásfejlesztéssel és technikai segítségnyújtással, mint például a PARIS21<sup>1</sup> partnerprogram, amelynek célja a szegénység csökkentése és a kormányzati munka

<sup>1</sup> Partnership in Statistics for Development in the 21st Century

javítása a fejlődő országokban a statisztika és a megbízható adatok beépítése révén a döntéshozatali folyamatokba. Ezen a területen azonban még rengeteg tennivaló van.

**TÁRGYSZAVAK:** fenntartható fejlődés, mutatók, statisztikai kapacitásfejlesztés, fejlődő országok, ENSZ Statisztikai Bizottsága, India

RÓZSA GÁBOR

92/2016

## A fenntartható fejlődés hálózatként kialakított célrendszere és az integrálás esélyei

David Le Blanc: Towards integration at last? The sustainable development goals as a network of targets

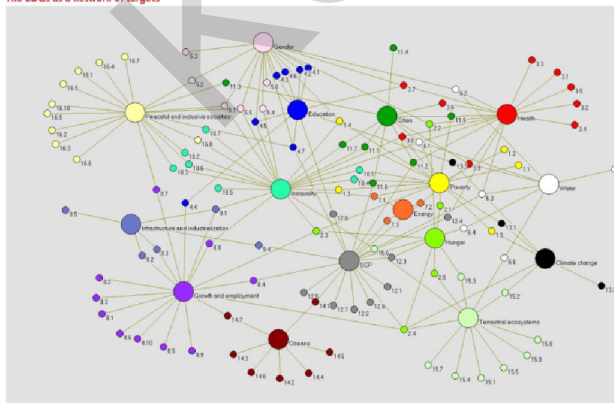
DESA Working Paper, 2015. 141. sz. 1–17. p.

URL: [http://www.un.org/esa/desa/papers/2015/wp141\\_2015.pdf](http://www.un.org/esa/desa/papers/2015/wp141_2015.pdf)

Az Egyesült Nemzetek Szervezetének (ENSZ) a 2015 és 2030 közötti időszakra összeállított SDG (Sustainable Development Goals) célrendszere a fenntartható fejlődésre vonatkozik. A 2014-ben elfogadott célrendszer az eddigi MDG (Millennium Development Goals) ezredfordulós célokat követte.

Az SDG hierarchiája összetett hálózatként kezelhető, a részcélok feltételezik és erősítik egymást a fenntartható fejlődés egymáshoz kapcsolódó rész céljának elérésében. A szakpolitikák ennek alapján nemzetközi és nemzeti szinten egyaránt összehangolhatók, hiszen például az energiagazdálkodás (az ábra szerint narancssárga, 7.) szervesen kapcsolódik több kiemelt célrendszerhez.

Figure 1  
The SDGs as a network of targets



1. ábra: Fenntartható fejlődési célok (SDG-k) mint célok hálózata

A vázolt komplex célrendszer a hálózatelemzés technikáival elemezhető. Egy éppen elemzett részrendszer szempontjai iránymutatók más rész célok

bemutatásában. Kiemelhetők azok az elemek, amelyek erősebben kapcsolódnak a vizsgált „saját” célokhoz.

A hálózatban sok olyan rész cél is van, amely gyengén hat (ha egyáltalán) az éppen vizsgált részrendszerre. Az ábra alján az „óceán” (barna, 14.) célrendszere például kellően „távol” van például az egészség megőrzésétől és nincs releváns kapcsolata például a szegénység leküzdésével.

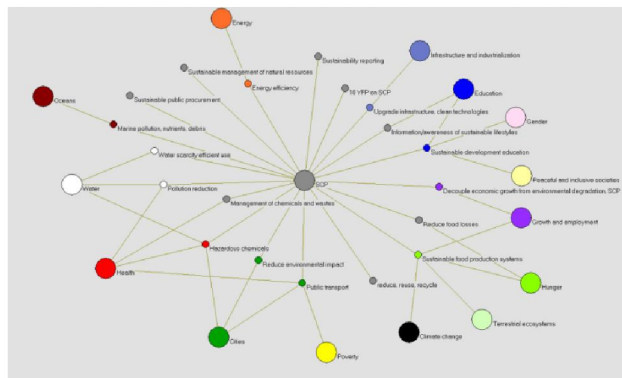
Az SDG teljes hálózatának része a békés, befoгодó társadalom (krémszín, 16.) célrendszere. Ez láthatóan kapcsolódik például a nemek szerinti (rózsaszín, 5.) célrendszerhez.

Szerteágazó kapcsolatok jellemzik az egészségmegőrzés (vörös, 3.), a vízgazdálkodás (fehér, 6.) célrendszerét, amelyek igazolhatóan összefüggnek egymással, továbbá például a szegénység leküzdése (sárga, 1.), az éhezés (zöld, 2.), az egyenlőtlenség (türkiz, 10.), illetve a városok (sötétzöld, 11) egyes rész céljaival.

