

EDUCATIO®

PEDAGÓGIA | SZOCIOLÓGIA | ÖKONÓMIA | HISTÓRIA | PSZICHOLÓGIA | POLITOLÓGIA

HUSZONNEGYEDIK ÉVFOLYAM, HARMADIK SZÁM

ADATBŐSÉG

Adminisztratív adatok társadalomkutatási kezelése | 3 | VEROSZTA ZSUZSANNA

A régi és új történeti adatok
társadalmi természete | 15 | NAGY PÉTER TIBOR

Adatok és kezelésük a hivatalos statisztikában | 27 | GÁRDOS ÉVA

Közzadatok és nyilvános
adatbázisok: a hozzáférés kérdései | 40 | SZÉKELY IVÁN

Adatokkal történő kormányzás:
dilemmák és perspektívák | 50 | HALÁSZ GÁBOR

Az adathozzáférés és felhasználás
nemzetközi gyakorlatai | 62 | SZÉLL KRISZTIÁN

Köznevelés az adatok bővületében –
az adattermelés forradalma | 73 | SALOMVÁRI GYÖRGY

Sokforrású adatbázis-építés –
buktatók, nehézségek, megoldási kísérletek | 86 | BOGÁROMI ESZTER
– MÁTH ANDRÁS

24

2015
ŐSZ

EDUCATIO®

INTERDISZCIPLINÁRIS SZEMLE AZOK SZÁMÁRA, AKIK
AZ OKTATÁS TÁRSADALMI ÖSSZEFÜGGÉSEIT KERESIK

HUSZONNEGYEDIK ÉVFOLYAM HARMADIK SZÁM | 2015 / ŐSZ | MEGJENIK NEGYEDÉVENKÉNT

Alapító főszerkesztő: KOZMA TAMÁS

Főszerkesztő: FEHÉRVÁRI ANIKÓ

E szám tanulmányait szerkesztette:
VEROSZTA ZSUZSANNA

Lektorálta: LAKATOS MIKLÓS

Szerkesztőbizottság:

BAJOMI IVÁN, BIRÓ ZSUZSANNA HANNA (VALÓSÁG),
CSÁKÓ MIHÁLY (ELNÖK), FEHÉRVÁRI ANIKÓ (KUTATÁS
KÖZBEN), FORRAY R. KATALIN, HRUBOS ILDIKÓ,
KOZMA TAMÁS, LUKÁCS PÉTER, NAGY PÉTER TIBOR,
POLÓNYI ISTVÁN, SÁSKA GÉZA, TOMASZ GÁBOR
(SZEMLE), VEROSZTA ZSUZSANNA (VALÓSÁG)

Szerkesztőség és kiadói hivatal:
1143 Budapest, Szobránc utca 6-8.
Telefon, fax: (06-1) 235-7200

AZ EDUCATIO NEGYEDÉVES FOLYÓIRAT,
ÉVENTE KÖZEL HATSZÁZ OLDALON, ÖTVEN
NYOMDAIÍVEN, MINTEGY NYOLCVANSZERZŐIÍV
KÖZREADÁSÁRA VÁLLALKOZIK.

LAPUNK KAPHATÓ A KIADÓBAN ÉS AZOKBAN
A FŐVÁROSI KÖNYVESBOLTOKBAN, AMELYEK
FOLYÓIRATOK ÁRUSÍTÁSÁVAL IS FOGLALKOZ-
NAK ÉS ELFOGADTÁK LAPUNKAT, ILLETVE
KÖZVETLENÜL IS MEGRENDELHETŐ AZ
EDUCATIO HONLAPJÁN KERESZTÜL:
[HTTP://WWW.EDU-ONLINE.HU](http://www.edu-online.hu)

ELŐFIZETHETŐ KÖZVETLENÜL A KIADÓ CÍMÉN.
AZ ELŐFIZETÉS DÍJA EGY ÉVRE 4896 FT.

EDUCATIO®

QUARTERLY REVIEW OF SOCIAL SCIENCES FOCUSED ON EDUCATION

Editor in Chief: ANIKÓ FEHÉRVÁRI

The journal is published four times a year (600 pages).
Postal address: H-1143 Budapest, Szobránc utca 6-8, Hungary
Annual subscription: \$ 300 on surface delivery or \$ 450 by air
(or the equivalent in another currency), to any address.
Orders may be placed to our postal address or directly
through our website: <http://www.edu-online.hu>

PLEASE, MAKE THE CHEQUE PAYABLE TO OFI, EDUCATIO.

AZ EDUCATIO FOLYÓIRAT MEGJELENÉSÉT A MAGYAR TUDOMÁNYOS AKADÉMIA TÁMOGATJA.
A TANULMÁNYOK BIBLIOGRÁFIAI ADATAI A MAGYAR TUDOMÁNYOS MŰVEK TÁRÁBAN RÖGZÍTÉSRE KERÜLNEK.

© OKTATÁSKUTATÓ ÉS FEJLESZTŐ INTÉZET, 2015.

MINDEN JOG FENNTARTVA.

ISSN 1216-3384

FELELŐS KIADÓ: AZ OKTATÁSKUTATÓ ÉS FEJLESZTŐ INTÉZET FŐIGAZGATÓJA

TIPOGRÁFIA: SALT COMMUNICATIONS KFT.

TÖRDELÉS, NYOMDAI ELŐKÉSZÍTÉS: PATTANTYUS GERGYEL

NYOMDAI MUNKÁK: KOMÁROMI NYOMDA ÉS KIADÓ KFT.

E szám tanulmányainak szerzői

VEROSZTA ZSUZSANNA – *kutató-elemző, Oktatási Hivatal* | LAKATOS MIKLÓS – *szakmai főtanácsadó, KSH* | NAGY PÉTER TIBOR – *egyetemi tanár, WJLF* | GÁRDOS ÉVA – *szakmai főtanácsadó, KSH* | SZÉKELY IVÁN – *kutatóprofesszor, CEU OSA* | HALÁSZ GÁBOR – *tudományos tanácsadó, OFI* | SZÉLL KRISZTIÁN – *kutató-elemző, OFI* | SALOMVÁRI GYÖRGY | BOGÁROMI ESZTER – *PhD-hallgató, BCE* | MÁTH ANDRÁS – *BI&BPR igazgató, Mediaworks* | KIRÁLY GÁBOR – *egyetemi docens, BGF, egyetemi adjunktus BCE* | GÉRING ZSUZSANNA – *tudományos munkatárs, BGF* | CSILLAG SÁRA – *főiskolai docens, BGF, egyetemi adjunktus, BCE* | GÁSPÁR TAMÁS – *tudományos főmunkatárs, BGF* | KÖVES ALEXANDRA – *egyetemi adjunktus, BCE* | SEBESTYÉN GABRIELLA – *BGF* | GYÖRFINÉ KÖRTVÉLYESI FRANCISKA – *ELTE PPK* | MOLNÁR BALÁZS – *adjunktus, DE* | PÁLFI SÁNDOR – *főiskolai tanár, DE* | SZEREPI SÁNDOR – *főiskolai docens, DE* | VARGÁNÉ NAGY ANIKÓ – *adjunktus, DE* | BAJOMI IVÁN – *szociológus, ELTE* | BÚS ENIKŐ – *PhD-hallgató, SZTE Neveléstudományi Doktori Iskola* | HRUBOS ILDIKÓ – *prof. emerita, BCE*

(a tartalommutató folytatása)

98 | VALÓSÁG

„Informatikai fejlesztés és információ fejlesztés AVAGY eljutunk-e oda, hogy az adatbázisok mesélni tudnak az oktatási rendszerről?”

Interjú Kádárné Fülöp Judittal

„Az adatok nem önmagukért beszélnek, mi nézünk valahogy rájuk”

Interjú Rudas Tamással

108 | KUTATÁS KÖZBEN

Iskola a jövőben? Egy jövőkutatási kísérlet bemutatása (KIRÁLY GÁBOR, GÉRING ZSUZSANNA, CSILLAG SÁRA, GÁSPÁR TAMÁS, KÖVES ALEXANDRA, SEBESTYÉN GABRIELLA)

A tanári tevékenység, mint a tanár-diák konfliktusok forrása (GYÖRFINÉ KÖRTVÉLYESI FRANCISKA)

Kisgyermekkori nevelés Magyarországon (MOLNÁR BALÁZS, PÁLFI SÁNDOR, SZEREPI SÁNDOR, VARGÁNÉ NAGY ANIKÓ)

129 | SZEMLE

Egy fontos könyv a statisztikai szemlélet történetéről (BAJOMI IVÁN)

Adat, mérés, fejlesztés – nemzetközi példák az oktatási adatok felhasználására (BÚS ENIKŐ)

Adatbőség az Európai Felsőoktatási Térségben (HRUBOS ILDIKÓ)

141 | ÖSSZEFOGLALÓK

SZERZŐINKHEZ

Az Educatio® minden tárgyilagos álláspontnak helyet biztosít. Minden közleményért szerzője felel. A beérkező kéziratokat megőrizzük.

Az elfogadott kéziratok felhasználási joga négy évre a folyóirat kiadójáé. A tartalmat nem érintő kisebb változtatások, a lap arculatához való igazítás, valamint a cím módosításának jogát a szerkesztőség fenntartja.

Az elektronikusan beküldött tanulmányokkal akkor áll módunkban érdemben foglalkozni, ha azok terjedelme nem haladja meg 35 ezer karaktert. Külön kérjük mellékelni a kézirat egy-két bekezdésnyi angol és magyar kivonatát, a kulcsszavakat, a szerző fontosabb adatait (ahogyan szerzőink között definiálni szeretné önmagát), valamint azt az elektronikus címet, ahová a kefelevonatot kéri). Ha a dolgozat ábrát is tartalmaz, kérjük külön lapon mellékelni a folyóirat tördelési méretének megfelelően (színes ábrákat nem közlünk), grafikon esetén az alapadatokat is kérjük. A hivatkozásokat és lábjegyzeteket a lap tipográfiájának megfelelően szerkesztjük.

A kefelevonaton a szerző javításait három munkanapon belül kérjük e-mailen visszajuttatni, e határidőn túl nem áll módunkban a szerzői javításokat elfogadni. Jelentősebb változtatásokra nincs már mód.

Adminisztratív adatok társadalomkutatási kezelése

Tanulmányunk célja számba venni az adminisztratív adatok társadalomkutatási felhasználásának elméleti és gyakorlati feltételeit és következményeit. Ennek során vizsgálódásunk elvi kiindulópontja az az állítás, miszerint az adminisztratív adat avagy adatbázis – mégha kutatási célra használják is – nem kutatási adat avagy adatbázis. Kutatási célú felhasználásához tehát számos validálási folyamaton szükséges végigmennie, hogy teljesítse a kutatási adatokkal kapcsolatos általános kritériumokat. Itt nyilvánvalóan nem hierarchikus minőségi különbségről van szó, hanem a társadalomkutatási módszertan által az adatok irányában megkövetelt szigorú feltételekről. A nehézséget ez esetben az jelenti, hogy míg a kutatási célú adatgyűjtések esetében ezek a validálási eljárások kidolgozottak, módszertanuknak története és kultúrája van, addig az adminisztratív adatok esetében – a megnövekedett mennyiség és a növekvő igény együttállása miatt – épp most válik egyre sürgetőbbé a kutatási befogadás útjának kiépítése. A tanulmány ehhez a jelenleg is zajló folyamathoz kíván néhány szempont átgondolásával hozzájárulni. Enélkül ugyanis az adatok és adatbázisok önmagukban valóságán alapuló kutatások egész sora indulhat újtúra és ez minden bizonnyal nem kedvez a társadalomkutatás nívójának, hosszú távon.

E szemlélettel és céllal foglaljuk össze az alábbiakban elsőként azokat a tényezőket, amelyek az adminisztratív adatok iránti megnövekedett figyelmet táplálják, majd megvizsgáljuk ezen adattípus természetét és viszonyát a kutatási céllal gyűjtött adatokhoz. Ezt követően az adminisztratív adatokon nyugvó társadalomtudományi kutatás szükséges validálási kritériumait és eljárásait tekintjük át, illusztratív példákat is beemelve. Gondolatmenetünket az adatbázisok összekapcsolásában rejlő lehetőségek áttekintéséig és néhány gyakorlat bemutatásig visszük el.

Adminisztratív adatbázisok

Az adminisztratív adat igazgatási céllal gyűjtött adattípus. Tekintettel arra, hogy az ilyen jellegű adatbázisba kerüléshez jog avagy kötelezettség társul, az adminisztratív adatbázisokat egyedi beazonosíthatóság és a célcsoport teljes lefedése jellemzi.¹ Előbbi tulajdonságuk adatvédelmi és etikai szempontból számos kérdést nyit meg a kutatási felhasználás

^{<1>} Az adattípusok leírását lásd Gárdos Éva és Salomvári György e számban megjelent tanulmányaiban.

nálás felé, ám ugyanennyi potenciált rejt az elemi szintű adatkapcsolás számára. Utóbbi jellemzőjük pedig sok módszertani lehetőséget és kezelendő sajátosságot nyújt a kutatás számára.

Vizsgálódásunk számára az adminisztratív adatok lényeges sajátossága az, hogy ezt az adattípust nem kutatási – hanem igazgatási, szabályozási, regisztrációs, szolgáltatási stb. – céllal gyűjtik. Mindez főként az olyan nagy állami rendszerek feladata, mint például az adózásért, társadalombiztosításért, oktatásért fenntartott szervezetek. Az adatok egy része személyekre, más része szervezetekre vagy egyéb elemzési egységekre vonatkozik. Az adatgyűjtésben meghatározott célon túli felhasználást összességben másodlagos felhasználásnak tekintjük. Ezen belül az adatok statisztikai jellegű feldolgozása már korábban megindult és sokszor hivatalos formát kapott, kutatási hozzáféréseikre azonban csak a törvényi szabályozási keretek változása után nyílt lehetőség.² Hasznosulási – de még nem kutatási – szempontból nézve az adminisztratív nyilvántartások másodlagos felhasználása mellett szóló legfőbb érv gazdaságossági. Ezen adatok közjóságnak számítanak, amennyiben gyűjtésüket, rögzítésüket az adatkezelő közpénzből finanszírozza még akkor is, ha az adatátadás (előbbihez képest elhanyagolható) költsége az adatkérőt terheli. A másodlagos felhasználás szakpolitikai döntésmegalapozás formájában is hasznos lehet. Az adminisztratív adatok alapján végzett hatásvizsgálatok mind a szándékolt intézkedés hatásainak előrejelzésében, mind a beavatkozás utáni állapot-összevetésben fontosak, ami a célpopuláción végzett előzetes és utólagos mérésekben ölt testet. Közvetett haszonként mindezt egy használat által megvalósuló minőségellenőrzési funkció is kiegészíti, javítva a rendszer által végzett kötelező ellenőrzések hatékonyságát (Cseres-Gergely – Scharle, 2008). Mindezeket felismerve számos országban állami szervezetek, ügynökségek jöttek létre kimondottan az adminisztratív adatbázisok összegyűjtése és másodlagos felhasználásának menedzselése céljából. E szempontból erős gyakorlatot és tradíciót találunk Dániában, Hollandiában, Svédországban vagy Finnországban, de az Egyesült Királyságban is kiemelt figyelemmel foglalkoznak a problémával (Dibben et al. 2009).³ E folyamatot számos nemzetközi szervezet (mint az Európai Bizottság, ENSZ, OECD, Világbank) is támogatja és szorgalmazza.⁴

Az adminisztratív adatok szerepének felértékelődése

E helyt nem foglalkoznánk részletesen az evidence based szemlélet nemzeti és nemzetközi szintű szakpolitikai előretörésével. Az adatbőség és adatéhség általános tendenciáját immár vizsgálatunk stabil háttérközegének tekinthetjük (Halász, 2009; 2010). Egy olyan alaphelyzetnek, melyben az adat főszereplővé vált a szakpolitikai döntéshozás legitimációjában, kitermelve mindeközben ennek kritikai megközelítéseit is (lásd erről Hammersley, 2001; Veroszta, 2011).

Az adatfelhasználás jelentőségét ezúttal a társadalomkutatás felől megközelítve azokra a folyamatokra utalnánk, amelyek a kutatók államilag gyűjtött mikroadatokhoz bizto-

² Az adathozzáférésről és annak szabályozásáról részletesen szól Székely Iván e számban megjelent tanulmánya.

³ Az adminisztratív adatok kezelésének nemzetközi gyakorlatáról ír Széll Krisztián e számban megjelent tanulmányában.

⁴ Ennek egyik alapküldetése a 2007-es Istanbul Declaration (OECD 2007).

sított hozzáférést dominálják⁵ (*National Research Council*, 2005). Ilyen folyamat a közvéleményben növekvő aggodalom a magánélet védelme és a titoktartás iránt, amely az önkéntes részvételen alapuló kutatások iránti bizalmatlansággal jár együtt, magas válaszmegtagadást és ezáltal romló minőséget eredményezve. Mindeközben a társadalomról való tudás igénye – az adatéhség – egyre növekvő mértékben és szinteken jelentkezik. A társadalomban zajló differenciálódással speciális, kislétszámú vagy nehezen elérhető csoportok sokaságára fókuszálunk, például célzott beavatkozások, támogatási rendszerek fejlesztése kapcsán. A gazdasági és társadalmi helyzet pontos feltérképezéséhez a részletező adatok mellett azok időbeni alakulása is fokozott figyelmet kap, melyet hatékonyan csak az adminisztratív rendszerek biztosítanak. Ezt a tendenciát az informatikai közeg gyors fejlődése hálózat kutatási, adatbányászati eszközökkel, adattárházakkal, fejlett elemzési és disszeminációs eszközökkel kellőképpen támogatja és rendkívül hatékonyra teheti. Ez persze csak egy megfelelő adatvédelmi-adathozzáférési szabályozási közegben tud jól működni, amely ennél fogva az elmúlt időszakban szintén jelentős változásokon ment keresztül. Az állami adatbázisok mikroadataihoz biztosított kutatói hozzáférést mindemellett az anonimizálási technikákban lezajlott fejlődés és a korlátozott hozzáférés kezelésére kidolgozott speciális módszertani és szervezeti fejlesztések (pl. adatközpontok, ellenőrzött távoli hozzáférési pontok, licenszálás) is támogatják (*National Research Council*, 2005). Az adathasznosítás témájában született programok és szakértői elemzések alapján jól látszik az is, ahogyan az adminisztratív adatok kutatási célú felhasználásáról való gondolkodás a szükségesség és a lehetőségesség lépésein keresztül tartalmi szempontból rendszerint az adatkapcsolásban rejlő fejlesztési potenciálra jut el (Hotz et al., 2008; *The UK Administrative Data Research Network*, 2012).

Szervezeti szempontból szintén egységesnek tűnnek a nemzetközi szakirodalomban és szakmai háttéranyagokban fellelhető ajánlások a kihívások kezelésére (OECD 2013; *The UK Administrative Data Research Network*, 2012; Mosley, 2012; Card et al., 2011). Ezek rendszerint olyan lépéseket jelentenek, mint a nemzeti avagy nemzetközi szintű adatkezelési szervezetek (ügynökségek) létrehozása, a speciális kutatóintézetek kiépítése, az adatfelhasználás és adatkapcsolás jogi kereteinek megteremtése, a működésben az etikai szempontok rögzítése és érvényesítése.

Adminisztratív adat versus kutatási adat

Az adminisztratív adatok kutatási felhasználásának módszertanai áttekintése előtt azokat a főbb jellemzőket határozzuk meg, amelyek ezt az adat- és adatbázistípust a kutatási céllal gyűjtött adattól/adatbázistól megkülönböztetik. Az összegzés alapjául több, a témával foglalkozó munka főbb megállapításait foglaljuk össze (Dixon, 2000; Roos, 2008; Smith et al., 2004; Hotz et al., 2008; McNabb et al., 2009; Garai-Veroszta, 2013; Elias, 2015)

⁵ Jóllehet vizsgálatunkban alapvetően az oktatási szférára és annak releváns adatbázisaira fókuszálunk, a kutatási felhasználás folyamata itt is – akárcsak az evidence based szemlélet elterjedésében (Halász 2009) – az egészségügyből indult el. Az átfogó egészségügyi adminisztrációs adatbázisok – beleértve, sőt kiemelten kezelve a finanszírozási adatokat – lehetővé tették a szolgáltatások, ráfordítások időbeni alakulásának vizsgálatát, az egyéni életutak elemzését, az ellátás regionális különbségeinek feltérképezését, melyeket aztán survey adatokkal kiegészítve, vagy éppen összekapcsolva részletesebb vizsgálatokkal tudtak kiegészíteni (Mor 2009).

Adatmennyiség

Az adminisztratív adatbázisok nagymennyiségű adatot tartalmaznak mind elemszám, mind változók szám tekintetében. A fő különbség a kutatási adatbázisokkal szemben ez esetben az, hogy a kutatónak nem szükségképpen kell korlátoznia az elemszám vagy változók szám alakulását, mérlegelve anyagi szempontokat, illetve a válaszadók tolerancia-küszöbét. A területi korlátok jóval gyengébbek voltak az adminisztratív adatbázisok esetében a „szükség esetén elhagyható” vagy „biztonsági” változók beemelése sem igényel olyan komoly kutatói mérlegelést, mint egy kérdőív esetében. Emellett a nagy elemszám lehetővé teszi relatíve kis létszámú alcsoportok vizsgálatát is.

Rugalmasság

A kutatás elindulása alapvető kutatói döntések előzetes meghozatalát feltételezi. A sorrendiség ennél fogva lényegesen jobban determinálja a survey kutatásokat, hiszen az elemzés céljára létrehozott, vagy szűrt adminisztratív adatbázisba bevont változók köre ismételt adatkéréssel legtöbbször akár utólag is módosítható (ez az integrált adatbázisok esetében nyilvánvalóan nem feltétlenül van így, erről a későbbiekben lesz szó). A rugalmasság időben is értelmezhető. A survey kutatások esetében a longitudinális vizsgálatok tudják ugyan biztosítani az időbeliséget, de például a vizsgálati időhöz képest múltira vonatkozó adatok már csak áttételesen ragadhatók meg. Ehhez képest az adminisztratív adatbázisok egy-egy személy adatait (az adatbázis létrejöttétől kezdődően és utána is kiegészülve) azonos struktúrában, azonos minőségben (megbízhatósággal és érvényességgel) tudják folyamatosan, esetenként nagyon gyakori időbontásban is biztosítani.

Lefedettség

Tekintettel arra, hogy az adminisztratív adatbázisok célja igazgatási, adatai elvben a célcsoport minden tagára kiterjednek. Ez a tulajdonság survey kutatáshoz képest lényegében eliminálja a mintavételi, illeszkedési kérdéseket, megszünteti a válaszadási torzítás mérésének és kezelésének problémáit. Ez az elvi teljeskörűség azonban gyakorlati szinten nem feltétlenül igaz. Semmiképpen nem jelenti például azt, hogy minden egyénről minden változó tartalmaz információt. Az adminisztratív adatbázisoknak is megvannak azok a „vakfoltjai” amelyek nem tudnak megfelelő mennyiségű adatot adni (a minőségről a későbbiekben esik majd szó). Ennek hátterében gyakran szervezési, szabályozási okok állnak (pl. hogy az adat rögzítése nem kötelező) vagy strukturális okok (pl. az adminisztrációs rendszeren, vagy input rendszereken, kategorizáción zajlott időbeni módosítás).

Költségek

A survey kutatások lebonyolításának magas költségigényéhez képest az adminisztratív adatok esetében egy már meglévő adatvagyonról van szó, ami kiiktatja az adatgyűjtés költségeit. Ugyanakkor az adatok kutatási felhasználásához szükséges szervezeti, informatikai és humán feltételek biztosítása – e célra létrehozott professzionális közvetítő szervezetben gondolkodva – tetemes beruházási és fenntartási költség. Utóbbi ráadásul

folyamatos, szemben a survey kutatási programok ad hoc finanszírozási jellegével (ez a longitudinális avagy ciklikus vizsgálatok esetében nyilván nem áll fenn.)