A hálózat része az oktatás (kék, 4.) és az ipari háttér, infrastruktúra (szürkés-kék, 9.) célrendszere, amelyek a fenntartható fejlődés sok területének az alapfeltételei. A gazdasági folyamatok középpontjában a fenntartható fogyasztás és felhasználás (SCF, szürke, 12.), valamint a növekedés és a foglalkoztatás (lila, 8.) céljai vannak. E részrendszereknek is szoros kapcsolata például az energiagazdálkodás (narancssárga, 7.), a térségek ökoszisztémája (világoszöld, 2.) és a klímaváltozás mérséklésének (fekete, 13.) céljaival.

Az előző MDG célrendszer tételei az ezredforduló idején kevésbé voltak komplexek. Az SDG itt vázolt céljainak komplex hálózatát célszerű több lépésben elemezni és egymástól elhatárolni a valóban lényeges kapcsolatokat a fenntartható fejlődés gyakorlati feladatainak összeállításához, előkészítéséhez.

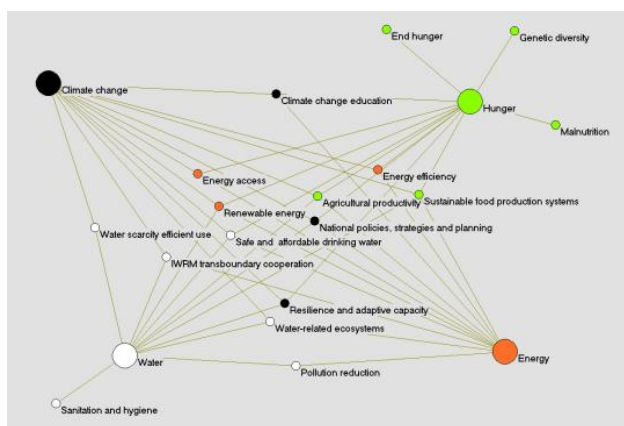
Kiemelve például a fenntartható fogyasztás és termelés (sustainable consumption and production – SCP) céljait, azok összefüggnek más célrendszerrel, a következő hálózatrész szerint.



2. ábra: Links among the goal 12 (SCP) and other goals

Az energiagazdálkodás és a vízgazdálkodás céljai kétségtelenül összefüggnek az anyagi javak fenntartható fogyasztásával, termelésével (SCF), azonban sok egyéb tennivaló is meghatározható a következő, hálózatnak megfelelő szoros kapcsolatokkal.

Az éhség (zöld) céljai például szervesen kapcsolódnak a víz (fehér) és klímaváltozás (fekete) céljaihoz. Az utóbbi rész céljai a vízgazdálkodás és energiagazdálkodás fenntarthatóságával összefüggésben elemezhetők.



3. ábra: Links among targets put forward by the ICSU-ISSC scientific review of SDGs

A részcélok között említendő (fehér jelöléssel) például a vízbe jutó szennyezőanyagok csökkentése, a vizes ökoszisztéma védelme, a szűkös vízkészlet hatékony felhasználása, a biztonságos ivóvízszolgáltatás. Az SDG azt is jelzi, hogy a vízgazdálkodás, valamint a klímaváltozás mérséklésének célrendszere nemzetközi együttműködést feltételez. Az ábra kiemeli azokat a pontokat, ahol a célok az országhatárokon átnyúló stratégiai tervekkel érhetőek el, például a folyók vízgyűjtő területén vagy a légkör terheléseiben.

A vázolt hálózatok – a sokrétű társadalmi, gazdasági kapcsolódás mellett – biofizikai célokat is megjelölnek az élővilág sokrétűségének megőrzése érdekében. A példaként bemutatott összefüggések alapján lényeges feladat elérni, hogy a monitorozás ne csupán az SDG gazdasági és környezeti főcéljaihoz rendelt részcélokra vonatkozzon, hanem „kiterjesztve” olyan kapcsolódó társadalmi célokra is, amelyek a fenntartható fejlődéshez (például szegénység elleni küzdelemhez, illetve az éhezéshez) tartoznak.

Ehhez a többféle kapcsolódó szakpolitika összehangolása, közös gondolkodás és cselekvés szükséges világszerte. Nem hagyható figyelmen kívül például a fenntartható fogyasztás és termelés (SCP) stratégiáiban a többi céllal fennálló szerves kapcsolat.

A víz és energiagazdálkodás, vagy a klímaváltozás mérséklése a hálózat többi célrendszerével összehangolt „saját” célokkal érhet el távlatilag eredményt a világ országaiban, az érintett termelőágazatokban, illetve az élővilág és a társadalmi csoportok fenntartható életfeltételeiben.

A mérés és megfigyelés tervezése is része az SDG megvalósításának, a javasolt mennyiségi és minőségi ismérvek összetett módszertani feladatokat generálnak. Hiányoznak, vagy a célnak nem megfelelőek a most elérhető információforrások. Jogos igény, hogy áttekinthető, konzisztens és közérthető adatközlések segítsék az előrehaladás méréseit és elemzéseit, például modellezéssel. Kialakítható lesz a mérés súlyrendszere, ezzel a bonyolult kapcsolatrendszer globális és nemzeti szinten történő kezelése a valóban lényeges részcélokra és kapcsolódásaikra összpontosíthat. Elfogadhatók bizonyos egyszerűsítő feltételezések is. A SMART angol betűszó (*specific, measurable, attainable, relevant, time-based*) szerint a célok monitorozásának eljárása legyen *specifikált, mérhető, megvalósítható, releváns és időadatokkal megalapozott*.

**TÁRGYSZAVAK:** a fenntartható fejlődés célrendszere, hálózatelemzés, kapcsolatrendszer, ENSZ

NÁDUDVARI ZOLTÁN

### 93/2016

#### A hivatalos statisztikához illeszthető hálózati információforrások esélyei, fejlesztései

Nicholas Robin – Thilo Klein – Johannes Jütting: Public-private partnerships for statistics: lessons learned, future steps. A focus on the use of non-official data sources for national statistics and public policy

OECD Development Co-operation Working Papers, 27. sz. Paris, 2016. február, OECD Publishing, 1–39. p.

URL: [http://www.oecd-ilibrary.org/development/public-private-partnerships-for-statistics-lessons-learned-future-steps\\_5jm3nqp1g8wf-en?crawler=true](http://www.oecd-ilibrary.org/development/public-private-partnerships-for-statistics-lessons-learned-future-steps_5jm3nqp1g8wf-en?crawler=true)

DOI: 10.1787/5jm3nqp1g8wf-en

A nemzeti és a nemzetközi statisztikai tájékoztatás mind hatékonyabban használhatja a világhálóon megosztott, tömegesen elérhető (Big Data) infor-

mációkat.<sup>1</sup> Ez a módszertani kihívás informatikai technológiákhoz kapcsolódik. A magánszféra elterjedten alkalmaz mobil kommunikációs eszközöket, intelligens érzékelőket, és a keletkező, a hálózaton nagy tömegben elérhető információk a hivatalos statisztika adatforrásaként is alkalmazhatók. A helyet is jellemző adatok a korábbinál részletesebb, kiegészítő, ellenőrző térinformációt szolgáltathatnak a területi statisztikákhoz is. A világháló közösségi oldalai nagy mintasokasággal sokféle gazdasági, társadalmi, környezeti és egyéb lényeges folyamat információit tartalmazzák, amelyek kikérdezések nélkül is elérhetők, bár többnyire nem strukturáltan („ömlesztve”, „adatáradatként”).

A magánszférára jellemző és a világhálón megosztott információk időszerűvé teszik a statisztika új koncepciójának kialakítását. A megközelítés egyik kulcsa a közösségi és a magánszféra partnersége (*public-private partnership – PPP*) a hivatalos statisztika digitalizált adatainak elérésében, a megfigyelt folyamatok mutatósorozatainak gyorsabb és szélesebb információs háttérrel végezhető összeállításában és elemzésében. Ez világméretű jelenség, stratégiát kell kidolgozni többek között

- a PPP technikai feltételeinek megteremtésére,
- a bizalmas adatok kezelésével járó felmért kockázatok kivédésére,
- a magánszféra szolgáltatásainak ösztönzésére,
- a tárolt digitális információk fenntartható elérésére.

A tanulmány áttekinti annak lehetőségét, hogy a hivatalos statisztika minőségét megőrizték az új megközelítések mellett is.<sup>2</sup>

A közszféra és magánszféra partnerségének a hivatalos statisztikában való megjelenésére vonatkozóan több megoldási forma alakítható ki:

- az adatszolgáltató saját maga állítja össze a statisztika követelményeinek megfelelő információkat; ennek esettanulmánya a mobilkommunikáció hálózatát üzemeltető cég rendszerének fel-

dolgozott adattömege a hívások részleteiről (*in-house analysis of call detail records – CDRs*),

- a feldolgozásra kiválasztott adatállományt a végző felhasználó elérheti, kezelheti, és elemezheti a nem érzékeny információkat,
- felkészült harmadik fél is közreműködik a keletkezett információk kezelésében a partnerek megbízásából, elemzést végezhet az adatállományok alapján; ennek esettanulmánya a turisztikai, utazási statisztika<sup>3</sup> mutatósorozatainak összeállítása (*travel statistics by trusted third party*),
- a hivatalos statisztikai intézmény vállalkozó beszállítást veszi igénybe meghatározott funkciók hatékonyabb, illetve nagy beruházással járó ellátására.