Érvényesség

A validitás kérdése különös jelentőséget kap az adminisztratív adatbázisok esetében. Amíg a kutatási adatgyűjtés esetében a konceptualizálás és operacionalizálás folyamatát előzetesen a kutató végzi el, addig az adminisztratív adatok operacionalizálása egy más – nyilvántartási, szabályozási – célszerűséget szem előtt tartva korábban már lezajlott. Ebből következően az adminisztratív adatok kutatási adatként történő felhasználása véleményünk szerint csak egy ismételt – kutatási célú – validálási szakaszt követően lehetséges, amelynek során feltárjuk az adott változó kontextusát, jelentését, forrását, és mérési problémáit az adott kutatás szempontjából. Ez az újra-operacionalizálás teszi lehetővé, hogy a bevont adatra kutatási adatként építhessünk a maga kutatási érvényességi korlátain belül (avagy vessük el kutatási felhasználását).

Hiányok

Az adminisztratív és kutatási adatbázisokat összevető elemzések mindegyike hangsúlyozza, hogy míg bizonyos tartalmak adminisztratív adatokkal jól lefedhetők, addig e források alapján a társadalomkutatás számos, gyakran vizsgált releváns területe bizonyosan homályban marad. Hogy mást ne említsünk, a motivációk, percepciók, értékek, tervek, vélemények és vágyak továbbra sem rögzíthetők „kartoték-adat”-ként. Az e kontextusban kevésbé evidensen előkerülő – nem szubjektív – területeken is sok a tartalmi hiány. Az adminisztratív adatbázisok mindegyike rendelkezik egyrészt a lefedettség során már említett vakfoltokkal, azaz nem teljes körűen rögzített adatsorokkal.⁶ Másrészt nyilvánvaló, hogy a társadalomkutatás számára elengedhetetlenül fontos objektív adatok jó része sem tartozik az adminisztratív nyilvántartások célkörébe – gondolhatunk itt elsősorban a származási adatokra. Más esetben a hiányt az okozza, hogy változószínten elérhető ugyan adminisztratív adatbázisból a kutatás számára szükséges adat, ám tartalmában nem illeszkedik a vizsgálat céljaihoz. Jellemzően ilyen eset a regionális vizsgálatához elengedhetetlen lakóhelyi adat, amely adminisztratív megközelítésben a hivatalosan bejelentett, állandó lakcímet jelenti, míg a kutatás számára a tényleges, életvitelszerű tartózkodás helye lehet informatívabb.⁷ Összességében a fő problémát nyilván az jelenti, hogy adminisztratív adatok felhasználása esetén a kutatói érdeklődésnek utólag kell hozzáidomulnia a rendelkezésre álló adatok által kínált információtartalomhoz és nem előre tervezetten és strukturáltan gyűjteni be azt, primer kutatás során.

Kontextus

Ahogy az előbbieken is rendre az adatok mögötti nem kutatási motiváció okozta a felhasználási nehézségeket, az adatok kontextusát tekintve is erre hivatkozhatunk. Az ad-

⁶ A diplomás pályakövetési célú adatbázis-felhasználás esetében ilyen jelentős hiány például a külföldi munkavállalás hivatalos rögzítetlensége.

⁷ Különösen így van ez olyan a fiatalabb, vagy átmeneti időszakban lévő célcsoportok esetében, mint például a frissdiplomások.

minisztratív adatbázisok kutatási felhasználása során ugyanis nem rendelkezünk az adatbázis létrejöttére vonatkozó kontextuális adatokkal, háttér-információkkal. Pusztán a már rögzített adatokat, adatbázisokat látva azok keletkezésének módja, története és sokszor oka sem ismert, holott alapvetően határozhatja meg az adat vagy adatbázis természetét. Sokszor a kutatási célú felhasználásnak az operacionalizálás részeként az adatok keletkezési közegének feltárása is evidens része. Adatbázisok esetében ez jelentheti a létrehozás és alkalmazás céljának és az adatgyűjtés menetének megismerését. Egyes adatkörök esetében szintén felmerülhet ilyen kontextualizálási igény. Jó példa erre a foglalkozásokat jelző FEOR kódok változója, amely egyfelől az alkalmazott statisztikai besorolás révén objektív mutatóként kínálja magát, az adat keletkezésének vizsgálata azonban számos esetlegességet, további feltárás iránti igényt mutat.⁸

Tervezettség

Az adminisztratív adatokat és adatbázisokat kutatási szempontból érintő érvényességi és kontextus-problémák összességében a tervezhetőség, sorrendiség sajátosságában azonosíthatók. Primer kutatás esetén az adatgyűjtés tartalma, menete és módja az előzetes elméleti keretrendszerbe ágyazódik be. Az adminisztratív adatok felhasználása esetén a kutató nem tudja kontrollálni az adatok tartalmát és létrejöttük folyamatát sem, ami óhatatlanul gyengíti a módszertani megalapozás minőségét. Hozzátehetnénk, hogy a már meglévő adattartalom mindemellett olykor gyengébb elméleti megalapozásra is csábíthatja a kutatót, de ezt nyilvánvalóan nem tekinthetjük törvényszerűnek. A sorrendiség felborulása egyébként az összevethetőségre is negatívan hat. A nemzetközi survey kutatási projekteket megelőző hosszas adatgyűjtési standardizálási folyamatok épp az egyes adatkörök összevethetősége felé törekszenek. Adminisztratív adatok esetében ez az összevethetőséget lehetővé tévő azonosság visszamenőleg kisebb mértékben ellenőrizhető, biztosítható. Az esetlegesség – annak kis esélye, hogy az adott probléma az egyes országokban egymással összevethető adatok alapján kutatható – nem csak adat, hanem adatbázis szinten is fennáll.

Kontroll

Az adminisztratív és kutatási adatok közti fontos eltérés emellett az ellenőrizhetőség is. Egy kontrollált kutatási helyzetben az adatok keletkezése folyamatosan kézben tartott, ezzel szemben az adminisztratív adatbázisok készen kapott, utólag nem ellenőrizhető adattömeget jelentenek. Nyilván ez esetben az adattisztítási folyamat is egészen más – ám nem elhagyható – eljárásokat követel meg.

Etika

Összevetésünkben nem feledkezhetünk meg az adatok felhasználásához és hozzáféréséhez kapcsolódó etikai vonatkozásokról, még ha a téma részletes kifejtésére ezúttal nem is nyílik tér. A kutatásetika alapelvei nyilván az adminisztratív adatbázisok felhasználása

⁸ Például annak tisztázását, hogy a munkaadói adminisztráció során helyi szinten kik és milyen megfontolások, eseti döntések alapján rögzítik e foglalkozási kódokat.

esetében is működnek, ám sok szempontból más a kritikus kérdések köre. Máshogy, más technikákkal szükséges kezelni és biztosítani például az anonimitást. Ugyanakkor az önkéntesség, a nyilvánosság, a kutatási felhasználás kapcsán lényegében a kutatási adatok másodelemzésekor alkalmazott általános elvek tűnnek alkalmazhatónak.

Hozzáférés

Tulajdonképpen szintén etikai kérdéseket vet fel az adatokhoz való kutatói hozzáférés biztosítása is. Ennek jogi szabályozása adott lehet ugyan, de fontos annak átgondolása, hogy az elvileg szabad hozzáférés ténylegesen mennyire függ az adatgazdától, információs egyenlőtlenségektől, kapcsolati hálótól. Sőt, a szabad hozzáférés nem feltétlenül jelenti az eredmények szabad publikálását.

Adminisztratív adat alapú kutatás

Az adminisztratív és kutatási adatok és adatbázisok közti fenti különbségek figyelembe vételével kell eljárnia a társadalomkutatónak akkor, amikor adminisztratív mikroadatokra alapozott kutatást tervez és végez. A (köz)igazgatás igényeihez igazodó adatbázisok esetében a kutatási újraértelmezés elengedhetetlennek tűnik, kiindulva abból a megfontolásból, hogy a kutatás számára ezen adatokat önmagukban véve üresnek kell tekintenünk és operacionalizálási igényteliséggel szükséges feljűk fordulnunk (Veroszta, 2011).⁹

A kutatói munka számára mindez tehát új feladatok ellátása és új kompetenciák fejlesztésének irányába mutat. Az adminisztratív adatbázisokban rejlő kutatási potenciál kiaknázásához megfelelő kapacitás szükségeltetik mind a tudás és szemléletmód, mind a készségek, mind pedig az eszközök szintjén (Elias, 2015). Tartalmi szempontból felértékelődik az adattudomány szerepe. Kompetencia-szempontból fontossá válik az adatgonddozás és újrafeldolgozás (content curation) képessége. Eszköz-szempontból pedig a nagy adattömegeket kezelni képes megfelelő informatikai háttér lesz elengedhetetlen.

Az adminisztratív adat alapú kutatások főbb lépéseit a brit Administrative Data Liaison Service (ADLS)¹⁰ és az amerikai National Research Council (2005)¹¹ iránymutatása alapján áttekintve számos olyan szükséges eljárást és megfontolást emelhetünk ki, amelyek a klasszikus empirikus társadalomkutatásnak nem részei. Részben ilyen a korábbi azonos tartalmú adatkérések előzetes feltáró vizsgálata (csak részben, mivel nyilván ezzel párhuzamba állítható a téma kutatási előzményeinek feltárása, amely minden kutatás része). Ezekből a korábbi eljárásokból sok előzetes információ gyűjthető a hatékony adatkéréshez és feldolgozáshoz. Az adminisztratív adatbázisok hozzáféréseinek kiterjedése sok szempontból új helyzetet teremthet a kutatásban. Az adatmegosztás általában is a kutatás és a kutatói közeg nyitottságát növelheti, bizonyos értelemben az átláthatóság irányába mozdítva a társadalomtudományi kutatást. Azonos empirikus bázison az eredmények megismételhetővé, ellenőrizhetővé válnak. Ebben az adminisztratív adatokhoz való szabad hozzáférés előmozdító folyamat.

⁹ Az adatok társadalomtudományi újrafelhasználásáról ír e számban megjelent tanulmányában Nagy Péter Tibor.

¹⁰ Az Egyesült Királyságban alapított, az adminisztratív adatok kutatási felhasználást támogató hivatal. Internetes elérhetősége: <http://www.adls.ac.uk/>

¹¹ <http://www.nationalacademies.org/>

Szintén szükséges előzetes döntést hozni arról, hogy a kutatás céljaihoz mely aggregáltsági szintű adatok szükségesek. A legtöbb esetben elengedhetetlen az egyedi szintű (mikro)adatok lekérése, ám kétségtelenül ez a leginkább nehezített eljárás. Ezért több ponton érdemes lehet mérlegelni a jóval könnyebben elérhető aggregált adatok, makrováltozók bevonását is, ahol ez nem sérti a várt kutatási eredményeket. A döntés e tekintetben igen nagy jelentőségű, hiszen a nagyobb részletezettségű adatok jobb használhatósága egyértelműnek tűnik, kifinomult többváltozós modellépítéshez a mikroadatokra feltétlenül szükség van. A kérdés ezért inkább az, hogy az egyéni szintű adathiányok milyen külső makrováltozók bevonásával pótolhatók. Az adatokra vonatkozó kutatói döntésben külön is tekintetbe kell venni a longitudinalitás lehetőségét, amely az adminisztratív adatbázisok esetében gyakran jól hozzáférhető opció.

Az adatokra vonatkozó kutatói döntés utáni fontos szakasz az adatbázist birtokló szervezetek megkeresése és adatgyűjtési és adatkérési dokumentációjuk tanulmányozása. Előbbi rendszerint a kutatás ismertetését, az érzékeny adatok kérésének indokoltságát, a vállalt biztonsági követelményeket tartalmazza. Utóbbi pedig annak vizsgálatát, hogy egészen pontosan milyen változók állnak rendelkezésre. Ennek természetesen elengedhetetlen feltétele a megfelelő transzparencia az adatgazda részéről. A kutatás előkészítése során azzal is számolni kell, hogy az adatgyűjtési – ez esetben adatkérési – szakasz a primer társadalomkutatási adatfelvételhez képest jóval kevésbé kiszámítható, tervezhető. Nagyban múlik ugyanis az érintett szervezet(ek) adatközlési feltételeitől és gyakorlatától (sőt gyakorlottságától).

Adatkapcsolások

Az adminisztratív adatok és adatbázisok kutatási felhasználásának áttekintése után a mikroszintű adatok összekapcsolásáról érdemes szót ejtenünk, hiszen ezen eljárások maguktól értetődően követik az adatbázisok egyre kiterjedtebb rendelkezésre állását. Az adatkapcsolás „data linkage” ezen áttekintésünkben mindazon eljárásokat jelenti, amelyben egy adminisztratív adatbázis valamely megfigyelési egységének adataihoz, illetve adatsorához más adatokat kapcsolunk. A technikai jellemzőkön túl tartalmilag az adatkapcsolásra olyan eljárásként tekintünk, amely több adatbázis információtartalmának integrálását célozza meg. Ez a lehetőség a készen kapott, limitált és rigid adattartalommal, ám teljes lefedettséggel rendelkező adminisztratív adatbázisok esetében különösen fontos szerephez jut és új módszertani eljárások felé nyit utat. Az összekapcsolás kulcsa – mint azt a továbbiakban jól látjuk majd – egy (vagy több) egyedi azonosítást lehetővé tévő változó avagy azok kombinációja.

Az adminisztratív adatbázis alapú adatkapcsolásoknak több típusa is lehet, főbb módjait az alábbiakban kategorizáljuk.

Adminisztratív adatbázison belüli összekapcsolás

Ezzel az eljárással ugyanazon adatbázis megfigyelési egységének egyedi adatsoraihoz kapcsolunk időbeliségében vagy tartalmában kiegészítő változókat. Ilyen lehetőség például a longitudinális, időben visszamenőleges vizsgálat egyazon adatbázison belül. Azonos adatbázison belül az egyedi azonosítók rendelkezésre állása nem okoz problémát (hacsak a feltöltöttségi időben nem különbözik).

Adminisztratív adatbázisok közti összekapcsolás

A több adminisztratív – azaz teljeskörű – adatbázis közti egyénsoros összekapcsolás feltétele az ezt lehetővé tévő, mindkét adatbázisban megtalálható egyedi azonosító kód (match-merge eljárás), avagy a személyes adatok olyan kombinációja, ami egyéni megkülönböztető adatképpé áll össze (deterministic linkage). Mindkét módon eljártak már a hazai társadalomtudományi gyakorlatban is. A hazai diplomás pályakövetési célú adatkapcsolás során például különböző időszakokban az összekapcsolás mindkét módját alkalmazták. A cél itt a felsőfokú tanulmányi adatok és a munkaerő-piaci kimeneti jellemzők összekapcsolása volt. Előbbi esetben a Felsőoktatási Információs Rendszerben a 2009-ben abszolutóriumot szerzettek teljes évfolyamára vonatkozóan rendelkezésre álló képzési és szociodemográfiai adatsorokat egészítették ki több állami adminisztratív rendszer adataival,¹² a kapcsolati kód kialakításához felhasználva – majd később törölve – a végzetek személyes azonosításra alkalmas adatait (Fodor – Veroszta, 2013). Ilyen személyes adatok jellemzően a név, születési dátum, lakcím stb. A pályakövetési célú adatkapcsolás későbbi – 2013¹³-as illetve 2014¹⁴-es – fázisaiban már a másik ismertített metódus, a több adatbázisban megtalálható egyedi azonosítók alapján történt az összekapcsolás (mely ez esetben a személyek TAJ számából, a kapott hashkód-képző eljárással előállított anonim kód volt). Ez az eljárás lényegesen jobb lefedettséget eredményez (Nyüsti-Veroszta, 2014a; Nyüsti-Veroszta, 2015).

Adminisztratív adatok összekapcsolása makrováltozókkal

Ezekben az adatkapcsolásokban az adminisztratív adatbázisban lévő egyénsoros (ill. a megfigyelési egységre vonatkozó) mikroadatokhoz nem egyéni szintű adatok kapcsolódnak más adatbázisokból, hanem az egyedi adatok magasabb aggregátsági szintjéhez kapcsolunk külső makrováltozókat. Ilyen kontextuális információt nyújthatnak például a hivatalos statisztikai adatgyűjtés mutatói, indexei. Ez az eljárás adminisztratív és survey adatbázisokon egyaránt gyakran alkalmazott a társadalomtudományi kutatásban, nagy szolgálatot tesz hiányzó kutatási dimenziók pótlásában.

Adminisztratív és survey adatbázisok közti összekapcsolás

Ezekben a keresztmetszeti avagy longitudinális adatkapcsolásokban a teljeskörű adminisztratív adatok megfigyelési egységre vonatkozó egyedi adatsorai egészülnek ki survey

¹² Az adatkapcsolásba bevont szervezetek: Felsőoktatási Információs Rendszer (FIR), Adó- és Pénzügyi Ellenőrzési Hivatal (APEH), Országos Egészségbiztosítási Pénztár (OEP), Foglalkoztatási és Szociális Hivatal (FSZH). A vizsgálat alapsokasága a 2008/2009-ben végzetek teljes köre, az adatgyűjtés a 2010-es státuszra vonatkozott.

¹³ Az adatkapcsolásba bevont szervezetek: Felsőoktatási Információs Rendszer (FIR), Nemzeti Adó- és Vámhivatal (NAV), Országos Egészségbiztosítási Pénztár (OEP). A vizsgálat alapsokasága a 2009/2010-ben végzetek teljes köre, az adatgyűjtés a 2012-es státuszra vonatkozott.

¹⁴ Az adatkapcsolásba bevont szervezetek: Felsőoktatási Információs Rendszer (FIR), Diákhitel Központ Zrt., Magyar Államkincstár (MÁK), Nemzeti Adó- és Vámhivatal (NAV), Országos Egészségbiztosítási Pénztár (OEP), Országos Nyugdíjbiztosító Főigazgatóság (ONYF), Nemzeti Munkaügyi Hivatal (NMH). A vizsgálat alapsokasága a 2009/2010-ben illetve 2011/12-ben végzetek teljes köre, az adatgyűjtés a 2013-as státuszra illetve visszamenőlegesen három évre vonatkozott.

kutatásból származó szintén egyedi adatsorokkal. A regiszter alapú összekapcsolás nagy értéke a fentiekben részletezett két típusú (kutatási, ill. adminisztratív) adattartalom érényekinek kombinálása. A megvalósítás történhet a teljes populációra vonatkozó adminisztratív adatbázison végzett, ahhoz kötött utólagos survey kutatással, amelynek adatai aztán az adminisztratív adatbázis egyedi azonosítói alapján visszavezethetők (*Sakshaug et al.*, 2012; *McNabb*, 2009).

Más esetben már meglévő különböző, ám azonos alappopulációt lefedő adminisztratív, illetve survey adatbázisok egyedi szintű összekötéséről van szó (*Groen*, 2012). Az ehhez alkalmazott eljárásaként rendelkezésre álló valószínűségi adatkapcsolás (probabilistic record linkage) statisztikai eljárással azonosítja a két, azonos alappopulációt lefedő adatbázis tagjai közti kapcsolatot. Az eljárás matematikai alapjainak kidolgozása Fellegi-Sunter (1969) nevéhez fűződik. A módszer (akárcsak a deterministic linkage) a mindkét adatbázisban rendelkezésre álló egyéni szintű, személyes adatok kombinációjára épül, de nem követeli meg a megfigyelési egység szintű teljes egyezést. Az eljárás megkülönböztet egyező, nem egyező és bizonytalanul egyező kapcsolódást a különböző forrású adatsorok között. Utóbbi esetben az egyezés statisztikai valószínűségével számol, és ezt rendeli hozzá az egyedi adatsorhoz. A probabilistic record linkage alkalmazása a társadalomtudományokban az erős statisztikai háttér mellett a számítástechnikai lehetőségek és ismeretek fejlődését is előfeltételezi (*Winkler*, 1999). Amellett, hogy az eljárás a különböző adatbázisok adattartalmának egyesítésével azok információtartalmát tudja megszoszorozni, fontos szereplő az adatminőség javítására – a lefedettség, mintavétel, válaszadás problémáinak kezelésére – alkalmazott statisztikai eszközök között (*Kreuter et al.*, 2010; *Schnell*, 2013). Ilyen jellegű és célú adatkapcsolást kísérleti jelleggel a Diplomás Pályakövetési Rendszer keretein belül is végeztek survey jellegű online adatfelvételekben feltételezett szisztematikus hiba azonosítása céljából. Ennek során ugyanazon végzett évfolyamra vonatkozóan vetették össze a survey kutatási adatok eltérésének mintázatait az adminisztratív adatbázis (FIR) alapján lefedett teljes célpopulációtól (*Nyüsti-Veroszta*, 2014b).

Az adminisztratív adatok kutatási felhasználásának elméleti és gyakorlati feltételeit és következményeit vizsgálva tanulmányunkban az adminisztratív és survey kutatási adatok eltéréseit és felhasználásuk módszertani lehetőségeit vetettük össze. Az adminisztratív adatok kutatási felhasználásához kapcsolódó szempontokat ezt célzó gyakorlati eljárásokkal egészítettük ki. Jóllehet munkánk során kevésbé elemző, sokkal inkább áttekintő megközelítést alkalmaztunk, elméleti kiindulópontunkhoz ennek során mindvégig igazodtunk. Eszerint a társadalomkutatásban egyre nagyobb figyelmet és súlyt kapó adminisztratív adatok és adatbázisok csak a megfelelő kutatási validálási eljárások után tekinthetők kutatási felhasználásra alkalmas adatállománynak. A validálási eljárás során pedig – az eltérő adatközegek ellenére – a társadalomkutatás hagyományos szemléletmódjához szükséges és lehetséges igazodnunk. Tanulmányunkban ennek a fontosnak ítélt adatmunkának a főbb megfontolásait és lépéseit vettük számba.

IRODALOM

- CARD, D., CHETTEY, R., FELDSTEIN, M. & SAEZ, E. (2011): *Expanding Access to Administrative Data for Research in the United States*. National Science Foundation. pp. 6.
- CSERES-GERGELY Zs. & SCHARLE Á. (2008): *Az államigazgatásban keletkező adatok nyilvánosságáról*. Kézirat. MTA-KRTK-KTI. pp.12.
- DIBBEN, C., GOWANS, H., ELLIOT, M., ANTTILA, C., BOYLE, P., KAYE, J., MCLENNAN, D., NOBLE, M., SMITH, G. & WILKINSON, K. (2009): *Encouraging the wider and more creative use of administrative data in the UK - the 'Administrative Data Liaison Service*. pp.10.
- DIXON, S. (2000): Using administrative data sources in labour market Research. *Labour Market Bulletin* 2000/02 Special Issue, pp. 26-30.
- ELIAS, P. (2015): *New forms of data – new opportunities for research*. Trans-Atlantic Platform Social Sciences and Humanities, 11th February 2015.
- FELLEGI, I. A. & SUNTER, A. B. (1969): A theory for record linkage. *Journal of the American Statistical Association*. 64. 328. pp. 1183–1210.
- FODOR SZ. & VEROSZTA ZS. (2013): Államigazgatási adatok pályakövetési célú integrációja a hazai gyakorlatban. In: GARAI O. & VEROSZTA ZS. (Eds.) *Államigazgatási adatbázisok a diplomás pályakövetésben*. Educatio Társadalmi Szolgáltató Nonprofit Kft. pp. 83-128.
- GARAI O. & VEROSZTA ZS. (Eds.) (2013): *Államigazgatási adatbázisok a diplomás pályakövetésben*. Educatio Társadalmi Szolgáltató Nonprofit Kft. pp. 128.
- GROEN, J. A. (2012): Sources of Error in Survey and Administrative Data: The Importance of Reporting Procedures. *Journal of Official Statistics*. 28. 2. pp. 173–198.
- GROGGER, J. (2013): Bounding the Effects of Social Experiments: Accounting for Attrition in Administrative Data. *NBER Working Paper* No. 18838.
- HALÁSZ G. (2009): Tényekre alapozott oktatáspolitikai és oktatásfejlesztés. In: Pusztai G. – Rébay M. (szerk.) *Kié az oktatáskutatás?* Csokonai Könyvkiadó, Debrecen. pp.187-191.
- HALÁSZ G. (2010): *Az oktatáskutatás globális trendjei*. Vitaanyag. A Magyar Tudományos Akadémia Pedagógiai Bizottsága. Kézirat. Budapest. pp. 67.
- HAMMERSLY, M. (2001): *Some Questions about Evidence-based Practice in Education*. Paper presented at the symposium on „Evidence-based practice in education” at the Annual Conference of the British Educational Research Association, University of Leeds, England, September 13-15.
- HOTZ, V.J., GOERGE, J., BALZEKAS, J. & MARGOLIN, F. (2008): *Administrative Data for Policy-Relevant Research: Assessment of Current Utility and Recommendations for Development*. A Report of The Advisory Panel on Research Uses of Administrative Data of the Northwestern University/ University of Chicago Joint Center for Poverty Research. pp. 116.
- KREUTER, F., MÜLLER, G. & TRAPPMANN, M. (2010): *Nonresponse and measurement error in employment research: Making use of administrative data*. *Public Opinion Quarterly*, Vol. 74, No. 5. pp. 880–906.
- MCNABB, J., TIMMONS, D., SONG, J. & PUCKETT, C. (2009): Uses of Administrative Data at the Social Security Administration. *Social Security Bulletin*, Vol. 69. No. 1.
- MOR, V. (2009): Administrative Data Systems in Social Science Research on Health and Aging. In: McKinlay, J.B. – Marceau, L.D. (Eds.): *Behavioral and Social Science Research Interactive Textbook*.
- MOSLEY, H. (2012): *Pes Data Products and Services for Labour Market Policy Evaluation in Germany*. Peer Review on “Evaluation of Labour Market Policies and Programmes: the use of data-driven analyses, Brussels. 19-20 November, pp. 11.