A kötet az ENSZ és az Eurostat hivatkozott forrásmunkáihoz kapcsolódó esettanulmányokat mutat be. A tömeges internetes információk megosztására és statisztikai célú átvételére alapozott négy példa közül három a mobil telefonok, egy a közösségi oldalak adatbázisára alapozott feldolgozási és elemzési eseteket tartalmazza. A közölt gyakorlati példák tanulságosak az internetes információk megosztását, statisztikai célú átvételét illetően.

A mobiltelefonok tömeges hívásinformációi világszerte felhasználhatók a hivatalos statisztikák adatforrásaként. A magánszféra részei a szolgáltatást végző szolgáltató vállalatok és a telefonáló személyek. A szerzők áttekintik a nemzeti statisztikai intézménnyel kialakítható partnerség fontosabb akadályait.

A magánvállalkozás érdekelt lehet a statisztika és annak inputjai<sup>4</sup> szolgáltatásában, a létrehozott (nyers) adatok birtokában. Megállapodhatnak a partnerek, hogy a vállalkozás aktív vagy passzív módon, önkéntesen végzi a statisztikai adatgyűjtéshez kapcsolódó feladatait. A közreműködés további elterjedt módja, hogy a saját üzleti (például reklám, termékajánlatot tartalmazó) adatgyűjtésének – egyeztetett módszertanú – eredményei statisztikai célra is felhasználhatók. A magánvállalko-

<sup>1</sup> A KSH és az MTA Társadalomtudományi Kutatóközpont (TK) 2016. május 27-i munkaiülésének *Adatok ára Data – Big Data* című tudósítása a függelékben.

<sup>2</sup> A szerzők forrása: Steve Landefeld: Uses of Big Data for official statistics. Privacy, incentives, statistical challenges, and other issues. *Discussion Paper*. United Nations Global Working Group on Big Data for Official Statistics. Beijing, 2014. október 31. 1–20. p. Elérhető: [http://unstats.un.org/unsd/trade/events/2014/beijing/default.asp/Uses of Big Data for official statistics - Steve Landefeld](http://unstats.un.org/unsd/trade/events/2014/beijing/default.asp/Uses%20of%20Big%20Data%20for%20official%20statistics%20-%20Steve%20Landefeld)

<sup>3</sup> A szerzők forrása: *A feasibility study on the use of mobile positioning data for tourism statistics* című Eurostat (Unit G-3 ‘Short-term business statistics and tourism’) kiadvány-sorozat. Elérhető: <http://ec.europa.eu/eurostat/documents/747990/6225717/MP-Consolidated-report.pdf/530307ec-0684-4052-87dd-0c02b0b63b73>

<sup>4</sup> Az „adatok megosztása” (*data-sharing*) vonatkozhat a „nyers” adattömegekre (*raw data*), valamint az egyeztetett szerkezetben összeállított mutatókat tartalmazó statisztikára (*exchange of statistics*).

zás közreműködhet az interneten elérhető nyers információinak egyeztetett feldolgozásában, kiadvány összeállításában és terjesztésében is.

A szerzők az ilyen partnerség kedvező lehetőségeként kiemelik a költséghatékonyságot, a rövidebb elérési időt, a mélyebb bontásokat. A turizmus statisztikája például a felkeresett célterület (*destinations of travellers*) részletesebb adatsorait veheti át mobilhívások jegyzéke alapján, ami a hagyományos adatgyűjtésekkel nagy terhek árán volna csak elkészíthető. Az átvehető mikroadatok elemzéseikhez alkalmazható az elemzés előírt csoportosítása, részletezése. A hivatalos statisztika mutatói más (egyéni) adatokkal is kombinálhatók. A magánszférától átvehető tömeges információk alapján olyan új témakörök is elemezhetők a PPP révén, amelyekre korábban nem voltak megfelelő minőségű adatforrások.

A tanulmány a kockázatok között említi, hogy sérülhet a felhasznált egyedi információk bizalmas jellege, például a megfigyeltek viselkedését, térbeli, illetve anyagi helyzetét stb. illetően, különösen az érzékeny személyes adatokat tekintetében. Új kihívásokat kell kezelni az adatmegosztás üzleti modelljével (*business model for data-sharing*), valamint a nemzeti adatvédelmi és egyéb jogszabály szerinti technikák alkalmazásával. A statisztikák ilyen tömeges információforrásai úgy maradnak elérhető hosszabb időszakban, ha a magánszféra kellően érdekelt a partnerség fenntartásában. Kockázatos, hogy az információforrások decentralizáltak, nem érvényesítik az elfogadott statisztikai szabványokat. A partnertől átvehető információk nem (vagy alig) strukturáltak. A statisztikai célú mintavétel új technikákat igényel, hogy a teljes sokaságot elfogadhatóan reprezentálja.

A tanulmány bemutatja egy mexikói mobilszolgáltató (Telefonica) cég partnerségét, itt megoldott a házon belüli feldolgozás. A gyakorlati példák a következők:

- a közegészségügy a fertőző betegségek riasztási információi, elemzéssel;
- a társadalmi-gazdasági helyzetértékelés, a fontosabb mutatók összeállításával.

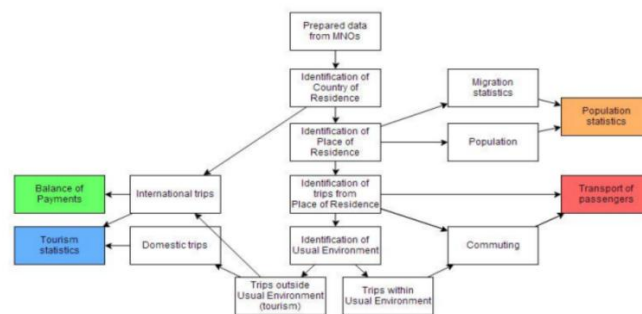
Egy másik mobilszolgáltató (Orange) elérhetővé tette a tárolt hívásinformációkat a végső felhasználó fejlesztéseinek támogatására. Esettanulmány készült egy nonprofit adatgazda (Flowminder) közreműködéséről, az intézmény a 2010-ben Haitin történt földrengés tárolt adataival segítette a koleraterjedésének elemzését.

A szerzők a PPP további lehetőségeként említik, hogy magáncég, vagy megbízható, felkészült szakértő intézmény, regionális statisztikai hivatal közreműködése támogathatja a hivatalos statisztikák előírt feldolgozási és publikációs feladatait. Az egyeztetés során a partnernek garantálnia kell, hogy jogsulatlan harmadik személy nem fér hozzá az információkhoz. Az ilyen partnerség egyik előnye, hogy az adatok feldolgozása kisebb befektetéssel végezhető. A hivatalos statisztikával konzisztens (külső) elemzés is igényelhető komplex feladatként, több szakterületre kiterjedően.

A szerzők az észtországi gyakorlatot említik, ahol a hivatalos statisztika egyes feladatainak ilyen „kiszervezése” révén 2009-ben csökkentették a saját infrastruktúra igényeit, például az ország mobilhívásai letöltött adatainak feldolgozáshoz. Itt azzal is számolni kell, hogy a közreműködő partner – és nem a hivatalos statisztikai intézmény – birtokolja a feldolgozás és elemzés szakmai hátterét.

A negyedik együttműködési lehetőség, hogy „beszállító” állít össze a partnerek által meghatározott mutatókört. Az Egyesült Államokban a fogyasztói árak szakstatisztikájához (*Billion Prices Project – BPP*)<sup>5</sup> az MIT (Massachusetts Institute of Technology) módszertanát fogadták el ellenőrzési célra.

Az OECD tanulmánya kitér a fejlődő országok sajátosságaira is. A PPP elősegítheti a szűkös erőforrásokkal rendelkező hivatalos statisztika adatgyűjtéseit és elemzéseit, a szakpolitikák hatásvizsgálatait. A szerzők esettanulmányai hivatkoznak a témák szakirodalmi forrásaira, kiemelve az általánosítható módszertani alapokat a statisztikai közvetlen adatgyűjtésből nyert és az elérhető tömeges nem hivatalos adatok összekapcsolására, az eredmények igazolására. Az 1. ábra vázolja a turisztikai és utazási adatállományok összekapcsolásának lehetőségeit és fontosabb eredményeit.



Az elért információk lehetséges felhasználása a turisztikai, utazási statisztikákban

<sup>5</sup> A szerzők forrása a projekt honlapja: <http://bpp.mit.edu/datasets/>

A vázlat jobboldalán a blokkok a demográfiai, valamint az utazási mutatósorozatok, mint a feldolgozás eredményei. A vázlat baloldalán jelzett eredmények: a fizetési mérleg és a turizmus statisztikájának előírt mutatói.

A tanulmány áttekinti az adatkezelés nemzetközi jogszabályait, ezen belül a személyes adatok fokozott védelmének európai előírásait. A szerzők a brit és a holland szakirodalmi források alapján közlik a közösségi oldalak tömeges információinak felhasználását például a fogyasztók elégedettségének statisztikai mérésére.