- NATIONAL RESEARCH COUNCIL (2005): *Expanding access to research data: Reconciling risks and opportunities*. In Panel on Data Access for Research Purposes, Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academies Press. pp. 120.
- NYÜSTI SZ. & VEROSZTA ZS. (2014a): *Diplomás pályakövetési adatok 2013*. Adminisztratív adatbázisok integrációja. Educatio Társadalmi Szolgáltató Nonprofit Kft. pp. 114.
- NYÜSTI SZ. & VEROSZTA ZS. (2014b): *Linking Administrative And Survey Data In Educational Research. Possibilities And Limitations Of Using Probabilistic Record Linkage In Graduate Career Tracking*. Poster presented at ECER 2014 „The Past, the Present and Future of Educational Research in Europe” Porto, 1 - 5 September.
- NYÜSTI SZ. & VEROSZTA ZS. (2015): *Adminisztratív adatbázisok integrációja 2014 - Gyorsjelentés*. Educatio Társadalmi Szolgáltató Nonprofit Kft. pp. 20.
- OECD (2007) *Istanbul Declaration*
- OECD (2013) *New Data for Understanding the Human Condition: International Perspectives*. OECD Global Science Forum Report on Data and Research Infrastructure for the Social Sciences pp. 52 .
- ROOS, L.L., BROWNELL, M., LIX, L., ROOS, N.P. WALLD, R. & MACWILLIAM L. (2008): From Health Research to Social Research: Privacy, Methods, Approaches. *Social Science & Medicine* 66(1) pp.117-129.
- SAKSHAUG, J.W., COUPER, M.W., OFSTEDAL, M.B. & WEIR, D.R. (2012): Linking Survey and Administrative Records. Mechanisms of Consent. *Sociological Methods Research*, 2012 vol. 4,1 no. 4. pp. 535-569.
- SCHNELL, R. (2013): *Linking Surveys and Administrative Data*. German Record Linkage Center, Working Paper Series.
- SMITH, G., NOBLE, M., ANTTILLA, C., GILL, L., ZAIDI, A., WRIGHT, G., DIBBEN, C. & BARNES, H. (2004): *The Value of Linked Administrative Records for Longitudinal Analysis*. Report to the ESRC National Longitudinal Strategy Committee. pp. 48.
- The UK Administrative Data Research Network (2012): *Improving Access for Research and Policy*. Report from the Administrative Data Taskforce. pp. 59. .
- VEROSZTA ZS. (2011): Adatok az oktatáspolitikára és a kutatás közti térben. *Educatio* 2011/4 pp. 521–534.
- WINKLER, W.E. (1999): *The state of record linkage and current research problems*. Statistical Research Division, US Census Bureau, pp. 15.

A régi és új történeti adatok társadalmi természete

Elég nagy biztonsággal kijelenthető, hogy nincs olyan történelmi, történetiszociológiai téma, ami objektíve ki lenne merítve, azaz amiről kizárható lenne, hogy valamikor a jövőben

- régi adatok új szempontból történő, új módszerrel készülő újraelemzése alapján,
- régi adatokból újjáalkotott metaadatok alapján,
- új adatok alapján,
- régi és újabb adatok összevetése alapján,
- más hasonló témákról szóló elemzésekkel való összevetés alapján, vagy akár
- egyetlen elemzés belső érvrendszerének, adatközléseinek újragondolása alapján újabb elemzés ne lenne írható.¹

Ha egy kvalitatív kutató vagy hermeneuta idáig eljut – azt további olvasásra biztatom: ugyanis a cikkben az „adat” nemcsak a kvantitatív vagy statisztikai alapú tanulmányok forrásául szolgáló ismeretet, hanem bármely tudományos mű forrását jelenti majd.

E cikkben minden egyes altémán belül először arra szeretnék kitérni, hogy az „adatok új világa” (mely valószínűleg a számítástechnikával, s különösen az internettel vette kezdetét) hogyan alakítja át – funkcionalistán szólva – a tudományos témákat és módszereket, majd arra, hogy – konfliktusos paradigmában fogalmazva – a tudományos közösségek milyen érdekviszonyok alapján kezelik, illetve ismerik el egyáltalán releváns adatnak az új információs helyzet „termékeit.”

Három – kétségtelenül eltérő méretű – altémát szeretnék kifejteni ezen a módon: a régi típusú, de új adatok keletkezését; az új típusú új adatok keletkezését; régi adatok újrahazsnosulását. Ez a tanulmány csak a „régidolgokról” szóló adatok új kontextusaival foglalkozik, nem tér ki arra, hogy napjaink jelenségeiről töménytelen mennyiségű új adat keletkezik, melyek egyrészt befolyásolják a régebbi dolgok, párhuzamos trendek „kifutásáról” való tudásunkat, ezért a napjainkról szóló adatok robbanásszerű bővülése a közelmúlt történetének folyamatos átírására késztet – másrészt a „gyorsuló időben” maguk is hamar történeti forrássá válnak. Nyugodtan feltételezhetjük, hogy amiképpen az első adatrobbanás korának történészei és levéltárosai nem látták át mindazo-

¹ A tanulmány illeszkedik a International cooperation in the social sciences and humanities. Comparative Socio-Historical Perspectives and Future Possibilities. c európai projektbe <http://interco-ssh.eu/en/> melynek magyar partnere a WJLF (<http://wesley.hu/sociology>).

kat a módszereket, ahogyan koruk adattömegét – pl. a frissen megszületett napisajtót, az alfabetizmus általánossá válása következtében nagyságrendekkel megnövekedett magánlevelezést, a városias ügyintézés, üzemszerű termelés exponenciálisan megnőtt írásbeliségét – a jövő tudósai használni fogják, (vagy használhatnák, ha megőrizték volna...) úgy valószínűleg ma sem látjuk pontosan, hogy spontán keletkező digitális adattömeg hogyan hasznosul majd a jövőben, illetve, hogy az „evidence based” döntéshozás korában – az egymással konkurráló álláspontok alátámasztására felépített strukturált, „tudományos” adattömegrész mennyire éli túl az alátámasztandó álláspont kisebbségben maradását.

Ha a terjedelem engedné, a „jövő történéseire gondolva” logikailag két – hipotetikus – fejezetet lehetne csatolni: az új dolgokról szóló új, de strukturálatlan adatok keletkezésének társadalmi természetét, végül a strukturált, illetve „tudományos” adat születését és visszatöltődését az adatok körforgásába.

Régi típusú, de új adatok keletkezése

Vannak olyan témák, ahol a szorosan vett kérdésre vonatkozóan „nem kerülnek elő” újabb adatok. Sok évtized óta nincs például újabb adat a római történelem királyságkori szakaszának politikatörténetével kapcsolatban, s nem kerülnek már elő újabb Arisztotelész szövegek sem. A „nincsenek már újabb források” mint diagnózis, a tudományos közösségben jóval megelőzi az adatrobbanás korát, így példáink jó része lesz kissé régebbi.

Először is szögezzük le: nem zárható ki, hogy előkerülhetnek – például újabb autópálya-építések során – újabb adatok ilyen régi dolgokról is. Az újabb autópálya-építések, sőt a kunyhók és bádoggvárosok helyére épülő, jelentősebb alapozást igénylő építkezések – Kelet Európában, Ázsiában és Afrikában – természetesen nem elsősorban a helyi gazdasági és politikai elitek ambícióin múlnak, így az ezzel kapcsolatos szabályozás sem egyszerűen a helyi maffia-eliték és gyenge tudományos elitek erőviszonyának kérdése,² hanem a nemzetközi tudományos közösség „dolga és felelőssége” is. Az „új adatok előkerülése” – pontosabban annak valószínűsége – bármely ókori, őskori, középkori témában annak a kérdése is, hogy a nemzetközi tudományos közösség belső erőviszonyai – az egyes témák „fontossága,” az irántuk való „érdeklődés” mértéke – mely témák esetében teszik lehetővé, hogy a közösség a szükségképpen korlátos lobbizási kapacitását arra fordítsa, hogy az építkezési projekt nagy nemzetközi finanszírozóin keresztül – finanszírozási feltétellel téve – rászorítsa a helyi kormányzatokat a régészeti munkálatok elvégzésére. Más esetekben maga a nemzetközi tudományos közösség mozgósítja az erőforrásait, hogy valamely építkezés következtében a régészet elől végképp elzáródó területek „soron kívül” feltárássra kerüljenek. (E folyamatok gyakran rejtve zajlanak, vagy épp csak elmaradásukat konstatálhatjuk, de vannak a tudományt kedvelők számára pozitív példák is. Ilyen például az asszuáni gát építése körül megvalósult gigantikus régészeti projekt.³)

A forrásokat, illetve potenciális forrásokat tudatosan pusztító külföldi folyamatokba való beavatkozást természetesen korlátozza, hogy a nemzetközi tudományos közösség szervezetei – mint ezt számos példa igazolja – általában az alacsony intenzitású (magyarán pacifista, mondhatnánk „defetista”) amerikai és európai külpolitikai vonalvezetés

² http://magyarnarancs.hu/belpol/a_regeszeti_torvenyi_ujraszabalyozasa_kapari_kurta-61361

³ <https://oi.uchicago.edu/research/projects/nubia-salvage-project>

hívei.⁴ Jelentős nemzetközi gazdasági, ill. katonai konfliktus vállalása nélkül pedig nyilvánvalóan nem lehet(ett) megvédeni a kínai kulturális forradalom áldozatául eső kincseket,⁵ a tálibok által elpusztított buddhista örökséget,⁶ vagy az ISIS csapatok útjába eső ókor-történeti szempontból egyedülálló rom-városokat,⁷ régészeti lelőhelyeket.

Igazi új adatok előkerülése azért sem kizárható, mert periférikus (vagy éppen autokratikus hatalmak által ellenőrzött) levéltárak, közgyűjtemények, magángyűjtemények időről időre hozzáférhetővé tesznek olyan korábban előkerült forrásokat, melyeknek néha létezéséről sem tudunk.

A modern történelem fordulatait megvilágító adatok titkosság alóli feloldása könnyen átlátható csatározások, vagy szabályozások tárgya. Ez előállhat diktatórikus rendszerek bomlásakor és demokratikus viszonyok között is. Előbbire példa a magyar-szovjet kapcsolatok történetét jelentősen módosító Jelcin dosszié megjelenése,⁸ utóbbira a második világháború hadtörténetét új megvilágításba helyező Enigma⁹ -story, vagy a németországi Fischer-vita alapjául szolgáló egyetlen dokumentum¹⁰ megtalálása. A szimbolikus politizálásra nagyobb gondot fordító hatalmak szemében évezredes iratok vagy leletek is jelentőséggel bírnak – gondoljunk csak a Holt-tengeri tekercsek Vatikánba került részeinek hosszú titkosságára,¹¹ vagy arra, hogy azok a régészeti adatok, amelyek az orosz állam megalapozásában betöltött viking szerepről szóltak, hosszú ideig államtitkot képeztek a Szovjetunióban.

Az pedig *gyakorlatilag* bizonyos, hogy adott történeti témával *egyidejűleg* történő események meglévő vagy újonnan felfedezett *relációi* előkerülnek. Azaz újabb adatok valamiféle, másutt egyidejűleg vagy kissé korábban zajló eseményről, amelyek a világban zajló események teljes komplexitását feltételezve valamiképpen kihathattak a vizsgált dolog lefolyására, vagy olyan későbbi jelenségekről szóló információk, melyek valamiképpen a vizsgált esemény hatását tükrözik. Itt nem arról az elméleti tézisről van szó, hogy minden mindennel összefügg, hanem arról, hogy a kutatás természete a szélesebb összefüggések keresése. A 16-17. század magyarországi bibliográfiájában például jól látszik, hogy régebben a kuruc-labanc, török-Habsburg erőviszonyok alakulásáról szólt elsősor-

⁴ <http://www.europeaninstitute.org/index.php/93-european-affairs/february--march-2010/964-dieter-dettkes-germany-says-no-the-iraq-war-and-the-future-of-german-foreign-and-security-policy>

⁵ <http://www.historia.hu/userfiles/files/2011-07/Jordan.pdf>

⁶ <http://www.nytimes.com/2001/03/19/world/19TALI.html>

⁷ <http://www.nytimes.com/roomfordebate/2014/10/08/protecting-syrias-heritage>

⁸ <http://beszelo.c3.hu/cikkek/a-%E2%80%9EJelcin-dosszie%E2%80%9D> E történeti forrásfeltárás – ahogy az az orosz levéltárgyvet azóta is jellemezni látszik – maga is nagyhatalmi politikai célokat szolgált.

⁹ <http://www.gcsu.edu/math/docs/2013Capstone/callahan.pdf> A német tengeralattjáró-háború irányítására szóló kódrendszer megfejtésének története moralista diskurzusok alapjául is szolgált: legitim volt-e tudatosan feláldozni jelentős számú angol tengerészt, még több tengerész későbbi megmentése, illetve a háború sikeres befejezése érdekében.

¹⁰ Az első világháborús német kormány által megtárgyalt, a hadicélokra vonatkozó dokumentum robbantotta fel azt az önképet, miszerint a német vezérkar volt agresszív, a polgári politikusok pedig mérsékelni akartak, az utóbbiak tehát lehetnek egy mai Németország pozitív hősei... vö: http://www.tankonyvtar.hu/hu/tartalom/tamop425/2011_0001_520_bevezetes_a_tarsadalomtortenetbe/ch03s04.html és Szabó Miklós: Nyugatnémet vita az újkori német történelemről. - In: Világtörténet, 1987. (9. évf.), 4. sz., pp.131-142..

¹¹ Moróczy Géza (1988): Kiáltó szó a pusztában – A Holt-tengeri tekercsek, Osiris Kiadó, Budapest. A Holt-tengeri esszénus tekercsek– minthogy az Újszövetség szavait más kontextusban tartalmazzák – az Újszövetség szavainak isteni sugallmazásának megkérdőjeleződését jelentették.

ban a szakirodalom. Ezt egészíti ki, váltja fel fokozatosan a török birodalom közép-keleti, illetve a Habsburg birodalom összeurópai, sőt latin-amerikai helyzetének elemzése, mint olyan tényezőké, melyek a végvári vitézségnél talán jobban hatottak a magyar történelemre. De a „kutatás természetén” kívül nyilvánvalóan célszerű megemlíteni, hogy a kutatók növekvő száma, a közelebbi források feltártsága a történettársadalmon belül prezstíznövekedésre vágyó kutatókat szociológiai értelemben is a tágabb összefüggések kutatása, távolabbi adatok beemelésének irányába nyomja.

Külön kérdés, hogy bármely kutatási tárgyul választott konkrét folyamat elemzésénél, megértésénél, megvilágításánál segít az *analóg* folyamatok elemzése, tehát *bármely másik hasonló* folyamattal kapcsolatban előkerülő újabb adat segíthet a vele térbeli, időbeli vagy oksági kapcsolatban *nem lévő* célba vett folyamat bemutatásánál. (Az időnként „interkulturálisnak” nevezett érvelésre a legújabb példa, hogy az Észak-Amerika történet szempontjából tényleg fontos Cahokia-projekt mellett azzal érveltek Illinois-i régészek, hogy eredményeiből más – közel keleti – ókori városok fejlődését lehet majd megérteni...) Nyilvánvaló, hogy az analógiás megközelítés és az adatbővülés kapcsolata sajátos dinamizmust hordoz. A kevés adat alapján analógnak látott – a kutató által egyetlen kategóriába sorolható – folyamatok megértését egy ideig segíti az adatbővülés. Ha viszont a kategóriába tartozó folyamatok egyikéről vagy másikáról váratlanul és esetlegesen nagyon sok adathoz jutunk, akkor felmerül a kérdés, hogy milyen mértékben extrapolálhatók az egyedi tulajdonságok a kategóriába sorolható többi folyamatra.

A kutatásról és tudományos elemzésről szóló fenti mondatok – nyelvi korlátainknál fogva – azt sugallják, mintha „az oksági kapcsolatok felállítása” vagy az „analógiák használata,” illetve ezek mértéke funkcionalista módon elhelyezhető lenne egy helyes-helytelen tengelyen, legalábbis módszertani értelemben. Csakhogy ez közel sincs így. A tudományos termelés természete¹² kétféle értelemben is relativizálja az egy-egy téma kapcsán egy konkrét időpontban „megengedhető” oksági kapcsolatok vagy analógiák mértékét.

Az egyik szempont a tudomány makrotársadalmi integrációjával kapcsolatos. A nemzeti elitek szimbolikus versengése pl. nemzeti ősmítoszok teremtésére ösztönözte (még az 1980- as években is!) a valláskutatókat.¹³ A nacionalizmus korában kuruc dalokat fedeztek fel a történészek, s hogy nehegy kiderüljön, hogy gyengébbek, mint a „labanc” nyugati zene értékei, még hamisítottak is ilyeneket. Thaly Kálmán professzor idevágó történetének a tanulsága¹⁴ nem egyszerűen alkotáslélektani: a lelepleződést követő széleskörű értelmiségi diskurzus mutatja, mekkora lehet a nyomás és milyen erős a mentség a tudományos etikát nyilvánvalóan sértő adatmanipulációra. Az a speciális körülmény, hogy az 1900-as évek iparának különös mértékű pétervári és moszkvai koncentrációja létrehozott egy erős ipari munkásságot és munkásmozgalmat, lehetővé tette azt a – min-

¹² A vikingek szerepe az orosz fejlődés önállóságának megkérdőjelezéséhez jelent municiót Szvák Gyula(2005): A rendszerváltó évtized magyar történeti ruszisztikája. - In: Világtörténet, (27. évf.), tavasz/nyár, pp. 104-111. Lásd még: Szili Sándor(2000): A „kolonizáció” problematikája az orosz és a szovjet historiográfiában. - In: Világtörténet, (22. évf.), tavasz/nyár, pp. 3-22.
Bourdieu, Pierre (2009): A tudományos mező; ford. Fáber Ágoston. In.:Replika 20. évf. 67. sz., pp. 11-36.
Uő: A tudomány tudománya és a reflexivitás : a Collège de France 2000-2001. évi előadás-sorozata / ; [szerk. és a szöveget gond. Lenkei Júlia] ; [ford. Házás Nikolett, Simon Vanda], 2005, Gondolat, Budapest.

¹³ Komoróczy Géza(1988): Tartalmában strukturalista, formájában nemzeti. Sz. A. Tokarev, szerk., Mitológiai enciklopédia. A magyar kiadás szerkesztője: Hoppál Mihály (Budapest: Gondolat), BUKSZ: Budapesti Könyvszemle, 2, No. 2 (1990. nyár), pp. 164-175.

¹⁴ http://epa.oszk.hu/00000/00011/00106/pdf/iskolakultura_EPA00011_2006_07_08_034-040.pdf

den jel szerint „téves” – tudományos elméletet, hogy más gyengén iparosodott országokban is lehetséges munkásmozgalom-történeti jelenségként magyarázni az országos folyamatokat – olyan kategóriákkal, melyeket a 19-20. századi német, angol társadalomra dolgozott ki a történetírás. A munkásmozgalom-történetírás így vált¹⁵ az egykori szocialista blokk egyik fontos közös tudományos projektjévé, erőforrásokat elvonva a bolgár, román vagy éppen albán történetírás talán relevánsabb kérdéseitől. (Csak az olvasó szórakoztatása céljából érdemes felidézni, hogy Werbőczy tettei között a köznemesi megyei hatalmat valóban évszázadokra meghatározó jogtudósi szerep és a második jobbágság kialakulásában valóban fontos politikai-közigazgatási szerepkör mellett, mint azokkal egyenrangú tettet idézték fel történelemkönyveink egy korabeli (tehát 16. századi!) bányásmozgalom letörését, hogy így Werbőczy a munkás-paraszt összefogás elleni úri-világ előképévé válhasson.¹⁶ E kétségtelenül mulatságos gondolatmenet komoly kutatói kapacitásokat is megmozgatott, melyek külföldi – mármint csehszlovákiai – levéltári kutatásokat tettek lehetővé. Így e történet már nem öncélúan kerül a „hogyan keletkezik az új adat” tárgyú alfejezetünkbe...)

A történelmi festészet és szobrászat, a könyvkiadás és a filmgyártás, sőt a turizmus értelmiségi és üzleti érdekkörei egészen konkrét nyomást fejtenek ki a kutatókra, hogy a szigorú történeti vagy régészeti kritika által csak nagyon töredékesen igazolt képeket a fantázia szabadabb használatával egészítsék ki. Az ősmagyarok ruháiból és használati tárgyaiból – ha múzeumban járva csak a saját szemünknek hiszünk – alig ismerünk valamit. Nemcsak a tudományos legitimációval semmiképpen nem rendelkező (de kiemelt kultúrpolitikai támogatást élvező) 1896-os Feszty körkép, de a régészprofesszor László Gyula 1980 körüli (!) rajzai is, mehökkentő „pontossággal” „ábrázolják” „őket.”¹⁷ (Ha az olvasó belegondol, az hogy a három szót külön-külön tesszük idézőjelbe világosan tükrözi: mi minden kérdőjelezhető meg egy efféle rajzban...) Aki Kréta hegyes útjain átvergődik, a déli parton feltárt Phestos-i palota bokáig érő romjaiból láthatja, mi az amiben intellektuális igényességgel „biztosak lehetünk,” azután visszamenekülhet az Evans által – gazdag képzelettel – újjáépített északi Knossos-i palotába,¹⁸ mely eleget tesz élményigényeinek, s ahol borsos díjjal növelheti a görög államkincstár bevételeit, reménykedve, hogy nemcsak a turisztikai ágazatot, hanem a görög tudományt is segíti befizetésével.

Nemzedékek szereztek vizuális képet a „római kocsiversenyről” a Ben Hurból, s a *tudomány népszerűsítést* célul kitűző – és erre például brit közpénzt is igénybe vevő – tévécsatornák is *játékfilmrészletekkel* növelik annak valószínűségét, hogy a néző őket választja, s ne a teljesen fiktív helyszíneken, fiktív időben játszódó fantasy-filmet a másik csatornán.¹⁹

A másik szempont a tudományos termelés természete. A kutatók növekvő száma, a kutatók közötti versengés a totálisan megbízható, de „teljesen érdektelen részletekbe” merülés Szküllája és a mindenképpen újat és „nagyot” mondás Kharübdiszé között mozgatja a kutatókat,²⁰ s számtalan egyéni körülmény mellett az egyetemek vagy a tudomány-

¹⁵ http://real-j.mtak.hu/614/1/TARSFILTORT_22.pdf

¹⁶ <http://mek.oszk.hu/02200/02228/html/01/160.html>

¹⁷ <http://www.tankonyvtar.hu/en/tartalom/historia/80-02/ch15.html>

¹⁸ http://mumia.art.pt.e.hu/periszkop/Musorok/history/ceram_-_a_regeszet_regenye.PDF

¹⁹ <http://www.bbc.co.uk/bbcfilms/catalogue>

²⁰ <http://filozofiaiszemle.net/wp-content/uploads/2015/01/Schwendtner-Tibor-%C3%89rtelmes-megismer%C5%91-v%C3%A1llalkoz%C3%A1s-e-a-filoz%C3%B3fia.pdf>

finanszírozás intézményes viszonyai határozzák meg, hogy mikor melyik stratégia bizonyul sikeresebbnek. Ráadásul a pillanatnyi siker a tudományban sem azonos mindig a „történelmi távú,” azaz tudománytörténeti kézikönyvek által visszaigazolt, s a kutatók Nyugat Európában látványosan növekvő élettartama²¹ miatt egyre inkább még személyesen is megtapasztalható sikerrel.

Új típusú új adatok keletkezése

A hagyományos szöveges vagy tárgyi-képi források mellett a történettudomány olyan új típusú adatokkal is dolgozik, mint az éghajlattörténet,²² vagy népek keveredését leírni képes genetikai vizsgálatok,²³ évezredes testmaradványok MRI vizsgálata stb., melyek una-
lomig ismert gazdaság vagy társadalomtörténeti források, sőt egyéni történetek újraértel-
mezésének nyitnak utat. A folyamatosan fejlődő számítógépes nyelvészeti módszerek²⁴
látszólag jól ismert forrásszövegek újabb és újabb rétegeit nyithatják meg, folyamatosan
fejlődnek a képelemzés módszerei, s azok a növénytani, ásványtani, ill. metallurgiai tech-
nológiák is, amelyek évszázadok óta múzeumokban porosodó tárgyak fizikai és kémiai
tulajdonságai alapján rekonstruálhatnak évezredes kereskedelmi útvonalakat, illetve bi-
zonyíthatják olyan technikai tudások létezését, melyről nincsen sem szöveges forrás sem
képi ábrázolás. A neurológia adatokat termel a nyelvészet, az etológia az őstörténeti ku-
tatások számára.²⁵

Mindez – ha a lelkes hangból tudománysszociológiai elemzőbe váltunk át – nem ke-
vesebbet jelent, minthogy a hagyományos humántudományi források bázisán kanonizá-
lódott állításokat, paradigmákat új kihívások érik, s az ezekkel kapcsolatos viták kifeje-
zik (ill. ha tárgyilag nem is értjük pontosan miről folyik a vita: számunkra tükrözik) az
egyes tudóscsoportok küzdelmét. A tudományos tudástermelés két korábban említett
társadalmi összefüggését – azaz a makrotársadalmi tudás-fogyasztást és az azonos terü-
leten működő kutatók egymás közötti versengését – kiegészíthetjük egy harmadikkal: a
humántudományi kutatók szövetséges-kereső aktivitásával. A 20-21. század befolyásos
természettudományai és élettudományai, sőt műszaki tudományai is részt vesznek a hu-
mántudósok közötti vitákban és küzdelmekben.

Ez a jelenség az egyik fő oka annak, hogy sosincs és sosem lesz olyan állapot, hogy
„nincs újabb adat,” de sosem lesz olyan állapot sem, melyben a különböző érdekű kutatók

²¹ A tudósok várható élettartamának növekedése ugyan elmarad a népesség egészében megfigyelhető növe-
kedéstől, de mégis egyértelmű, hogyha a huszadik század minden békés évtizedére megvizsgáljuk az adott
évtizedben elhunyt neves tudósok megélt éveinek átlagos számát.

²² http://nemzetikönyvtar.blog.hu/2012/03/03/hogyan_kutassunk_klimat_2_resz_historiografia

²³ <http://www.sciencemag.org/content/343/6172/747>

²⁴ Babarczy Anna(2006) : Számítógépes nyelvészet. In.:Látás, nyelv, emlékezet. - Budapest: TypoTEX.- pp.
111-124.

²⁵ Kovács Ákos András(2013): A történetírás mint tudomány . - In: Buksz ., (25. évf.), 2. sz., 140-144. , Z.
Karvalics László(1996): Az információtörténeti gondolkodás ágazatai, kiindulópontjai és kezdetei . - In:
Világtörténet, (18. évf.), ősz/tél, pp. 10-25. Gyáni Gábor(1986): Az „új történetírás” jelensége. - In: Világ-
történet, (8. évf.), 3-4. sz., pp. 3-6. Fogel, R. W.(1986): „Tudományos” és tradicionális történetírás . - In:
Világtörténet, (8. évf.), 3-4. sz., pp.7-39.; Zeldin, T(1986):. Társadalomtörténet és totális történet . - In:
Világtörténet. (8. évf.), 3-4. sz., pp. 40-52.;; Kocka, J.(1986): Elméletek és kvantifikáció a történetírás-
ban. - In: Világtörténet, . (8. évf.), 3-4. sz., pp.53-62.

megegyeznének abban, hogy az újabb adatok relevánsak-e a kérdésről szóló eddigi munkálatok megkérdőjelezéséhez, újíráshoz.

Régi adatok újrahasonosulása

Mindazok a történelmi, szociológiai, demográfiai, közgazdasági, irodalomtudományi, művészettörténeti stb. témák, melyek (Európában) az elmúlt nyolcszáz (!) évre koncentrálnak, s különösen azok, amelyek az alfabetizmus széles körű elterjedése, s főként a napisajtó megjelenése utáni évszázadokra fókuszálnak, folyamatosan növekvő adattömeggel néznek szembe.

Ennek a növekedésnek az egyik eleme a preszcientikus és pozitivisták korszakban kiadott óriási tömegű *eredeti forrás és adatalapú kézikönyv* digitalizálása, ami nem egyszerűen a gyorsabb – és a forrásművek eredeti szerkezetétől független – keresést tette lehetővé, hanem olyan összefüggések feltárulását – metaadattá válását – is, amelyek az eredeti mű szerkesztőinek fejében sem lehettek meg. Nyilvánvaló például, hogy több tízezer koraközépkori magyar, német, cseh, francia területeken született egyaránt latin nyelvű oklevél szófordulatait a szövegkiadások tudós szerzői nem hasonlíthatták szisztematikusan össze még egyetlen kézikönyvön belül sem, nem beszélve a különböző országokban kiadott források áttekinthetlenségéről. Furcsa mód tehát a középkori európai latin nyelvű adattömeg az eredeti – zömmel 19. századi – kiadások bázisán nemzetileg széttagolt adattömegként jelenik meg, míg a digitalizáció után megkezdődhet az „eredeti,” azaz a nemzetállami határoktól független kultúra kutatása.

De a nemzeti nyelvű, írott kézikönyvekre is igaz, hogy a digitalizálás olyan metaadatokat teremt, amelyek sem a szerkesztők, sem a szerzők számára nem álltak rendelkezésre. Például a világ tudományosságában számon tartott egyik leghosszabb terjedelmű kézikönyv, amit egyetlen ember (Színnyei József) írt a Magyar Írók Élete és Munkái 14 vaskos kötete,²⁶ illetve annak még terjedelmesebb, de torzóban maradt folytatása – mely Gulyás Pál nevéhez fűződik. E könyvek szerzői – akik cédulákat természetesen a nevek betűrendjében tárolták – sem lehettek tisztában például az egyes magyarországi települések nevének könyvükben való előfordulásával, vagy egyes szavak és fogalmak felbukkanásával az idézett könyvcímekben. A digitalizálás elemi formája mely szabadszavas keresést tesz lehetővé adatok millióit „teremti meg” e kézikönyvekből, s minimális adatbázisba szervezés, mely például a címszavakat, dokumentumokat egymástól elhatárolja, vagy akár elhatárolás nélkül az egymáshoz fizikailag közeli találatokat tekinti összetartozónak már kombinatív kereséseket tesz lehetővé.

A szöveges „adatbázisok” – alapvető forrásként szolgálnak a gyakoriságokat és szám-szerű összefüggéseket kutató kvantitatív és az egyetlen életpálya vagy jelenség potenciális kapcsolatait felgöngyölítő, egyidejűségeket, párhuzamosságokat mérlegelő klasszikus kvalitatív elemző számára is.

Az adatrobbanás másik eleme, hogy a korábban önmagukban csekély érdekességűnek minősített s ezért *kiadatlanul maradt (bár elméletileg ismert) források* fokozatosan hozzáférhetővé válnak. Azokban az országokban (pl. Angliában), ahol a történelem során keletkezett magániratok, helyi közigazgatási iratok nem semmisültek meg, a történelmi

²⁶ A bio-bibliográfia 1890 és 1914 között füzetenként jelent meg és összesen 29.553 magyar „író” (mármint mindenki, aki írt, köztük mérnökök, tudósok, orvosok stb) élettrajzát tartalmazza.

forráskiadást bőségesen dotáló 19. században is reménytelennek látszott a több millió oldalt kitevő kéziratos irattömeg papíralapú megjelenése. Így a késő-középkori kora-újkori forrásokat kutatók fő tevékenysége a 19-20. században a helyi források feldolgozása, ismertetése és *nem* publikálása volt. Az internet megjelenése óta azonban valóságos társadalmi mozgalommá vált a kéziratos források átírása, gépre vitele.²⁷ Ezek a kéziratos források elméletileg évszázadok óta rendelkezésre állnak, tömegesen használható adat mégiscsak a szabadszavas kereshetőség, illetve adatbázisba-szerveződés révén válik belőlük.

Az adatok exponenciális növekedésének harmadik eleme a *nyomtatott források* tömeges digitalizálása. Ezeknek a forrásoknak a fennmaradása – az iratok fennmaradásának esetlegességével szemben – az európai kultúrkörben az egyes korszakok és területek háborús veszteségeitől többé-kevésbé független. A napilapok, hetilapok folyóiratok, tömeges fogyasztásra szánt kalendáriumok és könyvek digitalizálása egy-egy írni-olvasni tudó emberre évtizedek alatt rázúduló információtömegben való gyors tájékozódást, egyedi esemény-hatás feltárást és szisztematikus tömeges elemzést egyaránt lehetővé tesz.

Míg az előző pontban jelzett példák – az éghajlattörténet vagy a metallurgiai vizsgálódások adatai – *teljesen más típusú* adatokkal egészítik ki meglévő (szöveges vagy tárgyi kultúrára épülő) ismereteinket, s teszik ezzel újraírhatóvá korábban kimerítően megírtnak gondolt témáinkat, az e pontban ismertetett adatszűkekedés a régi típusú (levéltári, könyvészeti) adatok *újfajta hozzáférhetőségén* keresztül fejt ki a hatását.

Ez az adatszűkekedés a korábbi korok adatpublikációihoz képest elképesztő nagyságrendű, de két szempontból mégiscsak a korábbi adatkiadás-adatpublikálás örököse.

Az egyik, hogy *csak* kulturális és tudományos célból keletkezik, tehát a „társadalmi idő” legnagyobb részét uraló gazdasági tevékenység, közigazgatási-közfoglalkozási tevékenység és magánéleti tevékenység nem termel ilyen jellegű adatokat. A levéltárak részei ugyan a közigazgatásnak, de e tevékenységük mégis inkább a tudományos alrendszerekhez köti őket. Nagy profitorientált cégek foglalkoznak régi folyóiratok vagy könyvek digitális kiadásával, de e cégek termékei iránti keresletet mégiscsak a tudományos kulturális alrendszerben kialakult közpénzek, illetve az értelmiség kulturális és tudományos fogyasztása biztosítja. Sok ember hobbi (vagy másképpen fogalmazva: szimbolikus tőkefelhalmozása) a digitalizálás, de az ő összesített időfelhasználásuk eltörpül pl. a chatelezéssel, vagy filmnézéssel töltött össztársadalmi időhöz képest...²⁸

A másik, a korábbi adatpublikációhoz hasonlóvá tévő tulajdonsága ennek az új adatvilágnak, hogy ez az adattömeg a némiképp történeti természetű kutatásokat támogatja. Nem a történettudományt érintő határról beszélünk, hiszen a jogtörténet a jogtudomány része, az irodalomtörténet az irodalomtudományé, a sajtótörténet a médiatudományé, a politikátörténet a politológiáé, a filozófiatörténet a filozófiáé, a gazdaságtörténet a közgazdaságtudományé stb. De lehetetlen észre nem vennünk, hogy *ez a fajta* adatszűkekedés szisztematikusan *kevésbé* módosítja a nem történeti természetű vizsgálódások eredményeit, s ekképpen tudományos erőviszonyait.

Ugyanis minél közelebb érünk a mához, annál inkább vannak olyan szerzői, könyvkiadói vagy kereskedelmi érdekek, melyek a nyomtatott művek digitális újrakiadását akadályozzák. Minél közelebb érünk a mához, annál inkább tiltja az adatvédelem, hogy egyedi ügyeket 100 évvel ezelőtti közigazgatási vagy bírósági iratok mintájára nyilvános-

²⁷ <http://www.leveltaros.hu/sites/default/files/Anglia.pdf>

²⁸ http://epa.oszk.hu/01500/01551/00065/pdf/EPA01551_educatio_13_03_405-412.pdf

ságra hozzunk. Minél inkább aktív egy személy, vagy működő egy vállalat és közintézmény, annál erősebb érdekek szabályozzák, hogy mi kerülhet irattermelésükből, vagy audiovizuális dokumentációjukból az internet nyilvánossága elé.

Azaz, míg a kulturális és tudományos célú tömeges digitalizálás a múltra vonatkozó adatok tömeges újrahasznosítását kifejezetten lehetővé teszi, addig a kortárs szempontú elemzéseket torzíthatja. (Ha például a mai amerikai és angol krimi-irodalom szóhasználatát akarná valaki összehasonlítani az interneten full textben elérhető szövegek alapján, könnyen eshetne abba a hibába, hogy a legnagyobb példányszámban megjelenő és ebben az értelemben legjelentősebb művek helyett épp azokból a periférikus művekből vonna le következtetéseket, melyre olyan kicsi a kereslet, hogy szerzőjük szemében a mindenkinek ingyenesen hozzáférhető digitális kiadás reklámértéke és presztízse felülmúlja a ténylegesen eladott könyvek után remélhető hasznót...)

Az elméletileg korábban is létező tudományos adatokhoz való újfajta hozzáférés többféle szempontból is átrendezi a tudományos erőviszonyokat.

Leértékel olyan tőkefajtaikat, melyek főképp az idősebb nemzedék rendelkezésére állnak. Például, ha valaki többet olvasott és jó volt a memóriája, fel tudta hívni ifjabb kollégája vagy tanítványa figyelmét egy, az ő témáját érintő, másról szóló könyvekben, iratokban megbúvó adatra. Természetesen erre nemzedéktársak között is volt lehetőség, sajátos szimbolikus cserealapként. A gondolatok, vélemények, tanácsok, érdemi tartalmi útmutatások sajátos tudományos piacán korábban sokkal nagyobb mértékben volt jelen az adatismeret. Az elmúlt évszázad tudományos hivatkozásai jelentős részben szövegesen is kitérnek arra, hogy valamely adatot egyik vagy másik kollégájának köszöni a szerző. Az óriási szövegbankokban való tudományos kereshetőség leértékeli ezt az adatismeretet a tudományos közösségben, s ennek mértékében mintegy automatikusan felértékeli az elemző vagy kritikai természetű tudást, s a kollégáknak nyújtott effajta tudásra támaszkodó segítséget is.

A tudományos célra használható adat fizikai hozzáférhetőségének a jelentősége csökken, s ez részben átrendezi az erőviszonyokat. A könyvtárak, levéltárak, múzeumok munkatársainak közösségen belüli súlya – vagy éppen a nagyobb magánkönyvtárral rendelkező kutatók szívességgpotenciálja – egyre kevésbé az adatok fizikai birtoklásán, a hozzáférés fizikai értelemben vett támogatásán vagy akadályozásán alapszik, sokkal inkább a digitalizálendő anyagok körének kijelölésén, illetve az ezzel kapcsolatos metaadat-kreáló tevékenységen. A nagy könyvtáraktól – vagy éppen a kutatót tárgytól – fizikailag távol élő kutató információs hátránya csökken, csökken tehát a hatalmi súlya a néhány évtizeddel ezelőtt még ügödöntő jelentőségű kiküldetési díjak, utazási költségtérítések feletti diszponálásnak. Húsz évvel ezelőtt teljesen lehetetlen volt pl. hogy egy lengyel kistelepülésen élő történelemtanár forrásszintű tanulmányokat folytasson a 19. századi amerikai sajtótörténetről – ma ennek semmilyen *technikai* akadályja nincs.

A technikai értelemben vett hozzáférés kiterjedése ugyanakkor bumeráangként is hat: egy konkrét tudományos teljesítmény nemzeti vagy nemzetközi értelemben vett sikerességének hiányát technikai nehézségekből – az adathozzáférés vidéki és fővárosi, perifériaország-béli vagy centrumország-béli eltérő esélyeiből levezető érvelés legitimitása

csökken. Láthatóbbá és leplezhetetlenné válik egyes személyek, intézmények, egész nemzeti tudományos közösségek érdemi, intellektuális partikularizmusa, provincializmusa.²⁹

A tudományos művek bírálóinak egyre kevésbé „kell” figyelembe venniük s „affirmative action”-ként kalkulálniuk azt a tényt, hogy egy mű a vizsgált jelenséggel azonos vagy távolabbi országban született-e. Ez pedig akár meg is növelheti az erős szelekcióval kiválasztott, a mainstream tudományos beszédmódot gyakorolni képesek versenylőnyét.

További hatás – bár a humántudományok körében valóban csak a legszegényebb országokban élő tudósközösséget sújtja, s nem teremt olyan igazi skálát, mint a természettudományok művelői körében – annak növekvő jelentősége, hogy ki fér hozzá és ki nem az internetes, *de nem ingyenesen hozzáférhető* szövegekhez.³⁰

Megnövekszik annak jelentősége, hogy a radikálisan megnövekedett adattömegben ki képes megfelelő módon keresni, ki tud személyes munkaidejével vagy megfizetett munkatársi munkaidőalappal rendelkezve adatot gyűjteni.

A digitalizálásra vonatkozó döntés, illetve a digitalizált forrásokkal kapcsolatos metaadat-készítésre irányuló döntés átrendezi a kutatók forráspreferenciáját. A disszertáció vagy cikktéma kiválasztásakor természetesen mindig szempont volt, hogy miről van és miről nincs hozzáférhető adat. A nagy könyvtárak, levéltárak közelében élő kutatók számára a kevésbé, vagy senki által fel nem dolgozott folyóiratok vagy iratcsomók csábereje nyilvánvalóan nagyobb volt. A digitalizáció után megnövekedett a szabadon kereshető, metaadatokkal segített, s nem utolsó sorban otthonról, vagy egyetemi irodából hozzáférhető források népszerűsége. Ez különösen abban az időszakban volt érzékelhető, amikor viszonylag kevés klasszikus forrás volt interneten hozzáférhető. A következő néhány évben „szinte minden” a netre kerül, a hozzáférhetőség, mint preferencia-átrendező hatás csökken, nem csökken azonban a metaadatokkal való ellátottság jelentősége. Mivel a metaadat-készítés igen jelentős életmunkaerőt igényel, az e felett diszponálós személyek és szervezetek továbbra is képesek befolyásolni a témaválasztást. Érdeemes megfigyelni pl. a Magyar Elektronikus Könyvtárba bekerülő forrásokat a bekerülés időrendjében, illetve az OSZK főigazgató aláírásával még 2004-ben Hungarológiai Alapkönyvtár Digitalizálási program címen megjelent könyvlistát.³¹ Ezeknek a forrásoknak a kiválasztása – mely több száz, talán több ezer digitalizálással foglalkozó tudományos intézmény vagy tudományos érdeklődésű magánszemély döntéseit tükrözi – az egyes témákkal kapcsolatos szimbolikus univerzumok közötti versengést mutatja.

Az interneten használható – s különösen az ingyenesen használható – kézikönyvek idézési valószínűsége lényegesen megnövekszik az interneten nem használható újabb könyvekkel szemben. A kiadói érdekek alapján számos modern kézikönyv nem, vagy csak igen korlátozott hozzáférhetőséggel kerül a netre (pl. a könyvet papíralapon is megvásárlók, vagy előfizetők számára). Ez növelte annak valószínűségét, hogy a régebbi, szabadon hozzáférhető kézikönyvek, ill. az ingyenes site-okra felhordott kivonatok szolgálnak adatforrássul. A hatás analóg azzal a jelenséggel, ami az 1990 előtti kelet-európai könyvtárhelyzet eredménye volt: a kutatók – különösen a legjelentősebb egyetemi könyvtáraktól távol élők – a devizahiány következtében be nem szerzett korszerűbb kéziköny-

²⁹ Lásd az angol felülreprezentációt, román lengyel alulreprezentációt a nemzeti kvóták nélküli európai tudományos források pályázatnyertesei között. <http://erc.europa.eu/projects-and-results/erc-funded-projects>

³⁰ <http://www.teachingcouncil.ie/research/ebasco.2324.html>

³¹ <http://mek.oszk.hu/hungalap/>

vek helyett régebbit vagy nyilvánvalóan gyengébb meglévőt használtak, idéztek. A digitalizálás jelensége — pontosabban a nem digitalizált művek makacs jelenléte — tehát felveti a korszerűtlen információk nagyobb terjedésének valószínűségét is. Az internet történetének egyik első nagy kézikönyv-digitalizációs programjának — az 1911-es Encyclopedia Britannicának — felületére belépve az olvasó azonnal figyelmeztetést kap, hogy a könyv adatai — elsősorban a természettudományos, orvosi, műszaki szócikkek — súlyosan elavultak, de hosszú évekig nem volt új Britannica a neten... Magyarországon ahol hosszú ideig csak a Pallas Lexikon volt elérhető, személyes tapasztalataim szerint a szakdolgozatokban gyakran köszöntek vissza a Pallas jellegzetes adatai, megközelítésmódjai, sőt stílusfordulatai.

Az új adatok keletkezése, a régi adatok újrahasznosulása még tovább bonyolítja azokat az erőviszonyokat, amelyek az egyes témákat kutató történészek, a lezárt dolgokat kutató nem történész identitású társadalom és bölcsészettudományi kutatók, a napjainkban is élő dolgok előzményeit és gyökereit bemutatni kívánó kutatók, a múlt társadalmi jelenségeit analógiaként és példatárként használó közvéleményformáló értelmiségiek között fennállnak. Ezek az erőviszonyok ugyanis természetesen kihatnak a forrásfeltárási, forráspublikálási folyamatokra, a tudományos művek megszületésére és elterjedésére — végső soron a tudomány egészére.

Az *adatfeltárás, adatrendezés és adatelemzés* jelenségvilága tehát egy komplex tudomány-szociológiai összefüggésrendszer része.

Funkcionalistán szólva: minden elemzés legnagyobb kihívása, hogy egyszerre kellene

- 1.) elemeznünk — úgy ahogyan vannak — a rendelkezésre álló adatokat;
- 2.) megállapítanunk, hogy a rendelkezésre álló adatok mennyire reprezentálják, mennyire tükrözik annak a társadalomrésznek a viszonyait, amelyről szólni akarunk, azaz van-e olyan szisztematikus körülmény, ami pl. a „mintába kerülést,” a történelmi adat fennmaradását akadályozta stb.;
- 3.) megállapítanunk (pl analógiák, vagy hasonló kutatások alapján) hogy a rendelkezésre álló adatok, változók közül nem hiányzik e valamilyen alapvető adat, melynek hiánya az összes magyarázatot irreálissá teszi;
- 4.) megállapítanunk, hogy az adat-termelődés közegének nyelvi valósága az adatok mértékének mekkora homogenitását biztosítja, azaz, hogy minden adatszolgáltató számára többé kevésbé ugyanazt jelentik-e az adatszolgáltatás során használt szavak;
- 5.) mennyiben módosítják adatainkat a másoktól átvett vagy magunktól kitalált elemzési kategóriák....

Konfliktusos paradigmában beszélve: az önreflexió és álcázás sajátos dinamikáját fenntartva szólnunk arról, hogy témaválasztásunk mennyire szabad és mennyiben megrendelt, adataink mennyire valóságok és mennyire konstruáltak, hogy kategóriáink mennyire rugalmasak és mennyire előre programozottak, hogy hipotézisünk, illetve tanulságaink mennyire voltak meg előre és mennyire keletkeztek az elemzés eredményeképp, hogy doxikusan használt oksáfgörmünk valójában mely tudományfilozófiai rendszer részeként legitim, illetve illegitim...

Egyszerre reflektáljuk és álcázzuk, hogy mennyiben vagyunk a tudományos gépezet fogaskerekei és mennyiben szabadon lebegő értelmiségiek.