Fontos tanulságok adódnak az esettanulmányok alapján, például a következők:

- Az adatmegosztás a magánszférának is előnyös lehet;
- a megosztott tömeges információk az előírt struktúrával illeszthetők a hivatalos statisztikához;
- a megfigyelt folyamat kategóriái a felhasználás előírt minősége, szintje szerint bonthatók;
- a megközelítéseikben új statisztikai mutatók összeállítása a hagyományostól eltérő szűrés, összegzés eredménye, ehhez speciális informatikai eljárások is szükségesek.
- A nem hivatalos adatforrások alkalmazásának minőségét elemezni kell, azok módszereit rendszeresen követve;
- a PPP esetén költséghatékony üzleti modellek választhatók, ehhez ajánlott a ráfordítások és ár-ajánlatok elemzése, a felkészült partnerek versenyeztetése.

A következő időszak fejlesztései rövidtávon az érzékeny, bizalmas adatok statisztikai célú felhasználásához kapcsolódnak, kellő garanciát nyújtva a jogosulatlan hozzáférés megakadályozására. Hosszabb távon az adatgazdákkal egyeztetett eljárások, protokollok garantálhatják a statisztikai célú adatmegosztás jól követhető, biztonságos, fenntartható módjait.

A holland közlekedésbiztonsági adatgyűjtésből például anonim jelleggel hosszabb távon átvehető nagy részletes térinformatikai adatok a hivatalos statisztikákhoz. A szerzők kiemelik, hogy az ilyen új együttműködési lehetőségek új ismereteket igényelnek a nemzeti statisztikai szolgálat munkatársaitól is, és ez a területi intézményekben is szükséges.

## Függelék:

### Adatok áraData – Big Data

A Központi Statisztikai Hivatal és a Magyar Tudományos Akadémia Társadalomtudományi Kutatóközpont (TK) *Big Data workshop* címmel 2016. május 27-én közös rendezvényt szervezett az együttműködési megállapodásuk keretében. A közismert angol elnevezés Big Data, és az „Adatok áraData” csak egy változat a jelenségkör említésére. A szakirodalom kifejti a jelenség lehetőségeit és egyben lehetséges kockázatait. A feltárt (bányászattal elért és letöltött) információk segítik ugyan a szakmai munkát, de a modellalkotás „adattudós” ismereteket igényel a kísérletek kezdetétől fogva.

A közösségi oldalak és hálózatok bejegyzései a hagyományos személyes kikérdezéses felmérések kiegészítői lehetnek. Új kihívás itt a kutatási és a statisztikai feladatok metodikai hátterének létrehozása, igazolása. Házigazdaként *Vukovich Gabriella*, a Központi Statisztikai Hivatal elnöke nyitotta meg a rendezvényt. A KSH és az akadémiai kutatóközpont (TK) első közös rendezvénye az adatok áradatát, mint gyorsan elérhető adatforrást vizsgálja. Előnye, hogy sok információ viszonylag nagy elemszámmal a világhálón elérhető, de nem nélkülözhetők az elméletileg megalapozott modellek sem. A kutatók és statisztikusok változatlan feladata az eredmények minőségének igazolása, elemezni szükséges az új (esetenként kiegészítő, vagy ellenőrző) megközelítés kockázatait is.

*Giczi Johanna* (KSH) *Big Data* című előadása ismertette az adatáradat módszertani alapjait. Az új paradigma jellemzője, hogy a hagyományos statisztikai munkamenet iránya megváltozik. Az adatok áradata eseteiben az új összefüggést „alulról fölfelé” haladva keresik, ehhez megfelelő algoritmusokat alkotnak, korrelációs számításokkal elemzik a vélt kapcsolatokat. Az eljárások elemző adattudóst (*data scientist*) igényelnek. Az adatáradat új lehetőségeinek felhasználását tervezi a KSH az időmérleg-elemzés adatgyűjtéseiben, valamint a lakásárak interneten elérhető tényadatainak és a felmért mutatók összekapcsolásában.

*Koltai Júlia* (TK SZI) *Okság és Big Data*, valamint *Adatmachináció Big Data adatbázisokon* című előadása a társadalomkutatás kísérleteit ismertette. Az elérhető nem strukturált információk csak elméletileg megalapozott modellekkel vezethetnek igazol-

ható oksági kapcsolatokra. A kutatói program a magyarországi (IWIW) közösségi oldal mintegy 4,5 millió adatára alapozott.

*Szőke Katalin (KSH) Big Data és a hivatalos statisztika módszertani kihívásai, tapasztalatok* című előadása említette a Big Data angol terminológiája szerint 4V (mennyiség, sebesség, változatosság, valamint valóság, minőség) jellemzőket, és összefüggéseiket a hagyományos statisztikai feladatokkal. A légi járatok digitális helyfoglalási rendszerének sok alkalmazója van világszerte, a letölthető, időben változó (online) információk felhasználhatók a repülőjegyek időszaki árindexének összeállításához is. A KSH kísérleti modellje a Budapestről Rómába, Berlinbe, Londonba és Párizsba, valamint onnan visszainduló járatok egységárait dolgozta fel.