IRODALOM

- BABARCZY A. (2006): Számítógépes nyelvészet. *Látás, nyelv, emlékezet*. TypoTEX, Budapest. pp. 111-124.
- BOURDIEU, P. (2005): *A tudomány tudománya és a reflexivitás: a Collège de France 2000-2001. évi előadás-sorozata* (szerk. és a szöveget gond. Lenkei Júlia) (Ford. Házas Nikolett, Simon Vanda) Gondolat, Budapest.
- BORRDIÉU, P. (2009): A tudományos mező. (Ford. Fáber Ágoston) *Replika* vol.20. No.67., pp. 11-36.
- FOGEL, R. W. (1986): „Tudományos” és tradicionális történetírás. *Világtörténet*, vol.8. No. 3-4. , pp. 7-39.
- GYÁNI G. (1986): Az „új történetírás” jelensége. *Világtörténet*, vol.8. , No.3-4. , pp. 3-6.
- KOCKA, J. (1986): Elméletek és kvantifikáció a történetírásban. *Világtörténet*, vol.8. No. 3-4. , pp. 53-62.
- KOMORÓCZY G. (1990): Tartalmában strukturalista, formájában nemzeti”. Sz. A. Tokarev, szerk., *Mitológiai enciklopédia*. (A magyar kiadás szerkesztője: Hoppál Mihály Budapest, Gondolat, 1988), BUKSZ:
- Budapesti Könyvszemle, 2, No. 2 (1990. nyár), pp. 164-175.
- KOMORÓCZY G. (1998): *Kiáltó szó a pusztában – A Holt-tengeri tekercek*, , Osiris Kiadó, Budapest.
- KOVÁCS Á. A. (2013): A történetírás mint tudomány. *Buksz*, vol. 25., No. 2. , pp. 140-144.
- SZABÓ M. (1987): Nyugatnémet vita az újabbkori német történelemről. *Világtörténet*, vol.9. No.4., pp. 131-142.
- SZILI S. (2000): A „kolonizáció” problematikája az orosz és a szovjet historiográfiában. *Világtörténet*, vol.22. , tavasz/nyár, pp. 3-22.
- SZVÁK GY. (2005): A rendszerváltó évtized magyar történeti ruszisztikája. *Világtörténet*, vol.27. , tavasz/nyár, pp. 104-111.
- Z. KARVALICS L. (1996): Az információtörténeti gondolkodás ágazatai, kiindulópontjai és kezdetei. *Világtörténet*, vol.18. , ősz/tél, pp. 10-25.
- ZELDIN, T. (1986): Társadalomtörténet és totális történet. *Világtörténet*, vol.8. , No.3-4. , pp. 40-52.

Adatok és kezelésük a hivatalos statisztikában¹

A statisztikai törvény² (Statvtv.) első paragrafusa szerint a „statisztika feladata és célja, hogy valósághű, tárgyilagos képet adjon a társadalom, a gazdaság, a tulajdonviszonyok, a környezet állapotáról és változásairól az államhatalmi és a közigazgatási szervek, valamint a társadalom szervezetei és tagjai számára.” Az elmúlt másfél évszázadban, mióta ez a tudományág önállósodott, ezen kívül számtalan más definíciót is megfogalmaztak, de minden meghatározás lényege, hogy a statisztika az adatgyűjtés és adatfeldolgozás általános tudománya. A statisztikai munka valójában a valóságnál kezdődik, azt képezi le adattá valamilyen definíció mentén, és a folyamat végén a közölt adat a felhasználónál válik információvá.

Az adat nagyon tág fogalom. A KSH módszertani dokumentációjának fogalomtárára (KSH, 2014a) szerint „az adat az információ formalizált módon való megjelenítése, amely alkalmas feldolgozásra, továbbításra, közlésre, értelmezésre.” Egy adatból akkor lesz információ, azaz objektumokra (tényekre, eseményekre, dolgokra, folyamatokra vagy a gondolati világ elemeire) vonatkozó ismeret, ha definiált, azaz pontosan meghatározott, hogy milyen objektum, melyik ismérévre vonatkozik, és milyen nomenklatúra, jegyzék szerint kell értelmezni. A statisztikai adat a „valós világ egyedeinek tulajdonságaira vonatkozó statisztikai megfigyelések, illetve további statisztikai műveletek eredménye” (KSH, 2014), vagyis annyiban specifikus, amennyiben előállítását és értelmezését a statisztikai módszertanok szabják meg.

Az adatok egy speciális fajtája az ún. metaadat, mely más adatokat ír le, illetve határoz meg. A statisztikai metaadatok tartalmazzák a statisztikai adatrendszerben használt fogalmakat, nomenklatúrákat, a felhasznált adatforrások leírását, az adatelőállítás módszertanát. A metaadatok alapján lehet eldönteni, hogy az adatok megfelelnek-e a nemzetközi és hazai standardoknak, kritériumoknak.

¹ Köszönetnyilvánítás: Ezúton szeretnék köszönetet mondani kollegáimnak, dr. Szép Katalinnak, dr. Nagy Eszternek és Nagy Beátának, hogy szakértelmükkel, észrevételeikkel és javaslataikkal hozzájárultak a tanulmány elkészítéséhez.

² 1993. évi XLVI. törvény a statisztikáról.

Jelen tanulmányban a szerző áttekinti a hivatalos statisztikai szolgálat (HSSZ)³ által használt adatforrás-típusokat, felhasználási területeiket és módjukat.

Az adatforrások típusai

Adatforrásnak nevezzük azt az adathalmazt, metaadat-halmazt, adatbázist vagy metaadat gyűjteményt, amiben a szükséges adatok elérhetők (OECD, 2008).

Az ENSZ 5. statisztikai alapelve kimondja, hogy „statisztikai célokra bármilyen adatforrásból – statisztikai adatfelvételekből vagy adminisztratív nyilvántartásokból – származó adatok felhasználhatók. A statisztikai szervezeteknek a forrás kiválasztásánál figyelembe kell venniük az adatok minőségét, időszerűségét, a költségeket és az adatszolgáltatókra háruló terheket.” (ENSZ, 2014)

Statisztikai adatigények megfogalmazását követően manapság, amikor a statisztikai munkában egyre nagyobb hangsúlyt kap a költséghatékonyság és az adatszolgáltatói – különösképpen a vállalkozásokra nehezedő – terhek csökkentése, első lépésként azt kell megvizsgálni, hogy létezik-e már olyan adatforrás, amelyből a szükséges információk átvethetők. Ennek hiányában indokolt csak új statisztikai adatokat gyűjteni.

Egy statisztikai tevékenységet végző intézmény szemszögéből tekintve attól függően, hogy mely szervezet végezte a kapcsolódó adatgyűjtést, az adatforrások rendszere két részre osztható:

- az elsődleges adatforrások és
- a másodlagos adatforrások csoportjára.

Elsődleges adatforrások

Az elsődleges adatforrások kategóriájába a HSSZ minden egyes szervénél a saját statisztikai adatgyűjtései tartoznak. A hivatalos statisztikai rendszeren belül az adatgyűjtések történhetnek kikérdezéssel (pl. kérdőívek alkalmazása) vagy megfigyeléssel (pl. árak összeírása). Abból a szempontból, hogy az adatgyűjtés a megfigyelendő statisztikai egységek milyen körét öleli föl, lehet teljes vagy részleges. Teljes lefedettség esetén a sokaság minden egységét számba veszik, míg részleges lefedettség esetén azoknak csak egy részéről gyűjtenek információt. A részleges megfigyelés egy fontos típusa a mintavételen alapuló, reprezentatív megfigyelés. Teljeskörű adatgyűjtésre a legismertebb példa a népszámlálás, az állatszámlálások és számos intézményi adatszolgáltatás. Sok esetben azonban, amikor nem lehetséges vagy nem célszerű a sokaság valamennyi elemét felmérni, a sokaságnak csak egy részéből vesznek mintát. Néha a a kétféle adatgyűjtési módot kombinálják, ha a sokaság egyes elemeit teljeskörűen felmérik (pl. bizonyos mérethatár felett az összes vállalkozást), a sokaság más elemeiből (pl. a kisebb vállalkozásokból) pedig mintát vesznek.

³ A hivatalos statisztikának formális definíciója nem létezik. Jelen tanulmányban ide értjük a statisztikai törvény hatálya alá tartozó tevékenységet. A statisztikai törvény 3.§ (1) szerint „E törvényben meghatározott statisztikai tevékenység ellátása a hivatalos statisztikai szolgálat feladata.”

Statisztikai célú adatgyűjtések rendszere

Magyarországon az adatszolgáltatási kötelezettséggel járó statisztikai adatgyűjtéseket – kivéve a törvényben elrendelteket – a hivatalos statisztikai szolgálat országos statisztikai adatgyűjtési programja (OSAP) tartalmazza. Az OSAP keretében gyűjtendő adatok körét a HSSZ bejelentett adatgyűjtési igényei alapján a KSH évente állítja össze az Országos Statisztikai Tanács (OST) bevonásával. A Sattv.-ben kapott felhatalmazás alapján a kormány az éves OSAP-ot rendelet⁴ keretében jelenteti meg. Az OSAP hatálya a jogi személyekre, a saját joga szerint jogképes szervezetekre, valamint a gazdasági tevékenységet folytató természetes személyekre terjed ki. A 2015-ös OSAP összesen 362 tétel⁵ tartalmaz.

Az Infotv⁶ 5. §-a szerint személyes adat akkor kezelhető, ha ahhoz az érintett hozzájárul, vagy ha azt törvény vagy törvény felhatalmazása alapján helyi önkormányzat rendelete kötelező érvénnyel elrendeli. Ennek megfelelően a népmozgalmi⁷ és a halottvizsgálattal kapcsolatos adatgyűjtéseket nem az OSAP, hanem a Sattv. rendeli el, míg a népszámlálásokat és az általános mezőgazdasági összeírásokat esetileg elfogadott törvények.⁸ A statisztikai adatgyűjtési program részeként minden egyéb, személyes adatra vonatkozó adatszolgáltatás önkéntes. A KSH honlapján és a Hivatalos Értesítőben közzé teszi az általa tervezett önkéntes adatgyűjtéseket. 2015-re tíz ilyen tétel szerepel a listán, melyek munkaerőre, háztartási költségvetésre és életkörülményre, magyar állampolgárságot szerzőkre, utazási szokásokra, infokommunikációs eszközök használatára, bevándorlókra és külföldiek magyarországi kiadásaira vonatkoznak.

Másodlagos adatforrások

Másodlagos adatforrásnak tekintünk minden olyan adatállományt, amely esetében az adatgyűjtő személy, szervezet nem azonos azzal, amelyik elemzi vagy használja. Ez különbözteti meg az elsődleges adatoktól, amelyek esetében a két funkció megvalósítója azonos. Egy adott, statisztikai tevékenységet végző szervezet szempontjából másodlagos adatforrások közé sorolandó valamennyi, más szervezettől átvett adatgyűjtemény függetlenül attól, hogy eredetileg mi volt az adatgyűjtés célja. A másodlagos adatforrásokon belül három nagy csoport különíthető el:

- HSSZ-en belüli (OSAP-ban rögzített) statisztikai adatgyűjtések,
- adminisztratív adatok,

4 288/2009. (XII. 15.) Korm. rendelet az Országos Statisztikai Adatgyűjtési Program adatgyűjtéseiről és adatátvételéről. A rendeleter évente módosítják, és a mellékletei tartalmazzák a hivatalos statisztikai szolgálat szerveinek aktuálisan érvényes közvetlen adatgyűjtéseit és statisztikai célú adatátvételeit, valamint a gyűjthető adatköröket.

5 Egy tétel alatt egy közvetlen statisztikai adatgyűjtést vagy egy, nem statisztikai célra gyűjtött adatállomány átvételét értjük.

6 2011. évi CXII. törvény az információs önrendelkezési jogról és az információszabadságról.

7 Születés, haláleset, házasságkötés, bejegyzett élettársi kapcsolat létesítése, élettársi nyilatkozat nyilván tartásba vétele, házasság felbontása, valamint érvénytelenné nyilvánítása, bejegyzett élettársi kapcsolat felbontása, megszüntetése, a nyilvántartott élettársi kapcsolat megszűnése, valamint lakcímváltozás tartozik a népmozgalmi események közé.

8 Pl. 2009. évi CXXXIX. törvény a 2011. évi népszámlálásról; 2010. évi XXIV. törvény a 2010. évi általános mezőgazdasági összeírásról.

- egyéb másodlagos adatforrások,
 - vállalkozások, által gyűjtött adatok,
 - HSSZ-en kívüli statisztikai adatgyűjtések,
 - big data és
 - bármi egyéb.

HSSZ-en belüli (OSAP-ban rögzített) statisztikai adatgyűjtések

A Stattu-ben megfogalmazott célok elérése érdekében a KSH nemcsak összeállítja az OSAP-ot, hanem egyúttal megszervezi a HSSZ-en belül az OSAP keretében gyűjtött adatok átadását, átvételét is. Ezzel elősegíti a statisztikai céllal gyűjtött adatok hatékony felhasználását. A 2015-ös OSAP-ban szereplő 362 tételből 293 eredményeit nemcsak az elrendelő szervezet hasznosítja, hanem a HSSZ más szerve is átveszi. A két legnépszerűbb adatgyűjtés a KSH 1845. Éves gazdaságstatisztikai jelentése és a 2235. Havi teljesítménystatisztikai jelentése. Ezeket 4-4 más intézmény veszi át. Az aktuális adatátvételei igények a KSH honlapján olvashatók.⁹

Meg kell azonban jegyezni, hogy a HSSZ-en belül csak korlátozottan léteznek olyan közös szabványok, kritériumok vagy ajánlások, amiket a statisztikai munkafolyamatoknak és termékeknek ki kell elégíteni, sőt a HSSZ tagszervezeteinél független statisztikai szervezeti egység is csak elvétve létezik. Ezért az OSAP-adatgyűjtések sem tartalmilag, sem módszertanilag nem képeznek koherens egységet. Másodlagos felhasználásuk során a teljeskörű statisztikai adatgyűjtéseket ugyanúgy kell kezelni, mint egy nemstatisztikai adatállományt.

Adminisztratív adatok

A legelterjedtebb meghatározás szerint az adminisztratív adatforrás közfeladatot ellátó szervezet által fenntartott adatgyűjtemény, amit azért gyűjtenek és használnak, hogy eleget tegyenek a közterhekkel, (nyilvántartási, engedélyezési, jogosultsági) eljárásokkal, szolgáltatásokkal, juttatásokkal kapcsolatos feladataiknak. Az adatokat tehát eredetileg jog és kötelezettség keletkeztetésével, az egyénnel szemben valamilyen joghatás kiváltásával összefüggésben gyűjtötték. Az adminisztratív adatok jellegzetessége, hogy gyűjtésük célja a célcsoport teljes lefedettsége, minden egyes egyed megfigyelése, továbbá, hogy a későbbi statisztikai hasznosítás ellenére az adatgyűjtés és a feldolgozás módszerét alapvetően továbbra is az adatgazda szervezet határozza meg (*Brackstone, G. J., 1987*). Az adminisztratív adatok gyűjtését és karbantartását mindig jogszabály rendeli el, ami részleteiben írja le a gyűjtendő adatok körét, a nyilvántartás által használt fogalomrendszert és sok esetben az alkalmazott osztályozási rendszereket is.

Az adminisztratív adat lényege az adat egyedi azonosíthatósága, hisz az adott egyedre származtathatnak következményeket. A statisztikai adat esetén viszont az egyedi azonosítás lehetősége csak eszköz az adatgyűjtés és feldolgozás során, mert a sokaságot kívánjuk jellemezni.

A statisztikai felhasználás szempontjából az adminisztratív adatok erősségét és gyengeségét mutatja az 1. tábla összehasonlítva a népszámlálással és az egyéb statisztikai adatgyűjtésekkel.

⁹ A hivatalos statisztikai szolgálat szerveinek egymás közötti 2015. évi adatátvételei http://www.ksh.hu/docs/hun/info/adatgyujtes/2015/HSSZ_egymuk_adatavetelei.pdf

1. táblázat: A népszámlálás, egyéb statisztikai adatgyűjtések és a népességre vonatkozó adminisztratív adatok mint statisztikai adatforrások összehasonlítása

Szempont		Népszámlálás	Egyéb statisztikai adatgyűjtés	Adminisztratív adatok
1.	Lefedettség	A népesség teljes lefedettségét célozza.	Bizonyos népességcsoportokat kizárhat.	A célpopulációt a közigazgatás igényei határozzák meg.
2.	Tartalom	Az adatok széles köre kiterjedt keresztesztályozást tesz lehetővé.	Rendszerint csak egy szűk témakörre szorítkozik, de sokkal részletesebben, mint egy népszámlálás.	A közigazgatási cél szempontjából igényelt adatokra szorítkozik.
3.	Fogalmak, definíciók	A társadalmi, gazdasági elemzések elvárásaira alapulnak.		A közigazgatás igényei határozzák meg.
4.	Kisterületi becslések	A teljes lefedettség eredményeként előállnak.	Általában nem állíthatók elő.	Előállíthatók, ha az egyedi rekordok a kisterületek szintjén beazonosíthatók.
5.	Minőség-ellenőrzés	A hibák minimalizálása érdekében megtervezhető.	A kisebb méret a népszámlálásánál szigorúbb ellenőrzést tesz lehetővé.	A közigazgatási hatóság végzi, és lehet, hogy a kulcsváltozók kivételével nem fordít rá elég figyelmet.
6.	Költség	Drága.	Egy adatgyűjtés viszonylag olcsó, de a két népszámlálás közötti időszak kumulált költsége nagy lehet.	Viszonylag olcsó, ha a kezdeti adatfeltöltés költsége a statisztikát nem érinti.
7.	Gyakoriság	5 - 10 év.	Témától függően lehet éves, negyedéves, havi.	A közigazgatási feladattól függően az éves és a napi többszöri gyakoriság között lehet, sőt akár folyamatos is.
8.	Időszerűség	Néhány hónaptól több év témától és részletzettségtől függően.	Rendszeres adatgyűjtés esetén néhány hét, ad-hoc adatgyűjtés esetén néhány hónap.	A közigazgatási eljárástól függ. Az éves adatok átvételre ellenőrzött formában gyakran csak sok hónap után van lehetőség.
9.	Stabilitás	A változások azoknak a statisztikusoknak a felelőssége, akik a felhasználói igények teljesítéséért is felelnek.	Ismétlődő adatgyűjtés esetén ritka a változás, hogy az időbeli összehasonlíthatóság biztosított legyen.	Jogi, szabályozási vagy eljárási módosítások következtében történhet változás.
10.	Adatszolgáltatói terhelés	Súlyos, de ritka.	Általában kicsi, de a mintába választottak számára súlyos.	Nincs további terhelés.

Forrás: Brackstone G J: Issues in the Use of Administrative Records for Statistical Purposes. Survey Methodology, June 1987 Vol. 13, No. 1, pp. 28-43 Statistics Canada

A két kulcsfontosságú előny, vagyis az adminisztratív adatokra épülő rendszer fenntartásának költséghatékonysága és az adatszolgáltatói terhek csökkentése mellett még továbbiakat is érdemes megemlíteni. Tekintve, hogy az igazgatási adatnak mindig naprakésznek kell lenni, ezért ezek alapján gyakrabban lehet a statisztikai mutatókat előállítani, ami különösen jelentős az olyan komplex és költséges, ezért csak ritkán végrehajtható adatgyűjtések tekintetében, mint a népszámlálás. A rövid távú mutatók esetében is

számottevő előny adódik abból, hogy sok idő nyerhető az igazgatási adatok használatával (ENSZ, 2011).

Egyéb másodlagos adatforrások

Vállalkozások és kutató helyek adatai

Az ENSZ 5. statisztikai alapelve azt jelenti, hogy a statisztika szempontjából a fentiekén kívül akár vállalkozások által gyűjtött adatok és kutató helyek statisztikai adatgyűjtéseinek eredményei is szóba jöhetnek. Az előbbi esetében megoszlik a nemzetközi gyakorlat az adatok elérhetősége szempontjából. A szlovén statisztikai törvény¹⁰ például úgy rendelkezik, hogy a hivatal és az egyéb hivatalos adat-előállítók felhasználhatják az állami és a magánszektor különböző hivatalos és egyéb adminisztratív adatgyűjtéseik (feljegyzések, nyilvántartások, adatbázisok stb.) egyedi azonosításra alkalmas adatait, ha azokat törvény alapján vagy az egyén írásbeli hozzájárulásával vezették. Észtország statisztikai törvénye¹¹ is általános felhatalmazást ad a magán adatforrásokhoz is. Ezeket az adatforrásokat azokban az esetekben lehet statisztikai célokra hatékonyan felhasználni, amikor – hasonlóan az adminisztratív adatokhoz – vagy általános nyilvántartásként egyetlen szervezet vezeti, vagy több adatgazda esetén jogszabály rendeli el a nyilvántartás tartalmát és a használt fogalmakat, definíciókat és osztályozási rendszereket. Magánvállalkozások által vezetett, statisztikai szempontból felhasználható nyilvántartások a hitel ügynökségek, üzleti elemzők, közművek adata, telefonkönyvek, telefonszolgáltatók, hűségkártya rendszert működtető kiskereskedők stb.

Kutatóintézetek és egyéb kutató helyek adatai (pl. szociológiai felmérések, epidemiológiai felmérések, közvélemény kutatások) ellenőrizhető és megfelelő minőség esetén a statisztikai szervezetek adatforrásai között szintén felhasználhatók, ha a jogszabályi feltételek adottak, illetve az adatvédelmi szabályok nem sérülnek.

Big data

A statisztikai hasznosítási potenciálokat tekintve nem lehet kikerülni, hogy megemlítjük itt az utóbbi idők legnagyobb kihívását, a naponta keletkező, óriási nagyságú digitális adattömeget, a „big data”-t. Az elnevezés arra utal, hogy az egyes adatállományok mérete terabájtban, sőt petabájtban mérhető, és ezért az általánosan használt szoftver eszközök nem képesek elfogadható időn belül átvenni és feldolgozni őket. A big data hatékony kezelése különleges technológiát igényel.

Ez a típusú adat egyrészt a hagyományos üzleti tevékenységek nyomán keletkezik, másrészt olyan újfajta adatforrásokban, mint például a közösségi médiumok. Tipikus big data az elárusítóhelyeken felszerelt terminálok (POS) és a banki ATM-ek által generált adatok, de a Google-keresések, blog- és Facebook-bejegyzések, lájkok, mobiltelefonok GPS-koordinátái és a YouTube videók is.

A big data túlnyomóan az emberi viselkedés nyomait követi figyelemmel, és az adatok összekapcsolásával és elemzésével a különböző viselkedések előre jelezhetők, ezért főként

¹⁰ National Statistics Act. <http://www.stat.si/dokument/5186/NationalStatisticsAct.pdf>

¹¹ Official Statistics Act. <https://www.riigiteataja.ee/en/eli/ee/506012015002/consolide/current>

társadalomkutatásra használható. Meg kell azonban jegyezni, hogy a big data-ból természeti jelenségekre vonatkozó információk is nyerhetők, pl. úrfelvételek, meteorológiai álmások mérései vagy egyéb műszeres megfigyelések.

Az adminisztratív adat mint statisztikai adatforrás

Nemzetközi jogszabályok, ajánlások

A skandináv országokban az adminisztratív adatok rendszerszerű felhasználása a statisztikában egészen a 60-as és 70-es évek fordulójáig vezethető vissza (Statistics Finland 2004). A nemzetközi szervezetek (ENSZ, Eurostat) az 1990-es évek közepén kezdtek azzal foglalkozni, hogy hogyan lehetne az adminisztratív adatok statisztikai felhasználását koordinálni. Ezeknek az erőfeszítéseknek az első jeleként értékelhető, hogy az ENSZ-alapelveknek az 1994-ben megjelent első verziója már úgy említi az adminisztratív adatokat, mint a statisztika lehetséges adatforrásait. A közösségi statisztikákról szóló 322/97/EK rendelet és a helyébe lépő európai statisztikákról szóló 223/2009/EK európai parlamenti és tanácsi rendelet egyformán szabályozta a közigazgatási nyilvántartásokhoz való hozzáférést. Bár a jogszabály úgy rendelkezett, hogy a tagországok érintett szervezeteinek olyan mértékben kell hozzáférniük a közigazgatási rendszerükön belül az adminisztratív adatforrásokhoz, amilyen mértékben ezek az adatok az európai statisztikák fejlesztéséhez, előállításához és terjesztéséhez szükségesek. Ezzel együtt a helyi szabályozást a tagállamok hatáskörébe utalta.

A rendelet 2015. áprilisi módosításában az egyik legfontosabb változás az adminisztratív nyilvántartásokhoz való statisztikai célú hozzáférést érintette. A hatályos jogszabály szerint az érintett intézmények jogosultak a közigazgatási nyilvántartásokhoz való azonnali és ingyenes hozzáférésre és azok felhasználására. Adatok átadásakor csatolni kell a metaadatokat is. Kötelezővé teszi a statisztikai szervek bevonását a nyilvántartások kialakításába, fejlesztésébe és megszüntetésébe, valamint az európai statisztikák előállítása szempontjából lényeges közigazgatási nyilvántartásokra vonatkozó standardizálási tevékenységekbe. A statisztikai hivataloknak és a közigazgatási nyilvántartások tulajdonosainak ki kell alakítaniuk a szükséges együttműködési mechanizmusokat.

Az Európai Bizottság 2005-ben ajánlasként jelentetett meg egy szabályzatot „Az európai statisztika gyakorlati kódexe” (ESR, 2011) címmel. A gyakorlati kódex célja az európai statisztikák iránti általános bizalom biztosítása azáltal, hogy megállapítja az európai statisztikák fejlesztésének, előállításának és terjesztésének módját a statisztikai rendeletben meghatározott statisztikai elvekkel és a statisztika legjobb nemzetközi gyakorlataival összhangban. A kódexet 2011-ben megújították, és az új változatban hangsúlyossá vált az adatgazdák és a statisztikai szervezetek közötti együttműködés az igazgatási nyilvántartások statisztikai célú felhasználhatósága érdekében.

A gyakorlati kódex 15 alapelvét 80 ismervben fogalmazták meg részletesen, és az adminisztratív adatforrásokra vonatkozó 9 ismerv a 2., 8., 9. és 10. elvénél jelenik meg az alábbiak szerint.

2. alapelv: Felhatalmazás adatgyűjtésre.

2.2. ismerv: A statisztikai szervezetek jogszabályi felhatalmazással rendelkeznek az adminisztratív nyilvántartások adataihoz való statisztikai célú hozzáférésre.

8. alapelv: Megfelelő statisztikai eljárások.

8.1. ismerv: Ahol az európai statisztikák adminisztratív nyilvántartáson alapulnak, ott az adminisztratív célú alkalmazások definíciói és fogalmai jól közelítik a statisztikai célú fogalmakat.

8.7. ismerv: Az adminisztratív adatok statisztikai céloknak való jobb megfelelése érdekében a statisztikai szervezetek részt vesznek az adminisztratív adatok tervezésében.

8.8. ismerv: A statisztikai szervezetek az adminisztratív adatok tulajdonosaival megállapodásokat kötnek, amelyek tartalmazzák a felek közös elköteleződését az adminisztratív adatok statisztikai célra való használatára.

8.9. ismerv: Az adatminőség biztosítása érdekében a statisztikai szervezetek együttműködnek az adminisztratív adatok tulajdonosaival.

9. alapelv: Az adatszolgáltatói teher nem túlzott mértékű.

9.4. ismerv: Az információgyűjtés megkettőzésének elkerülése érdekében, amennyiben lehetséges, adminisztratív forrásokat használnak.

9.5. ismerv: A többszörös információgyűjtés elkerülése érdekében a statisztikai szervezetek között általánosan elfogadott az adatmegosztás.

9.6. ismerv: Az adatszolgáltatói teher csökkentése érdekében a statisztikai intézmények olyan lépéseket szorgalmaznak, amelyek lehetővé teszik az adatforrások összekapcsolását.

10. alapelv: Költséghatékonyság.

10.3. ismerv: Előremutató lépések történnek az adminisztratív nyilvántartások használhatóságának javítására és a közvetlen adatfelvételeket szolgáló források korlátozására.

Hazai tapasztalatok

A hazai folyamatokat jellemzi, hogy a Stattu. a KSH feladatai között már az 1993. évi elfogadásától felsorolta ugyan az adminisztratív adatgyűjtések fogalmi és osztályozási rendszerének kialakításában való közreműködést,¹² de a hivatal jogosultsága az azonosításra alkalmas, nem saját gyűjtésből származó adatok felhasználására¹³ csak az 1999-es módosítással került be a jogszabályba.

A nemzetközi tendenciákkal összhangban az utóbbi másfél évtizedben hazai szintén is felerősödött az adminisztratív adatok rendszerszerű, statisztikai célú felhasználása. Ezeknek az adatforrásoknak az előtérbe kerülését elsősorban az informatikai fejlődés segítette elő, hiszen kiszorultak a papíron vezetett nyilvántartások az elektronikus rendszerek javára, ami megkönnyítette az adatok mozgatását, kezelését.

A hazai jogrendszer hozzáigazítása a módosított európai statisztikai rendelethez a közeli jövő feladata. Az adminisztratív adatok statisztikai felhasználása szempontjából elsősorban a nyilvántartásokat elrendelő jogszabályokban kell feloldani a statisztikai felhasználás korlátait.

Az OSAP-tételeknek közel háromnegyede közvetlen statisztikai adatgyűjtés, a többi ún. statisztikai célú adatátvétel, amiknek a döntő része adminisztratív adat átvételét jelenti. Ebből azonban nem az következik, hogy a hivatalos statisztikai rendszerben összesen használt adminisztratív adatkörök száma a 100-at sem éri el, mert sok esetben az

¹² 6.§ (1)g) pontja.

¹³ 21. § (4) bek.

átadás-átvétel jogi kerete nem az OSAP, hanem egy szerződés vagy megállapodás. Pl. a KSH a nem-statisztikai adatátvételeinek csak 45%-át teszi közzé az OSAP-ban.

Az adminisztratív források minőségének rendszerszerű megközelítése azt igényli, hogy beazonosítsuk és nyilvántartásba vegyük őket. Jelentősen hátráltatja az adminisztratív nyilvántartások statisztikai célú felhasználását, ha ezeknek az adatrendszereknek nincs egy regisztere, ami információt szolgáltat az igazgatásban fellelhető adatok köréről, azok tulajdonságairól.

A nyilvántartások regisztere azonban nemcsak a statisztikai hasznosítás szempontjából fontos, hanem az állami és önkormányzati nyilvántartások együttműködése tekintetében is. Ezért a 2015. január 1-jén hatályba lépett interoperabilitás törvény¹⁴ 2016 januárjára létrehozta a nyilvántartások elektronikus információs rendszerét. A végrehajtásról szóló kormányrendelet¹⁵ szerint a regiszter tartalmazza a nyilvántartás és a nyilvántartó azonosító adatait, a nyilvántartás vezetéséről és adattartalmáról rendelkező jogszabályra hivatkozást, a nyilvántartott adatok megnevezését és a nyilvántartás együttműködésének módját a többi nyilvántartással.

A törvény egyúttal létrehozta az adatmegnevezések jegyzékét is, amely a megnevezéseken kívül tartalmazza a hozzá tartozó értelmezés leírását is. Így a közigazgatás szintjén fokozatosan felépül egyfajta metarendszer.

Mindezek megteremtik a lehetőséget a nyilvántartások monitorozására és rendszeres értékelésére az egyes adatállományoknak a statisztikai adat-előállításban betölthető értéke szempontjából. Ha a tartalmilag alkalmas adatok minősége nem megfelelő, a statisztikusok és az adatgazda közösen kialakíthatják a nyilvántartás fejlesztési koncepcióját azért, hogy az a statisztikai felhasználásra alkalmas legyen.

Az adminisztratív adatok felhasználási területei a statisztika előállításának folyamatában

Adminisztratív adatok gyakorlatilag az összes statisztikai munkafázisban alkalmazhatók.

- ♦ Statisztikai regiszter, felvételi keret

A statisztikai regiszter egy adott sokaság teljeskörű nyilvántartása, mely az egyes egységek azonosító és elérhetőségi adatait, valamint a felhasználástól függően egyéb adatokat is tartalmaz. A regiszterek karbantartása főként adminisztratív adatforrásokon alapul. A statisztikai regiszter egyik fő felhasználási módja, hogy mintavételi keretként szolgál a statisztikai adatgyűjtésekhez. Sok esetben az adminisztratív nyilvántartás már önmagában is alkalmas erre a célra.

- ♦ Statisztikai adatgyűjtések teljes vagy részleges kiváltása

Az adminisztratív nyilvántartásoknak ez a felhasználási módja eredményezi a legtöbb erőforrás megtakarítását, hiszen a legköltségesebb munkafázist, az adatgyűjtést egyszerűsíti vagy rövidíti le, illetve teszi szükségtelenné. Ha adminisztratív adatokkal váltunk ki statisztikai adatgyűjtést, arra is lehet számítani, hogy az elsődleges adatgyűjtő az adatokat bizonyos mértékig már ellenőrizte, tehát az adatelőkészítési munkafázis is olcsóbb

¹⁴ 2013. évi CCXX. törvény az állami és önkormányzati nyilvántartások együttműködésének általános szabályairól.

¹⁵ 142/2015. (VI. 12.) Korm. rendelet az állami és önkormányzati nyilvántartások együttműködésének általános szabályairól szóló 2013. évi CCXX. törvény végrehajtásáról.

lesz. Azt, hogy az adatgyűjtés helyén milyen ellenőrzések és javítások történtek meg, az adatgazdával történő együttműködés keretében lehet megismerni.

- **Editálás és imputálás**

Mind az adminisztratív, mind pedig a statisztikai adatgyűjtési folyamatban történhet olyan hiba, ami rontja az adatok minőségét, vagy adathiányt okoz. Ezek szükség szerinti javítása, pótlása adminisztratív adatállományokból származó kiegészítő információk bevonásával történhet.

- **Közvetett használat a becslésekben**

Az adminisztratív adatok további felhasználási lehetősége a becslések finomítása, rendelkezésre álló adatok súlyozása és az eredmények érvényességének ellenőrzése, validálása.

- **Tájékoztatás kiegészítése, gazdagítása**

Az adminisztratív adatgazda gyakran maga is számol statisztikai mutatókat, amiket a statisztikai szervezet az alapadatok ismerete nélkül a saját tájékoztatási rendszere bővítésére vesz át.

A közvetlen statisztikai adatgyűjtés és a másodlagos adatok felhasználásának feltételei¹⁶

Közvetlen statisztikai adatgyűjtések

A statisztikai adatgyűjtés szabályait a Stattv. tartalmazza. Statisztikai adatgyűjtéseket adatszolgáltatási kötelezettség előírásával, vagy önkéntes adatszolgáltatás alapján lehet végrehajtani. Amint a fentiekben szó volt róla, adatszolgáltatási kötelezettséget az adatszolgáltató típusától függően törvényben vagy az OSAP-ban lehet előírni.

Másodlagos adatok felhasználása

A másodlagos adatforrásokat eredetileg vagy nem statisztikai célra hozták létre, vagy legalábbis nem az újrahasznosító szervezet céljaira. Ezért az újrafelhasználásnak speciális feltételei vannak.

Jogalap

A jogszabályoknak biztosítani kell, hogy a statisztikai szervezet hozzáférjen a más szervezetek által gyűjtött, megfelelő részletezettségű adatokhoz. A statisztikai adatok esetében a Stattv. szerint¹⁷ a hivatalos statisztikai szolgálathoz tartozó szerv az OSAP végrehajtásából rendelkezésre álló adatokat egyedi adat kivételével a HSSZ másik szervének kérésére köteles továbbítani. Hozzá kell azonban tenni, hogy a KSH egyedi adatot is átvehet.

A Stattv. ugyanitt¹⁸ felhatalmazást ad a KSH-nak egyedi adminisztratív adatok statisztikai célú átvételére, „ha törvény eltérően nem rendelkezik.” Ez utóbbi kitétel miatt

¹⁶ A fejezet tartalmilag támaszkodik a szerző szerkesztésében a „Másodlagos adatforrások használata a statisztikában” címmel készülő KSH kiadvány kéziratára.

¹⁷ 21. §

¹⁸ A Stattv.-nek ez a rendelkezése összhangban van az Inforv. 13. § (1) bekezdésével.

az adminisztratív adatok szükséges mértékű eléréséhez a módosított európai statisztikai rendelet biztosítja ugyan a jogalapot, de a jogforrásként megjelenő törvényeket módosítani kell ahhoz, hogy megszűnjenek az európai statisztikai rendelettel jelenleg fennálló ellentmondások, és az információk egyedi azonosításra alkalmas módon, statisztikai felhasználás céljából átadhatók legyenek.

Az egyéb adatforrások tekintetében, ha azok személyes adatot tartalmaznak, az Infotv. korábban idézett 5. §-a a mérvadó, más esetben a Polgári Törvénykönyv magán, üzleti és egyéb titkokra vonatkozó rendelkezéseit kell figyelembe venni.

A közvélemény támogatása

A statisztikai munkában arra törekszünk, hogy a másodlagos adatok egyre nagyobb körre szolgálhasson adatforrásként, amihez a megfelelő jogi szabályozás mellett a közvélemény támogatása is szükséges. Tudatosítani kell a szakmai és civil adattulajdonosokkal, hogy a másodlagos adatforrások statisztikai célú felhasználásával elkerülhetők az átfedő adatgyűjtések, ami költséghatékony és csökkenti az adatszolgáltatási terheket. Ezenközben biztosítani kell a közvéleményt, hogy a statisztikai újrahaznosítás nem jelent megnövekedett adatvédelmi kockázatot.

Egységes azonosító kódrendszerek, adat-összekapcsolás

Mint erről már korábban is volt szó, a másodlagos adatforrások eredeti célja többnyire nem azonos azzal a statisztikai céllal, amire felhasználják, ezért gyakori, hogy egyetlen adatforrás nem biztosítja a szükséges információkat, hanem több adatforrást kell rekord szinten összekapcsolni, amihez elengedhetetlen a célpopuláció elemeinek egyértelmű azonosítása. Egy egységes azonosító rendszer megkönnyíti az adatállomány lefedettségének, teljességének vizsgálatát is. Statisztika szempontból egy ideális helyzetben a nyilvántartások egységes azonosító rendszereket használnak. Ha nincs ilyen, gyakran rendkívül nehéz és fáradságos, olykor lehetetlen hatékony módon összekapcsolni a különböző nyilvántartásokhoz tartozó adatokat.

A statisztikai alkalmazásban alapvetően három fő azonosító rendszer szükségességét szokták kiemelni: a személyeket azonosító kód, az épület-, lakásazonosító kód és a gazdasági szervezeteket azonosító kód, mert ezek a népszámlálási adatok előállításához nélkülözhetetlenek. Valójában azonban minden statisztikai egységnek (település, megye, gépjármű, termék stb.) kell rendelkezni ilyenekkel. Ezek közül a gazdasági szervezetek egységes azonosítására a törzsszám használható, ami összekapcsolja a különböző államigazgatási szervek adminisztratív nyilvántartásait. Az épületek és lakások esetében a legkézenfekvőbb a pontos cím, amit a 2015. január 1-jével létrejött címregiszter biztosít. A személyek azonosítása már lényegesen bonyolultabb, hiszen mint ismeretes, ebben a vonatkozásban több általános kódrendszer létezik (személyi szám, társadalombiztosítási azonosító jel, adóazonosító jel), de a különböző rendszerek elemei nem feleltethetők meg egymásnak, vagyis a különböző állományok rekordjai leggyakrabban a természetes azonosítók (név, nem, születési hely és idő stb.) mentén kapcsolhatók csak össze, amik írása, helyesírása nem mindig következetes.

Átfogó és megbízható igazgatási nyilvántartási rendszerek

Az igazgatási nyilvántartások, adatbázisok országos vagy regionális szinten hasonló tulajdonságokkal rendelkeznek, teljes lefedettségűek, továbbá módszertanilag és minősé-

gileg is általában egységes állományok. Az adminisztratív adatok természetük szerint többnyire pontosak és a célnak megfelelően naprakészek. A statisztikai célú újrahasznosítás éppen ezeket a tulajdonságokat használja ki, mert az ilyen típusú adatforrásokra lehet hatékony módon statisztikai rendszert építeni.

A gazdálkodó szervezetek által az adminisztratív és a statisztikai célra szolgáltatott adatok egyre kevésbé térnek el egymástól, hiszen a nagyobb szervezetek adatszolgáltatása mindkét irányban a könyvelési szoftverrel történik. Jelentősebb eltérésre a lakossági adatok esetén lehet számítani, amiket számos, a nyilatkozás pillanatában fennálló szubjektív tényező is befolyásolhat.

A közigazgatási hatóságok együttműködése

A közigazgatási szervek közös erőfeszítése egy olyan statisztikai rendszer kiépítésére, amely a lehető legnagyobb mértékben használja fel forrásként a nem-statisztikai céllal gyűjtött adatokat, csak akkor lehet valóban eredményes, ha azt egy kormányzati szintű határozott és egyértelmű kötelezettségvállalás is támogat, valamint szoros együttműködés van az érintett hatóságok között. Különösképpen a másodlagos adatforrásokra épülő átfogó, nagy statisztikai adatrendszerek kiépítéséhez (pl. népszámlálás) van szükség arra, hogy a statisztikai hivatal megfelelő alkupozícióval rendelkezzen a többi hatósággal szemben.

Az elmúlt nyolc évben két kormányhatározat¹⁹ szorgalmazta az adminisztratív adatok statisztikai célú hasznosítását. Ezek ráirányították a HSSZ tagszervezeteinek figyelmét az adminisztratív adatok jelentőségére a statisztikai munkában, bővítették a témához kapcsolódó közös ismereteiket, növelték a szakmai egyetértésük mértékét, és hozzájárultak az adatgazda és a felhasználó szervezetek közötti párbeszéd elindulásához, illetve felerősödéséhez.

Statisztikai tájékoztatás és adatvédelem

Stattv. által az adatok publikálására vonatkozóan megfogalmazott alapszabály szerint a hivatalos statisztikai szolgálathoz tartozó szervek által végrehajtott adatgyűjtések eredményei alapvetően nyilvánosak. Nem lehet azonban nyilvánosságra hozni minősített adatot, valamint az adatszolgáltatóval kapcsolatba hozható adatot (egyedi adat).

A statisztikai adatokat hagyományosan statikus táblázatok formájában publikálják. Ennél korszerűbb mód, amikor a felhasználó egy adatkör változóiból önállóan válogatja össze a tábla dimenzióit. Manapság azonban egyre inkább fokozódik a kutatók igénye az ún. mikroadatok – a közvetlen azonosítóitól megfosztott egyedi rekordok – iránt. Míg a táblázatos megjelenítés esetében az adatkiadás előtt lehet biztosítani a felfedés elleni védelmet, a mikroadatokhoz való hozzáférés már bonyolultabb adminisztrációs eljárást tesz szükségessé, illetve biztonságos környezetben előállított kutatási eredmények utólagos adatvédelmi szempontú ellenőrzését.

A különböző felfedés elleni védelmi módszerek eltérő hatást gyakorolhatnak a védendő adatállományokra, az adatok minőségére. A felfedés elleni nagyon szigorú eljárások csökkentik az adatok pontosságát, részletezettségét. A megfelelő módszer megválasztá-

¹⁹ 2212/2007. (XI.14.) kormányhatározat az államigazgatási nyilvántartások adatainak statisztikai célra történő felhasználásáról (hatályon kívül helyezve); 1133/2011. (V. 2.) kormányhatározat a vállalkozói adminisztratív terhek csökkentésére irányuló rövid- és középtávú kormányzati programról.

sakor arra kell törekedni, hogy az felfedési kockázat mértéke és az adatminőségben bekövetkező veszteség egyensúlyban maradjon.²⁰

IRODALOM

- KSH (2014): Módszertani dokumentáció/Fogalmak, definíciók. http://www.ksh.hu/apps/meta.menu?p_lang=HU&p_menu_id=220&p_param=S&p_session_id=82019363
- OECD (2008) OECD Glossary of Statistical Terms. <https://stats.oecd.org/glossary/about.asp>
- ENSZ (2014): ENSZ-alapelvek a hivatalos statisztikára http://www.ksh.hu/nemzetkozi_ajanlasok
- BRACKSTONE, G. J. (1987): Statistical Issues of Administrative Data: Issues and Challenges. In: Statistical Uses of Administrative Data - An International Symposium. Statistics Canada, 23-25 November 1987 (Proceedings published by Statistics Canada, Ottawa, December 1988).
- ENSZ (2011): Using Administrative and Secondary Sources for Official Statistics: A Handbook of Principles and Practices. UNITED NATIONS. New York and Geneva. <http://unstats.un.org/unsd/EconStatKB/Attachment442.aspx>
- STATISTICS FINLAND(2004): Use of Registers and Administrative Data Sources for Statistical Purposes Best Practices of Statistics Finland http://www.stat.fi/tup/julkaisut/kasikirjoja_45_en.pdf
- ESR (2011): Az Európai Statisztika Gyakorlati Kódexe. Az Európai Statisztikai Rendszer Bizottsága által 2011. szeptember 28-án elfogadott változat https://www.ksh.hu/az_europai_statisztika_gyakorlati_kodexe

²⁰ A KSH az adataihoz különböző hozzáférési módokat biztosít, hogy a különböző felhasználói igényeket kielégítse pl. előre megadott szerkezetű táblázat (STADAT), felhasználó által definiált tábla (tájékoztatási adatbázis és egyéni kérések), nyilvános mikroadat, szerződéssel kiadott anonimizált mikroadat, továbbá kutatószoba, távoli hozzáférés, illetve távoli végrehajtás. A kutatószoba és a távoli hozzáférés esetén a KSH biztonságos környezetben ad hozzáférést kutatás céljából előkészített mikroadat állományokhoz kutatók részére. Ezekben az esetekben mikroadat állományt nem, csak a belőlük származó kutatási eredményt kapja meg a kutató a kutatási eredmények utólagos adatvédelmi szempontú ellenőrzését követően.

Közzadatok és nyilvános adatbázisok: a hozzáférés kérdései

Közhelyszerű megállapítás, hogy a közinformációkhoz való hozzáférés a demokratikus jogállamok alkotmányos alapjoga, a demokrácia egyik alappillére, az uralkodó politikai rendszerek és szövetségeseik elszámoltathatóságának sarkalatos eszköze, számos további információs jog érvényesíthetőségének előfeltétele, a nyilvánosság záloga, a tájékozott döntéshozatal megalapozója, a közügyekben való állampolgári részvétel serkentője, a szabad média táptalaja – és még sorolhatnánk. Talán legemelkedettebb, és egyben szakszerű elnevezése az *információs szabadság*, angol elnevezéssel Freedom of Information (FOI), gyakran így, nagybetűvel írva. Röviden azt jelenti, hogy mindenki szabadon hozzáférhet a közzsféra információihoz, és ezt a jogát csak szűk körben, törvény korlátozhatja.

Mégis, ennek az alapértéknek és számos országban tételes jognak¹ az érvényesíthetősége folyamatos politikai, üzleti és civil érdekütközések színtere, jogviták tárgya, gyakorlati problémák felvetője. Ennek egyrészt az az oka, hogy változnak a közügyek menedzselésének és információgazdálkodásának keretei, és vele együtt változnak a potenciális felhasználók igényei és információs szükségletei, így az információs szabadság nem tekinthető egy kívánatos *állapotnak*, amit el lehet érni és meg lehet őrizni, sokkal inkább egy *folyamatnak*, amelynek mindig követnie kell a változásokat, vagy akár elébe kell mennie a várható változásoknak. Közelebből, oka ennek a helyzetnek az információs szabadság gyakorlati érvényesíthetőségének kereteit meghatározó politikusok hozzáállása is, amit David Flaherty, Brit Columbia (Kanada) korábbi adatvédelmi és információs szabadság-biztosa² egyszer így fogalmazott meg, igen találóan: a politikusok egyenesen imádják az információs szabadságot, *mielőtt és miután* hatalmon vannak. Nem utolsó sorban oka pe-

¹ 2014. szeptemberében Paraguay a világ századik országaként iktatta törvénybe az információs szabadság jogát, és azóta is csatlakoztak újabb országok a közinformációk hozzáférhetőségének jogát garantáló országok táborához. Megjegyzendő, hogy nemcsak hagyományos demokráciák vagy autoritárius politikai rendszereiktől nemrég megszabadult rendszerváltó országok vannak ebben a táborban, hanem fejlődő országok is, amelyek a korrupció visszaszorítását, politikai és társadalmi rendszerük demokratizálását remélik a jóléti országok információs jogának implementálásától.