*Ságvári Bence (TK SZI) Big Data alapú kutatás/megismerés módszertani, tudományfilozófiai dilemmái és a szociológia útkeresése* című előadása kiemelte, hogy elméleti alapok hiányában a tömegesen elért digitális adat csak zaj. Szükséges az adatáradat kezelésére felhasznált algoritmus auditálása. Javítható a minőség körültekintő szűrésekkel, amelyek elfogulatlan válaszokat szolgáltatnak az érzékeny, üzleti vagy más érdekekhez is kapcsolódó kérdésekre.

*Gilyán Csaba (KSH) Online pénztárgépadatok hivatalos statisztikai célokra való felhasználásának kérdései, tapasztalatai* című előadása kifejtette a kapcsolatokat az adóhatósághoz 2014 őszétől folyamatosan érkező értékesítési adatok, valamint az áfa-bevallások adatátvételei és a KSH hagyományos statisztikai adatgyűjtése között. Elemezték a NAV adatbázisából átvehető aggregált mutatók minőségét, érvényesítve a kapcsolódó szakstatisztikák nemzetközi követelményeit. A megfigyelt adóalanyok (nem csupán kiskereskedelmi üzletek) aggregált havi forgalmi adatai elérhetők és más statisztikai adatforrásokkal kombinálhatók a tárgyidőszakot követő 15. naptól. Ez az új adatforrás megalapozta a célsokasághoz tartozó mintegy nyolcezer vállalkozás adatszolgáltatási terheinek mérsékelését. A KSH kezdeményezte az online pénztárgépadatok kiegészítését a megfigyelési egység főtevékenységének ágazati (TEÁOR) kódjával. A minőségbiztosítás megkövetelte az előforduló kiugró tételek (outlier) egyedi vizsgálatát.

*Lovics Gábor (KSH) A turisztikai határfelvételek NÚSZ Zrt.-től érkező járműforgalmi adatai felhasználásával kapcsolatos tapasztalatok* című előadása a korábbi kézi mérés és a digitális képfelvétel kombi-

nálását mutatta be. Az adatgazda Nemzeti Útdíjfizetési Szolgáltató Zrt. (NÚSZ Zrt.) kameráival rögzített adatok 2008-tól alkalmazhatók statisztikai célra. A nem őrzött határszakaszokon az országból kilépő autók számának becslése regressziós modell alapján történik. A statisztikai adat minőségéhez az átvett nyers adatok szűrése, kiigazítása is szükséges.

*Antal Z. László (TK SZI) Közösségi válaszok az ökológiai kihívásokra (A lakosság fogyasztása CO<sub>2</sub> kibocsátásának mérése)* című előadása ajánlásokat tartalmaz új kombinált indikátorok létrehozására, a klímatudatos, ökotudatos magatartás erősítése érdekében. A diagnózis jellegű adatsorokkal kellene a terápiát megalapozni, de nem a jelenlegi (területi szempontú) indikátorokkal, hanem az adatáradat lehetőségeit felhasználó fogyasztási modell kialakításával. A KSH rendelkezik ehhez a népszámlálás, a mikrocenzus, a háztartásstatisztika mutatóival, amelyek a koncepció szerint kombinálhatók a fontosabb áruk, szolgáltatások fogyasztásából számítható szén-dioxid-terhelés mértékével. A kutatók becsléssel meghatározták, hogy az informatika tömeges, globális felhasználása, annak áramfogyasztása (például internetes művelet távoli szerverével) mekkora mennyiségű szén-dioxid kibocsátását okozza a felhasználás következtében.

*Vancsó Anna (TK PTI) Szövegbányászat az online médiában (corpus and automatic sentiment analysis)* című előadása esettanulmányokat mutatott be az adatáradat társadalomtudományi felhasználására, az internetes bányászat, illetve az elért szöveg (korpusz) elemzéseinek gyakorlati példáival. Figyelembe vették a beszélt és szaknyelvek sajátosságait, az előforduló kifejezések gyors változásait, továbbá az érzelmek jellemző fokozatait. Magyarozatként fontos a vizsgált közösségi háló „beállítottsága”, vagyis az értékítéleteiben a „pozitív”, illetve „negatív” vélemények előfordulási trendje, például a bevándorlással, vagy a droghasználattal összefüggésben.

*Sebők Miklós (TK PTI) A szöveg, mint adat a politikatudományban. Egy interpellációkkal foglalkozó kutatás tapasztalatai* című előadása összefoglalta egy projekt tapasztalatait, amelyben a *Népszabadság* és a *Magyar Nemzet* újságok, valamint az országgyűlés jegyzőkönyveiben elérhető interpellációk korpuszaira gépi szövegelemzést alkalmaztak.