² Flaherty 1997-ben, amikor még Magyarország a rendszerváltó országok élvonalának számított az információs jogok elismerése és érvényesítése terén, itteni látogatása után követendő példaként állította Magyarországot a kanadai törvényhozók elé, magyarul lásd Majtényi (1998:424-425.), <http://abi.atlatszo.hu/index201.php?menu=beszamolok/1997/F/6>

dig az információs és kommunikációs technológiák fejlődése, amely önmagában is átalakítja a közinformációk keletkezésének, rögzítésének, elemzésének és felhasználásának közegét, az egyre inkább információ-alapú hatalomgyakorlás erőviszonyait, és egyúttal szétfeszíti a korábbi információs viszonyokra alkotott fogalomrendszer és jog kereteit. Ma egy számítógépes adatállomány vagy „adatbázis”³ esetében nem mindig könnyű feladat meghatározni a közadat-hozzáférési jog alapján, hogy voltaképpen mi is a hozzáférés konkrét tárgya, és mire vonatkozik a hozzáférés joga.

Közinformáció, közadat

A „közinformáció” kifejezés a laikusok számára azt sugallhatja, hogy valamiféle közös jószágról, mindenki által szabadon használható és művelhető területről van szó. A digitális bennszülöttek generációja számára az internet maga is ilyen közös jószágnak tűnik a rajta fellelhető információkkal együtt (és eközben sokukban nem tudatosul, hogy a hálózati világ információhálózata amúgy kőkemény üzleti modelleken alapul, amely mellesleg lehetőséget ad társadalmilag előnyös nonprofit tevékenységek végzésére is). Az információs társadalomban vagy tudástársadalomban – ahogy a jelenkor teoretikusai szívesen jellemzik a posztindusztriális korszakot – különösen népszerűek az offline világ közös használatú javainak metaforáit alkalmazó kifejezések, mint a „szellemi közjavak”,⁴ a „digitális közjavak” vagy a „creative commons.”⁵ Hozzátehetjük: vannak olyan, a társadalom alapvető értékeit kifejező közjavak is, amelyek nem csupán közös használatú és használati értékű, nem-vertékelődő „jószágok”, amelyek használatából senki sem zárható ki (mint amilyen például a friss levegő vagy a nemzetvédelem), hanem maga a „közjó” (public goods). Ha pedig kiterjesztjük a közjó értelmét a szabad, demokratikus társadalmak olyan alapértékeire, amelyek e társadalmak elidegeníthetetlen lényegét alkotják – ahogy Loader és Walker (2007, és másutt) javasolja – a *constitutive public goods* vagy „alkotó közjó” fogalmához jutunk.⁶ Az információs szabadság szószólói hajlamosak a közinformációkhoz való szabad hozzáférést is ilyen alkotó közjónak tekinteni a nyugati demokratikus társadalmakban.

Az azonban a laikusok számára is nyilvánvaló, hogy a nyilvánosság ténye nem elegendő kritérium a közinformáció vagy közadat minősítéshez: nyilvánossá válhatnak magántulajdonú adatok, személyes adatok, jogellenesen nyilvánosságra hozott adatok, vagy a sokat vitatott szerzői és szomszédos jogok által védett adatok is – ezek nyilvánossága nem minősíthető a közjó részének.

³ Felhasználói szemszögből adatbázison strukturált adatok összetartozó együttesét értjük, amelyeket különféle szempontok szerint lehet kiválogatni, összefüggéseiket feltárni, és amelyek állománya jellemzően folyamatosan változó: elavult vagy hibás adatokat törölnek, javítanak, újbabbakat felvesznek – következésképpen egy adatbázisnak meghatározhatatlanul nagy számú állapota van. Az adatbázisnak nemcsak numerikus adatok, nevek, címek vagy tárgyszavak, hanem akár teljes dokumentumok, képek, hang- és videófájlok is képezhetik az alapelemeit.

⁴ A szellemi közjavak, vagy rokon értelemben a digitális közjavak legismertebb példái a nyílt forráskódú szoftverek, vagy a közösségi előállítású tartalmak, mint a Wikipedia, lásd Ferge (2007), Rátai (2008).

⁵ A „kreatív közjavak” vagy korábbi, kissé földhözragadt fordításban „kreatív közlegelők” koncepció lényege, hogy a szerzők szabályozott formában és meghatározott kategóriákban eleve lemondhatnak jogaik egy részéről, és ezzel felszabadíthatják alkotásaik felhasználhatóságát (<http://creativecommons.org/>). Magyarország 2005-ben csatlakozott hivatalosan a Creative Commons rendszerhez.

⁶ Reziliencia-tárgyú tanulmányukban lásd erről Raab, Jones és Székely (2015) okfejtését.

Mi tehát akkor a közinformáció, vagy – amennyiben rögzített, strukturált és visszakereshető információról van szó – a közadat? A magyar jog 1992 óta⁷ azokat az adatokat, amelyek kezelésének főszabálya a nyilvánosság és bárki számára való hozzáférhetőség, *közérdekű adatoknak* hívja. A jelenleg hatályos törvény⁸ szerint közérdekű adat a „közfeladatot ellátó szerv vagy személy kezelésében lévő és tevékenységére vonatkozó vagy közfeladatának ellátásával összefüggésben keletkezett, a személyes adat fogalma alá nem eső, bármilyen módon vagy formában rögzített információ vagy ismeret.” Az adatok másik – a fenti definícióban is hivatkozott – alapkategóriája, amelyet ugyancsak az 1992-es adatvédelmi törvény definiált először, a *személyes adat*. Lényegét tekintve ez nem más, mint minden, az érintett személlyel kapcsolatba hozható adat, valamint az adatból levonható, az érintettre vonatkozó következtetés.

A közérdekű adatok kezelésének főszabálya a nyilvánosság, a személyes adatoké pedig az önrendelkezés. A két főszabály jól kifejezi az állam-állampolgár, általánosságban a magán-köz dichotómiát, közelebbről a demokratikus jogállamok azon alapkövetelményét, hogy az állam legyen átlátható és elszámoltatható (kivéve, ha valamely speciális törvény indokolja a titkosságot), az állampolgár pedig legyen átlátszatlan, kivéve, ha ő maga teszi magát részben átláthatóvá (vagy valamely speciális törvény indokolja az átláthatóvá tételét). Ezt az utóbbi jogot – vagyis azt, hogy alapvetően mindenki maga dönthet az adatai sorsáról – nevezi a magyar jog, német mintára, információs önrendelkezésnek. Noha a két alapkategória dichotómiája ma is érvényes, számos más kivétel és egymást részben átfedő speciális információ- és adattípus létezik, amelyeket a jog különböző ágai próbálnak több-kevesebb sikerrel összehangolni, például statisztikai adat, „nyilvános adat,” nevesített titokfajták (üzleti, orvosi, adó- stb. titok), környezeti adat stb.⁹

Más okokból is finomítani kellett a kétpólusú személyes adat/közérdekű adat koncepciót. Az korábban is világos volt, hogy a közfeladatot ellátók vagy közszereplők e tevékenységével összefüggő adatok, ugyan definíció szerint személyes adatok, mégis kezelésükre nem az önrendelkezés, hanem a nyilvánosság főszabálya vonatkozik. Ezt erősítendő, a jog létrehozta a *közérdekű nyilvános adat* kategóriáját. Eszerint közérdekű nyilvános adat a közérdekű adat fogalma alá nem tartozó minden olyan adat, amelynek nyilvánosságra hozatalát, megismerhetőségét vagy hozzáférhetővé tételét törvény közérdekből elrendeli.¹⁰ Amint látható, ez a kategória nemcsak személyes adatokra, hanem bármilyen más adatra is vonatkoztatható. A nemzeti adatvagyon körébe tartozó állami nyilvántartások fokozottabb védelméről szóló törvény pedig a „nemzeti adatvagyon” kategóriáról rendelkezik, ami a közfeladatot ellátó szervek által kezelt közérdekű adatok, személyes adatok és közérdekből nyilvános adatok összessége.¹¹

A jognak szabályoznia kellett a korszerű adatállományok, adatbázisok hozzáférhetőségének kérdését is. Az adatbázis, mint fogalom már az 1980-as évek végén előfordult a magyar jogban, és elterjedése a rendszerváltás utáni évek jogalkotásában sem a nyilván-

⁷ Ekkor született a személyes adatok védelméről és a közérdekű adatok nyilvánosságáról szóló 1992. évi LXIII. törvény, röviden „adatvédelmi törvény”, Avtv. A rövid elnevezése sokakat megrévesztett, akik úgy vélték, hogy Magyarországon nincs is információszabadság-törvény.

⁸ Az információs önrendelkezésről és az információszabadságról szól 2011. évi CXII. törvény (Infotv.), amely a korábbi törvény hatályon kívül helyezésével vált irányadóvá.

⁹ Összefüggéseikről lásd bővebben Majtényi (2006).

¹⁰ Infotv. 3. § 6. pontja.

¹¹ 2010. évi CLVII. törvény 1. § 1. pontja.

nos hozzáférés szabályaival volt összefüggésben (lásd statisztikai törvény,¹² nyilvántartási törvény,¹³ döntéselőkészítési adatokról szóló törvény¹⁴). Az adatbázisok nyilvános elérhetőségéről először a közművelődési törvény¹⁵ és a közbeszerzési törvény¹⁶ rendelkezései voltak jelentősek, de az igazi áttörést két nemzetközi jogi dokumentum implementálása, illetve egy hazai kidolgozású törvény elfogadása jelentette. Az első csoportba tartozik a környezeti ügyekben az információhoz való hozzáférésről, a nyilvánosságnak a döntéshozatalban történő részvételéről és az igazságszolgáltatáshoz való jog biztosításáról szóló, ún. Aarhus-i Egyezmény kihirdetése, amely előírja: „Valamennyi Fél biztosítja azt, hogy a környezeti információk egyre inkább rendelkezésre álljanak olyan elektronikus adatbázisokban, amelyek a nyilvánosság számára a nyilvános telekommunikációs hálózaton keresztül könnyen hozzáférhetők,” valamint hogy a rendszer „egy strukturált, számítógépes és a nyilvánosság számára hozzáférhető adatbázison alapul, mely adatbázist egy egyszerűsített beszámolási rendszer táplál.”¹⁷ Még jelentősebb e tekintetben a közadatok újrahasznosításáról szóló törvény (Közzadat tv.), amely az Európai Unió tagállamai számára kötelezően implementálandó irányelv magyar megfelelője.¹⁸ Ehhez egyrészt meg kellett teremteni a „közzadat” definícióját („az információk önrendelkezési jogról és az információszabadságról szóló törvényben meghatározott közérdekű adat és közérdekből nyilvános adat”¹⁹); másrészt pedig ki kellett terjeszteni a hozzáférés szabályait a korszerű adatállományokra is: eszerint a közzadat újrahasznosítás céljából történő rendelkezésre bocsátása „a közzadathoz az igénylő részére biztosított olyan hozzáférés, amely lehetővé teszi az igényelt közzadat újrahasznosítását az igénylő számára, ideértve különösen a közzadat adathordozón vagy elektronikus úton történő egyszeri vagy rendszeres átadását, a közzadatot tartalmazó adatbázishoz történő közvetlen hozzáférés biztosítását...”²⁰

A legjelentősebb fejlemény a jogalkotási vonulatban a 2005-ben hatályba lépett elektronikus információszabadság törvény (Eitv.),²¹ amelynek előkészítésében, koncepciójának kidolgozásában széleskörű szakmai egyeztetés zajlott, nem-kormányzati szereplők érdemi bevonásával.²² Ugyan a törvény azóta beolvadt az Infotv.-be és így megszűnt önálló jogszabályként funkcionálni, legfontosabb rendelkezései ma is érvényesek.²³ Három fő területen rendelkezik a közinformációk megismerését elősegítő, korábban nem létező eszközök és szolgáltatások létrehozásáról: a közérdekű adatok kötelező internetes közzététele (beleértve a közérdekű adatok központi elektronikus jegyzéke és az egységes közzadatkereső rendszer kiépítését), a jogszabályok és a jogszabályalkotás nyilvánossága (beleértve a hatályos jogszabályok elektronikus gyűjteményének létrehozását) és a

¹² 1993. évi XLVI. törvény a statisztikáról.

¹³ 1992. évi LXVI. törvény a polgárok személyi adatainak és lakcímének nyilvántartásáról.

¹⁴ 2007. évi CI. törvény a döntéselőkészítéshez szükséges adatok hozzáférhetőségének biztosításáról.

¹⁵ 1997. évi CXL. törvény a muzeális intézményekről, a nyilvános könyvtári ellátásról és a közművelődésről.

¹⁶ 2011. évi CVIII. törvény a közbeszerzésekről.

¹⁷ 2001. évi LXXXI. törvény 5. Cikk 3. és 9. pontjai.

¹⁸ 2012. évi LXIII. törvény a közzadatok újrahasznosításáról.

¹⁹ Közzadat tv. 4. § 2. pontja. – A törvény 2015. júliusi módosítása külön definiálja a „kulturális közzadat” fogalmát is, amely átfedésben van több létező adatkategóriával is.

²⁰ Közzadat tv. 4. § 5. pontja.

²¹ 2005. évi XC. törvény az elektronikus információszabadságról.

²² Részletesebben lásd erről Majtényi et al. (2005).

²³ Lásd az Infotv. 33–37/B. paragrafusait és 1. mellékletét.

bírószági határozatok nyilvánossága (beleértve a bírósági határozatok elektronikus gyűjteményének felállítását). Noha az Eitv. mindhárom fő rendelkezése adatbázis felállítását és internetes elérhetőségét írja elő,²⁴ voltaképpen célja az egyedi adatok, illetve dokumentumok internetes hozzáférhetőségének elősegítése az adatbázisok felállítása és működtetése révén. (Ezzel egyébként voltaképpen megsérti a technológiasemlegesség elvét, de ezt azzal magyarázza, hogy jelenleg az internetes honlapok használata olyan sajátos és kiemelt módja a polgárok tájékozódásának, hogy ez indokolja külön szabályozását.²⁵)

Az Eitv. hatályba lépése mégis a vártnál kisebb érdeklődést eredményezett, elsősorban a sajtó részéről,²⁶ azonban az évek során a jogszabálykereső, a bírósági határozatok tára, sőt az egységes közadatkereső is, minden korlátja ellenére, elfoglalta helyét az érdeklődő polgárok és szakemberek információszerzési spektrumában.

Ehhez képest a szerzői jogi törvény (Szjt.)²⁷ más oldalról szabályozza az adatbázisok felhasználhatóságát, és ebben ugyancsak egy európai uniós irányelv²⁸ implementálása tükröződik. Az „adatbázis-jog” – fontos hangsúlyozni – nem az adatbázisokban tárolt és rendszerezett adatok és dokumentumok felhasználhatóságát szabályozza, nem is az adatbázis összeállítójának, mint gyűjteményes mű létrehozójának szerzői jogait védi, hanem az adatbázisok készítőinek ad jogokat, amivel elismeri „a homlok verítékét,” vagyis azt a ráfordítást, amit a készítők az adatbázis kifejlesztésére használtak. Az adatbázis egészét (beleértve a működtető logikát) vagy tartalmának jelentős részét lemásolni és újrahasznosítani csak az előállító engedélyével és feltételeivel lehet. Az újrahasznosítás ismert esete az ún. deep linking, ami azt jelenti, hogy egy internetes portál egyik ablakában egy másik weboldalon nyilvánosságra hozott anyag (például szerkesztett hírek vagy egy adatbázis) jelenik meg, azt a látszatot keltve, hogy az a portál üzemeltetőjének terméke. Ezt a gyakorlatot egyébként a Magyarországi Tartalomszolgáltatók Egyesületének etikai kódexe kifejezetten tiltja (MTE, 2009:5.6 pont)

Látható, hogy a jog koncepciója és fogalomrendszere az iratközpontú megközelítéstől elmozdult az adatközpontú megközelítés felé,²⁹ bár sok szempont indokolja az iratközpontú szabályozás fenntartását is. Az is látható, hogy számos jogszabály rendelkezik megengedő vagy korlátozó módon az információk hozzáférhetőségéről. Ehhez képest megtevesztő, hogy a laikus felhasználó szemszögéből az interneten „minden” fent van és gyakorlatilag szabadon elérhető – sőt, a sokat vitatott online szerzői jogok területén kifejezetten etikátlan stratégiának minősíthetjük azt, hogy egyfelől ingyenes eszközöket kínálunk a tartalmak könnyű letöltésére, megosztására, remixelésére, majd pedig elítéljük a

²⁴ A jogszabályokat a magyarorszag.hu kormányportál jogszabálykeresőjén, illetve közvetlenül a Nemzeti Jogszabálytár honlapján (www.njt.hu) lehet lekérdezni, a bírósági határozatokat a bírósági portál szolgáltatásaként (<http://birosag.hu/ugyfelkapcsolati-portal/anonim-hatarozatok-tara>), a közérdekű adatokat pedig az egységes közadatkereső honlapján (www.kozadat.hu).

²⁵ A technológiasemlegesség részbeni fenntartását célozza ugyanakkor az, hogy közérdekűadat-kérést nem lehet elutasítani azzal az indokkal, hogy a kért adat „fent van a honlapon.”

²⁶ Ennek egyik oka lehet az újságírás azon sajátossága, hogy a bárki számára rendelkezésre álló információknak kisebb az értéke a médiában, mint a személyes kapcsolatok révén kiszivárogtatott (még ha kevésbé megbízható, de privilegizált módon megszerzett) információknak.

²⁷ 1999. évi LXXVI. törvény a szerzői jogról.

²⁸ „Szerzői és szomszédos jogok az információs társadalomban” (Európai Unió, 2001).

²⁹ Ennek példája a titokvédelmi szabályozás, ahol a pecsétet ugyan az iratra kell tenni, de a titkosítás tárgya nem az irat, hanem az iratban lévő adat; vagy példa lehet a levéltári szabályozás, amely iratokról (köziratokról) rendelkezik, de az iratok kutathatósága nagymértékben az adattartalmuktól függ.

használóikat, mint a szellemi javak tolvajait. Lessig meg is mondta: Code is law, vagyis a mai világban a kód a törvény – nem az számít, hogy mit mondanak a jogászok, hanem az, hogy mit kódolnak a rendszerekbe az informatikusok (Lessig, 1999).

Az ingyenesség kérdése

A felhasználók többségét azonban nem érdeklik a jogszabályi rendelkezések vagy a jogdogmatikai megfontolások, sem pedig a belekódolt üzleti és politikai érdekek, inkább az érdekli őket, hogy kell-e fizetni az adatért vagy az adatállományért.

A közinformációk (konkrétan: adatok, a közérdekű adatok, a közadatok, vagy a nemzeti adatvagyon) ingyenessége tekintetében két klasszikus érv csap össze. Az egyik szerint az adatokat előállító és kezelő szervezeteket mi, adófizetők tartjuk fenn, és közfeladatot látnak el – miért fizessünk akkor még egyszer az adatért? A szervezetek ellenérve viszont az, hogy ha a polgárok kéréseit kell kiszolgáltatniuk, akkor ez elvonja az erőforrásaikat a közszolgáltatától, tehát fizessen ezért a polgár.

Az állam mintegy arra kényszeríti közfeladatot ellátó szerveit, hogy „üzleteljenek,” vagyis hogy pénzt kérjenek még az alapfeladataik körébe tartozó szolgáltatásaikért is, mivel az állami költségvetésük még a működtetésükre sem elegendő, a fejlesztésről nem is beszélve. Az információs szabadság alkotmányos értéke és törvényi előírásai rendre veszítenek a közfeladatot ellátó szervek lobbistái és az őket támogató pénzügyi kormányzat szövetségével szemben. Az olykor elriasztó mértékű díjak mellett a szakmai szóbeszédben fel-felbukkantak olyan – meg nem erősített – feltételezések, hogy például a KSH a népszámlálási alapadatait csak olyan leválogatásban hajlandó átadni a közvéleménykutatóknak, amiből azok nem tudnak kellően részletes elemzéseket végezni, mert ezzel „riválisai lennének az állami statisztikáknak.”

Azt természetesen senki sem várja el a közfeladatot ellátó szervektől, hogyha alapfeladatukon túlmenően hozzáadott értéket hoznak létre, például a KSH a népszámlálás adatainak elemzéséből kiadványokat publikál, akkor ezeket a kiadványokat ingyenesen osztogassa. A vita az alapfeladaton, illetve az adat előkeresésének és átadásának, másolásának költségein folyik.

Ehhez az állapothoz képest új helyzetet jelentett a közzsféra információinak további felhasználásáról szóló uniós irányelv kötelező áttételése a hazai jogba. Az irányelv³⁰ ugyanis az uniós információs piac fejlesztését célozza, és ennek érdekében díj szedését, sőt egy „ésszerű nyereséghányad” érvényesítését is megengedi a közfeladatot ellátó szervezeteknek. Ugyan az Infotv. és a Közadat tv. legújabb, 2015. júliusi módosítása³¹ szerint a közfeladatot ellátó szervezetek nem nyilatkoztathatják az adatigénylőt újrahasznosítási szándékáról, és el is tekinthetnek a díj szedésétől, a gyakorlatban a hatályos irányelv és a módosított magyar törvény voltaképpen széles körben legalizálja az adatkiadás díjait. Azt láthatjuk, hogy egy információs alapjog, amely bárki számára ingyenes (csak a másolás díját felszámító) hozzáférést nyújtott a közérdekű adatokhoz, ütközik egy piaci szemléletű szabályozással, és az ütközet kimenetele a jelenlegi igazgatási környezetben nem kérdéses...

³⁰ 2003/98/EK (Európai Unió, 2003).

³¹ A sok kritikus hangvételű sajtóbeszámoló közül lásd például http://index.hu/belfold/2015/06/26/irt_az_uj_sarc_az_allam_az_adatigennylore_terheli_a_koltsegeket/ vagy http://hvg.hu/itthon/20150713_Ader_alairta_akadalyozzak_a_kozerdeku_ada

Az egységes uniós információs piac kialakítására és az Unió versenyképességének növelésére irányuló kezdeményezéseknek egyébként van nem-piaci szemléletű hozadéka is: a nyílt hozzáférésű adatok EU-portálja (<https://open-data.europa.eu/hu/data>) egyablakos hozzáférést biztosít az Európai Unió intézményei, ügynökségei és egyéb szervei által létrehozott és folyamatosan bővített adatkészletek széles köréhez. Az adatportál nemcsak „a nyílt hozzáférésű adatokra vonatkozó európai uniós stratégia kulcsfontosságú eleme,” hanem egyúttal ingyenes hozzáférést nyújt az adatokhoz és nem vizsgálja felhasználásuk célját.

Közzétett-hozzáférést segítő technológiák és alkalmazások

A technológia demokratizálódása (vagyis az, hogy mindenkinek lehet okostelefonja), nem egyenlő az információs hatalom demokratizálódásával (vagyis például azzal, hogy hozzáfér-e az okostelefonjával a közinformációkhoz, adatbázisokhoz). Kézenfekvő tehát, hogy az adatkezelési technológia fejlődésével – a közzétett adatokat kezelő szervezetek informatikai és információmenedzsment fejlesztéseivel párhuzamosan – a felhasználóknak is új eszközökre, alkalmazásokra van szükségük a közinformációk eléréséhez.

Mint említettük, a hagyományos információs szabadság fogalmi keretei között nehéz értelmezni az adatbázisokat, az elektronikus információs szabadság törvényi előírásai pedig, bár adatbázisok felállításáról rendelkeznek, végső soron egyedi adatok, adategyüttesek, dokumentumok elérhetőségét célozzák, ennek megfelelően a közzétett-hozzáférést segítő számítógépes alkalmazások egy része is egyedi kérések bonyolítását könnyíti meg.