*Zódi Zsolt (TK JTI) A bírósági ítéletek szövegtörzsa, mint Big Data* című előadás a *Bírósági Határozatok Gyűjteménye* digitális állományában véggezhető műveletek gyakorlati tapasztalatait ismer-

tette. A projekt a gazdasági és a polgári perek anyagait kutatta a 131 ezer Bírósági Határozat (BH) tárolt információval. Kapcsolat van a KSH igazságügyi statisztikáinak pertárgy-kategóriái, valamint a szövegtörzsekből a kutatással kiemelt megjelölések között. Ez utóbbiak sokféle változattal emlétenek azonos kategóriát, például a „kártérítés” ügyeit. Ez indokolja a releváns és homogén „ügycsoportok” meghatározását a Big Data (vagyis az adatok áradata) alapján összeállított anyagokban.

**TÁRGYSZAVAK:** Big Data, statisztikai alkalmazások, esettanulmányok, külső adatforrás kockázata

NÁDUDVARI ZOLTÁN

## E számunk szerzői

**CSÁSZÁR JÁNOS**

könyvtáros, Budapest

**CSEH GABRIELLA,**

a Központi Statisztikai Hivatal Könyvtár  
kutatástámogató könyvtárosa

**HOLKA LÁSZLÓ,**

a Központi Statisztikai Hivatal  
ny. vezető főtanácsosa

**HORVÁTH BALÁZS,**

a Központi Statisztikai Hivatal Könyvtár  
tájékoztató könyvtárosa

**KÁLMÁN RITA,**

a Központi Statisztikai Hivatal Könyvtár  
tájékoztató könyvtárosa

**NÁDUDVARI ZOLTÁN,**

a Központi Statisztikai Hivatal  
ny. főtanácsosa

**ORBÁN-SZIRBUCZ ZSÓFIA,**

a Központi Statisztikai Hivatal Könyvtár  
tájékoztató könyvtárosa

**ROVÁCS BARNA,**

a Központi Statisztikai Hivatal Könyvtár  
könyvtárosa

**RÓZSA GÁBOR,**

a Központi Statisztikai Hivatal  
ny. főosztályvezető-helyettese

**SZULY RITA,**

a Központi Statisztikai Hivatal Könyvtár  
kutatástámogató könyvtárosa

**SZUROMI ORSOLYA,**

a Központi Statisztikai Hivatal munkatársa

**TÓTH JÁNOS MÁRK,**

a Központi Statisztikai Hivatal Könyvtár  
kutatástámogató könyvtárosa

## MEGJELENÉS ELŐTT ÁLLÓ KÖNYVÜNK

Segyevy Dániel: *Térképművek Trianon árnyékában. Magyarország néprajzi térképe (1918)*

A kötet a valaha készült egyik legnagyobb méretű és egyik legrészletesebb Magyarországot ábrázoló anyanyelvi térképet mutatja be, amelyet – társaival együtt – levéltári kutatások nyomán nemrégiben azonosított a szerző. A mű, Magyarország néprajzi térképe a magyar térkép-történet egyik kiemelkedő alkotása, az első térkép, amely már egyértelműen az első világháborút lezáró béketárgyalásokra készült Teleki Pál javaslatára, 1918 őszén. A térképsorozat a következő évtizedekben talán a közismertebb vörös térképnél is jelentősebb módszertani hatást fejtett ki mind itthon, mind külföldön. Az olvasó a könyvet lapozgatva végigkövetheti ezt a sokszor rendkívül szerteágazó hatástörténetet, bepillantást kapva a sokszor egymást tagadó korabeli nemzeti érvrendszerek mögött húzódó szövevényes nemzetközi tudományos háló rendkívül tanulságos működési mechanizmusai.

## Nemzetközi Statisztikai Figyelő

ISSN 2416-0814

Megjelenik évente négyszer  
II. évfolyam 2. szám

**Főszerkesztő**

Rózsa Dávid

**Főszerkesztő-helyettes**

Pásztor Angelika

**Rovatszerkesztők**

Holka László (Társadalom)  
Nádudvari Zoltán (Gazdaság)  
Rózsa Gábor (Demográfia)

**Olvasószerkesztő**

Szuly Rita

**Lapterv**

Domonkos László

**Kiadó**

Központi Statisztikai Hivatal Könyvtár  
Budapest II., Keleti Károly u. 5.

**Felelős kiadó**

Fülöp Ágnes

**A szerkesztőség elérhetősége**

Telefon: +36 (1) 345-6406

E-mail: david.rozsa@ksh.hu

**Honlap**

[http://konyvtar.ksh.hu/index.php?s=kiadvany\\_nsf](http://konyvtar.ksh.hu/index.php?s=kiadvany_nsf)