Az Átlátszó.hu internetes oknyomozó, tényfeltáró portál (www.atlatszo.hu) 2012-ben online közérdeklődő-adat-igénylő szolgáltatást indított KiMitTud néven (<http://kimittud.atlatszo.hu/>). Az igénylők egy előre elkészített levélsablonba írhatják kérdésük lényegét, kereshetnek az adatgazdák között; a rendszer figyeli a válasz törvényes határidejének letelvét, válasz híján automatikusan újraküldi az adatigénylést, valamint közzéteszi a feltevéseket és a válaszokat. 2015 júliusáig mintegy 5.200 adatigénylést kezelt a szolgáltatás és 4.704 adatgazda kapcsolati adatait tette elérhetővé.³² Az Átlátszó, tizenkilenc további országban működő testvér-szolgáltatásával közösen, az Alaveteli nevű, finn fejlesztésű, ingyenes és nyílt forráskódú, közzétett-kérés segítő szoftvert használja (www.alaveteli.org). Számos előnye mellett a szoftver kritikájaként említhető, hogy ugyan az, hogy az egyéni adatigénylők mit kérdeznek és arra a hatóságok mit válaszolnak, maga is közérdeklődő adat, de az, hogy *ki* kérdezi, az nem – az Alaveteli rendszer azonban ezt is nyilvánosságra hozza. (A fejlesztők országában feltehetően kevésbé merül fel annak lehetősége, hogy csupán a közérdeklődő-adat-kérései miatt valakit hátrány érhet, mint a rendszert alkalmazó egyes országokban.)

Más módon segíti a magyarországi közérdeklődő adatok iránt érdeklődőket az Eitv. által előírt egységes közzétett-adatkereső rendszer. Az elektronikus közzétételre kötelezett adatgazdáknak regisztrálniuk kell a www.kozadattar.hu weboldalon, le kell tölteniük az ingyenes szoftvernek a saját rendszerükhöz illeszkedő nyelvű verzióját, majd pedig a közzétett közérdeklődő adatok és dokumentumok leíró adatait (metaadatait) – például a dokumen-

³² Összehasonlításképpen: a KiMitTud brit testvér-szolgáltatása, a WhatDoTheyKnow (www.whatdotheyknow.com) 2015 júliusáig több mint 278 ezer közzétett-kérést menedzselt és 16.385 szervezet adatait teszi elérhetővé.

tum típusát, keletkezésének időpontját, pontos webcímét – be kell írniuk egy egyszerű táblázatba, amit aztán az internetre kötött rendszer automatikusan „learat”.³³ A polgárok a www.kozadat.hu weboldalon tulajdonképpen ezek között a metaadatok között keresnek, és a találataik az adatgazda weboldalára viszik őket (Galambosi, 2008). A regisztrációs és adatfeltöltési fegyelem serkentésében és ellenőrzésében sem a korábbi adatvédelmi biztosok, sem a jelenlegi hatóság nem jeleskedtek, ezért a rendszer adattartalma hiányos. Ezzel együtt nyilvánvalók az előnyei az internetes keresőgépek találatáival szemben: amíg az általános keresők találatainak relevanciája, időszerűsége és megbízhatósága kérdéses, addig a közadatkereső ellenőrizhető adatokat szolgáltat, az adatgazdák felelősségével.

Természetesen magánszervezetek is nyilvánosságra hozhatnak adatokat, adatbázisokat, ahogy ez tömegesen történik is a weboldalakon. Az üzleti szektorban esetenként nehéz elkülöníteni a marketing célú és az „objektív” adatszolgáltatást (bár ez érvényes a kormányzati kommunikációra is): nemcsak az adat hitelessége számít, hanem teljessége, kontextusa is. Emellett a korrekt tájékoztatással közölt adatok sem értelmezhetők minden esetben a megfelelő szakértelem és jártasság híján lévő felhasználók számára, ezért létrejöttek olyan civil szervezetek, amelyek közvetítő szerepet töltenek be az adatforrás és a felhasználó között, és mintegy emészthetővé teszik az információkat a laikus felhasználók számára. A közvetítő szervezetek között sajátos szerepet tölt be az amerikai National Security Archives, amely – meglehetősen ellenére – a polgárok képviseletében az USA titkosított dokumentumainak visszaminősítéséért küzd az amerikai információszabadság-törvény alapján, és a felszabadított dokumentumokat nyilvánosságra hozza, kereshetővé teszi és jegyzetekkel látja el saját honlapján (<http://nsarchive.gwu.edu/>).

Teljesen más alapon működik a *Data.gov* honlap, amely az Egyesült Államok kormányhivatalainak adatállományait teszi elérhetővé (<http://www.data.gov/>). Nem egyedi adatokat vagy dokumentumokat lehet tehát itt lekérdezni, hanem adatállományokat, adatbázisokat, statisztikákat, és ehhez a laikus felhasználó vonzó vizuális eszközöket is kap, a fejlesztők pedig fejlesztő eszközöket, amelyek segítségével a honlapon elérhető adatokat beágyazhatják saját szolgáltatásaikba, alkalmazásaikba. A rendszerhez magánszervezetek is csatlakozhatnak, sőt más országbeli adatgazdák is.³⁴ Egy ilyen rendszer felállításához és működtetéséhez egységes metaadat-szabvány következetes alkalmazására van szükség, ami minden interoperábilis adatmegosztó rendszer alapvető feltétele.

Ki mire használja az adatokat?

A közadatok és nyilvános adatbázisok hozzáférhetősége csak az első lépés a hasznosulásuk folyamatában. Az adatállományok magán, üzleti, politikai, katonai, titkosszolgálati vagy tudományos célú felhasználásának számtalan vetülete van, amelyek taglalására a területi korlátok e helyütt nem adnak módot. Csupán néhány fontosabbat megemlítve közülük, olyanokat, amelyek gyakran hiányoznak az elemzésekből:

A közinformációk hozzáférhetősége az információs hatalommegosztás egyik módja, vagyis a demokratikus társadalmi berendezkedés egyik alapeleme; éppen ezért nem elfo-

³³ A közadatkereső technikai megoldása az Open Archives Initiative internetes tartalmegosztásra és archiválásra szolgáló szabványán alapul, ennek része a metaadatok aratása (harvesting), lásd <https://www.openarchives.org/>

³⁴ A *Data.gov* honlapon jelenleg mintegy 158 ezer adatállomány érhető el.

gadható az az érvelés egyes adatgazdák részéről, hogy azért nem adnak ki közérdekű adatokat és adatállományokat, mert az emberek „visszaélnék az adatokkal.”

A közadatokhoz nemcsak a polgároknak vagy a magánszektor szervezeteinek kell hozzáférniük, hanem maguknak a közfeladatot ellátó szerveknek is (ezt célozzák a döntéselőkészítési adatokról szóló törvény hivatalok közötti adatáramlást elősegítő rendelkezései is). Az azonosítható egyénekre vonatkozó nagy tömegű adat központi hozzáférhetősége és használata azonban az ún. aktuárius társadalom³⁵ víziójához vezet, ahol az egyéni és csoportszintű cselekvés (köztük a deviáns cselekedetek) okai már nem érdekesek, csupán bekövetkezésük statisztikai valószínűsége. Ezen alapul az „új büntetés-tan” koncepciója is,³⁶ amelyben a prevenció kiterjesztésével olyan cselekmények büntetése is lehetővé válik, amit az érintett el sem követett, csupán valószínűsíthető az elkövetése. Ez alapjaiban kérdőjelezi meg a demokratikus jogállami értékrendet, az ártatlanság vélelmét, vagy a kriminológia azon törekvését, hogy megismerje a bűncselekmények elkövetőinek indítékait, és hogy végül rehabilitálja és visszavezesse őket a társadalomba.

Hasonlóan ellentmondásos felhasználási lehetőségeket kínál napjaink divatos kifejezésével, a Big Data, vagyis a nagy adattömegek elemzése is, amennyiben az adatok azonosítható egyének viselkedésére, véleményére, megjósolható cselekedeteire vonatkoznak. A nagy adattömegek elemzése ennél fogva a társadalomtudományokban is új lehetőségeket, és egyúttal új kutatási kérdéseket vet fel a kutatások eredményeinek közvetlen üzleti és politikai visszacsatolásának lehetőségei miatt (dual use).³⁷

Az adatbőség más oldalról is csábít a szakmailag, etikailag vitatható felhasználásra: általánosan elterjedt gyakorlat az interneten vagy más úton elérhető kvantitatív vizsgálatok „kimazsolázása,” vagyis az, hogy az adatokból az érdekeltek csak azokat hozzák nyilvánosságra, amelyek saját érdekeiket szolgálják, vagy saját nézeteiket támasztják alá. E megállapítások érvényességét természetesen csak akkor lehet cáfolni, ha mások is hozzáérnek az eredeti adatokhoz.³⁸

A felhasználáshoz természetesen nem elég hozzáférni az adatokhoz, elemezni is kell tudni az adatállományokat, különösen a nagy adattömegeket. Azt mondhatjuk, hogy korábban az információs monopóliumok kialakítása, az információk elzárása adott privilégiumokat és hatalmi előnyt az információk birtokosainak, ma az adatelemzési lehetőségek birtoklása nyújtja ugyanezt, ami nemcsak tudást, hanem megfelelő szoftvereket és számítógépes kapacitást is igényel. Ez természetesen társulhat az információk elzárásával is; elég csak az Edward Snowden által nyilvánosságra hozott globális titkosszolgálati lehallgatásokra és adatelemzésekre gondolnunk.

35 Az aktuáriusok biztosításmatematikuskok, akik a káresemények bekövetkezteként és költségeinek valószínűségeit számolják.

36 New Penology, klasszikus forrása Feeley és Simon tanulmánya (1992).

37 Lásd ezekről a kérdésekről Desewffy és Láng valamint Székely (2015) írásait a Replika c. folyóirat 92. számában, amely a Big Data társadalomtudományi vetületeinek szentel tematikus összeállítást.

38 Egy Európai Unió által támogatott nemzetközi kutatási projekt, a PRISMS keretében a kutatók mintegy 260 nyilvánosságra hozott közvéleménykutatás adatait vizsgálták a magánélet, a biztonság, a megfigyelés és a bizalom témakörében az 1985–2012 időszakból, és megállapították, hogy a kutatások megrendelői, végzői és publikálói között egyaránt jelentős számban vannak olyanok, akik – vélhetően szándékosan – nem közlik sem a nyers adatokat, sem a kutatás megítéléséhez szükséges metaadatokat (Watson és Wright, 2013). Egyébként az EU előírja, hogy az EU által támogatott kutatási projektek eredményeit nyilvánossá kell tenni, és ezek között vannak adatbázisok is, mint például a PRISMS által az EU 27 tagállamában végzett vizsgálat (www.prismsproject.eu).

IRODALOM

- DESSEWFFY T. & LÁNG L. (2015): Big Data és a társadalomtudományok véletlen találkozása a műtőasztalon. *Replika*, No. 92.
- EURÓPAI UNIÓ (2001): Az Európai Parlament és a Tanács 2001/29/EK irányelve (2001. május 22.) az információs társadalomban a szerzői és szomszédos jogok egyes vonatkozásainak összehangolásáról. <http://eur-lex.europa.eu/legal-content/HU/TXT/?uri=URISERV:l26053>
- EURÓPAI UNIÓ (2003): Az Európai Parlament és a Tanács 2003/98/EK irányelve (2003. november 17.) a közzsféra információinak további felhasználásáról (módosítva a 2013/37/EU irányelvvel). <http://eur-lex.europa.eu/legal-content/HU/ALL/?uri=C ELEX:02003L0098-20130717>
- FEELEY, M. M. & SIMON, J. (1992): The New Penology. Notes on the Emerging Strategy of Corrections and its Implications. *Criminology* 30(4), pp. 449–474. <http://scholarship.law.berkeley.edu/cgi/viewcontent.cgi?article=1717&context=fa cpubs>
- FERGE S. (2007): Digitális közjavak. *Eszmélet*, No. 75. http://eszmelet.hu/ferge_sandor-digitalis-kozjavak/
- GALAMBOSI A. (2008): A közérdekű adatok központi elektronikus jegyzéke és az egységes közadatkereső. In: Székely I. és Szabó M. D. (szerk.) *Szabad adatok, védett adatok 2. Információs Társadalomért Alapítvány*, Budapest. pp. 121–152.
- LESSIG, L. (1999): *Code and other laws of cyberspace*. New York, Basic Books.
- LOADER, I., & WALKER, N. (2007): *Civilizing security*. Cambridge University Press, Cambridge.
- MAJTÉNYI L. (szerk.) (1998): *Az adatvédelmi biztos beszámolója 1997*. Adatvédelmi Biztos Irodája, Budapest. <http://abi.atlatszo.hu/index201.php?menu=beszamolo1997>
- MAJTÉNYI L. (szerk.) (1999): *Az adatvédelmi biztos beszámolója 1998*. Adatvédelmi Biztos Irodája, Budapest. <http://abi.atlatszo.hu/index201.php?menu=beszamolo1998>
- MAJTÉNYI L. (2006): *Az információs szabadságok.*, CompLex Kiadó, Budapest.
- MAJTÉNYI L., MOLNÁR P., PETRI L. Á. & SZABÓ M. D. (szerk.) (2005): *Az elektronikus információs szabadság*. Eötvös Károly Intézet, Budapest.
- PÉTERFALVI, A. (2005): *Az adatvédelmi biztos beszámolója 2004*. Adatvédelmi Biztos Irodája, Budapest.
- RAAB, C., JONES, R. & SZEKELY, I. (2015): Megfigyelés és reziliencia elméletben és gyakorlatban. *Replika* 94, pp. 63–93.
- RÁTAI B. & SZEMES B. (2008): Szellemi közjavak: a nyílt forrású szellemi alkotások jövője. *Információs Társadalom*, vol.VIII. No. 1. pp. 35–44. http://www.infonia.hu/digitalis_folyoirat/2008_1/2008_1_ratai_balazs_szemes_balazs.pdf
- SZÉKELY I. (2015): Az adatmentes zónák szükségessége és esélye. Helytelen reflexió Dessewffy Tibor és Láng László írására. *Replika*, No.92.
- WATSON, H. & WRIGHT, D. (szerk.) (2013): Report on Existing Surveys. Deliverable 7.1, PRISMS Project. <http://prismsproject.eu/wp-content/uploads/2013/03/PRISMS-D7-1-Report-on-existing-surveys.pdf>

Adatokkal történő kormányzás: dilemmák és perspektívák

2014 májusában több mint száz, többségében jól ismert és a világ legkülönbözőbb régióit képviselő tudós nyílt levelet intézett *Andreas Schleicher*hez, az OECD oktatási igazgatójához, melynek ez volt a címe: *Az OECD és a PISA tesztek világszerte kárt okoznak az oktatásnak*. Az angol *The Guardian* által nyilvánosságra hozott nyílt levél¹ így fogalmazott: „Az új PISA rezsim, a visszatérő és folyamatossá vált globális mérésekkel árt gyermekeinknek azáltal, hogy elkerülhetetlenül támogatja a több és hosszabb feleletválasztásos tesztek s az „előregyártott” tanórak terjedését és hozzájárul a pedagógusok önállóságának beszűküléséhez, elszegényítve ezzel az osztálytermek világát. A PISA ily módon tovább erősítette az iskolákra már amúgy is jellemző stresszt, veszélyeztetve a tanulók és a pedagógusok jóllétét.” A tudósok, ahogy fogalmaztak, „a tanulás örömet elpusztító” PISA-vizsgálatok leállítását követelték.

Ez reakció, és mindaz, ami kiváltotta, különösen jól illusztrálja az adatok különleges erejét. A PISA-mérések elindítása óta az OECD folyamatosan és szisztematikusan elemzi e méréseknek a világ oktatási folyamataira gyakorolt hatását, és igen sok publikáció jelent meg ebben a témában.² Ugyancsak kiterjedt irodalma keletkezett annak a jelenségnek, amelyet hol „számokkal történő kormányzásnak” (Grek és Ozga, 2008; Grek, 2009), neveznek az elemzők, hol „adatokkal” (Grek és Ozga, 2010), esetenként „információval” (Bogdandy és Goldmann, 2008), „standardokkal” (Lawn, 2011), vagy éppen „összehasonlítással” történő (Nóvoa et al., 2014) kormányzásnak. Az Európai Unió tudatosan építette ki az adatokkal vagy indikátorokkal történő kormányzásnak azt a mechanizmusát, amelyet Nyitott Koordinációs Módszernek nevezünk,³ és amelyet 2000 óta alkalmaz a nemzeti oktatáspolitikák összehangolására. E módszer intézmény- és eszközrendszerre az ezeket megalapozó közösségi jogalkotás és fejlesztési programok révén fokozatosan alakult ki az elmúlt másfél évtizedben.⁴

¹ A *The Guardian* cikkét lásd itt: <http://www.theguardian.com/education/2014/may/06/oecd-pisa-tests-damaging-education-academics>

² Ezek áttekintését segíthetik az Európai Unió által támogatott Know&Pol (lásd: <http://www.knowandpol.eu>) és Transpol (lásd: <http://www.ces.ed.ac.uk/research/TransPol>) projektek keretében készült elemzések. Az OECD által készült elemzésekhez lásd *Breakspear* (2012)

³ Erről részletesebben lásd Halász (2012)

⁴ Erről részletesebben lásd Halász (2015)

Az adatokkal történő kormányzás dilemmái

Az adatokkal történő kormányzás szorosan összekapcsolódik a tényeken alapuló döntéshozatallal: ezek kölcsönösen erősítik egymást. Ott, ahol a tényeken alapuló döntéshozatal vagy szakpolitika-alakítás kiemelt figyelmet kap, nagy valószínűséggel megtaláljuk azokat a kormányzási mechanizmusokat is, amelyeket itt az adatokkal történő kormányzás kifejezéssel jelölünk, és fordítva: ahol megfigyelhetjük az adatokkal történő kormányzást, nagy valószínűséggel nagyra értékelik a tényeket és az ezekre épülő döntéshozatalt. Az adat és a tény fogalmát azonban nem szabad összekeverni, és különösen a tény fogalma igényel külön kifejtést.

A *Stanford Encyclopaedia of Philosophy* „tények” (evidence) szócikke⁵ e fogalom négy jellegzetes használatát különbözteti meg. Az első az, amikor a tényeket a vélekedésekkel vagy hiedelmekkel állítjuk szembe, illetve amikor ezek igazolására használjuk őket. A második szóhasználat a „racionálisan gondolkodó” emberek attitűdjéhez vagy viselkedéséhez köti e fogalom használatát: az ilyen emberek általában „tisztelik” a tényeket, azok előtt „meghajolnak”. A harmadik szóhasználatban e fogalom az igazság vagy a megismerés felé vezető úton található „jelként” értelmezi e fogalmat. Végül a negyedik szóhasználatban e fogalom az objektivitásra, a pontosságra és az egyes szubjektumoktól függetlenségre utal. Ez utóbbiban az igazságosság eleme is megjelenik, amikor a tényekre épülő ítékezés kap figyelmet.

A tény ebben az összefüggésben nem azonosítható a mérhető vagy mért adattal. A számszerűsített adatok éppúgy lehetnek hamisak, megtévesztőek és félrevezetőek, mint ahogy a vélekedések, és a kvalitatív információk is lehetnek pontosak, a valóságban szerzett tapasztalatok által visszaigazoltak, a racionális elme számára meggyőzőek. A tények és adatok összemosása mégis gyakran előfordul, még a tudományos diskurzusban is. A korábban említett lap egy másik cikke idézi az OECD oktatási igazgatójának egyik gyakran hangoztatott jelmondatát: „adatok nélkül csak egy újabb ember vagy, akinek véleménye van.”⁶ Ebben a mondatban az adat szó nyilvánvalóan a tény szót helyettesíti, egyúttal jelezve azt, hogy e két fogalom még egy olyan szervezet gyakorlatában is összekeveredhet, mint amilyen az OECD. Ezt talán semmi nem mutatja jobban, mint az, hogy a PISA keresztmetszeti adatai alapján gyakran fogalmaznak meg oksági összefüggéseket, noha ezt az adatok alapján nem szabadna tenni.

Dilemmát e tanulmány kontextusában az okoz, hogy kormányozni nemcsak tényekkel, hanem adatokkal is lehet, hiszen az adatok – különösen azok, amelyekkel országokat, intézményeket vagy programokat hasonlítanak össze – hatnak az érzelmekre, és a velük való szembesülés megváltoztathatja egyének, intézmények vagy az országokat irányító politikai elitek viselkedését. Az adatok hatását az érzelmekre, az önértetre, a presztízsrre a média rendkívüli módon felerősítheti. Ennek oka elsősorban az, hogy az adatok azt az érzetet keltik bennünk, mintha tények lennének, amelyek előtt „illik meghajolni”, amelyek objektivitást sugallnak, amelyek igazolnak dolgokat. Az adatokkal való kormányzás és a tényekre épülő döntéshozatal összekapcsolódásának jelentősége éppen az, hogy az utóbbi elvárásokat támaszt az adatokkal szemben, és nem tekint minden adatot ténynek.

⁵ Lásd itt: <http://plato.stanford.edu/entries/evidence>.

⁶ „Without data, you are just another person with an opinion” – idézi a *The Guardian* 2013. november 26-i száma (lásd itt: <http://www.theguardian.com/education/2013/nov/26/pisa-international-student-tests-oecd>)

Ugyanakkor éppen ebből fakad a kockázat is: a tényekre épülő, döntéshozatalra történő fókuszálás erősíti azt a késztetést, hogy az adatokat tényeknek tekintsük.

Azt, hogy az információk, ezen belül az adatok mennyire tekinthetők ténynek, azaz a vélekedéseket alátámasztó evidenciának, részben megegyezés kérdése, részben lehetnek objektív kritériumai. Abban a világban, ahol a tényekre épülő döntéshozatal ideája kialakult, azaz az orvostudomány és a gyógyítás világában, vannak ezzel kapcsolatban általánosan elfogadott normák, amelyeket részben maga a tudomány, részben a gyógyszerek és beavatkozások okozta károk miatt indított perekben hozott bírósági döntések alakítottak. Egy hivatalos amerikai egészségügyi szakértői testület például a kilencvenes években olyan hierarchiát állított fel, amelynek legalján az „elismert tekintélyek klinikai gyakorlaton alapuló véleménye, leíró jellegű tanulmányok vagy szakértői bizottságok jelentései” állt, és a csúcán az „olyan bizonyító erejű tény (*evidence*), amely legalább egy megfelelően randomizált kontrollált kísérleti kipróbálásból (*randomized controlled trial*) származik” (*Harris et al., 2001*).

Új gyógyszert vagy egyéb orvosi beavatkozást hivatalosan bevezetni a fejlett világban hosszú ideje csak úgy lehet, ha annak hatásosságát és mellékhatásait randomizált kontrollált kísérleti kipróbálással ellenőrizték. Az Egyesült Államokban a randomizált kontrollált kísérleti kipróbálás követelményének érvényesítése igen előrehaladott az oktatás területén is, aminek a 2002-ben elfogadott *No Child Left Behind* oktatási törvény és az ehhez szorosan kapcsolódó, a neveléstudományi kutatásokról hozott törvény (*Education Sciences Reform Act*) adott különösen nagy lökést. Ennek nyomán norma lett, hogy az amerikai szövetségi kormány csak olyan oktatási reformokhoz vagy fejlesztési beavatkozásokhoz ad pénzügyi támogatást, amelyek hatásosságát tényekkel igazolták. E folyamat globális kisugárzását jelezte az, amikor – nagyrészt amerikai nyomásra – a kétezres évek elején az OECD elindított egy olyan programot, amelynek eredetileg az lett volna a célja, hogy a randomizált kontrollált kísérleti kipróbálás elvét elfogadtassa az oktatáspolitikai és oktatásfejlesztési döntéshozatalban. E program⁷ végül is „felpuhult”: az oktatás világának szereplői nem tudták befogadni a „tényeknek” azt a fajta értelmezését, amely a randomizált kontrollált kísérleti kipróbálást helyezi a csúcsra: ezt sokan, közöttük az oktatáskutatás meghatározó súlyú szereplői is vitatják (pl. *Berliner, 2002; Borman, 2009*). A randomizált kontrollált kísérleti kipróbálás, és ami talán még ennél is értékesebb forrása a tényeknek tekinthető adatoknak, a hosszabb ideig folyó, longitudinális nyomon követő vizsgálatok jelentőségét nem lehet eléggé hangsúlyozni. De éppígy nem lehet eléggé hangsúlyozni az ezzel kapcsolatos fenntartások jelentőségét sem. A tényeken alapuló döntéshozatal talán legnagyobb kockázata az, hogy gyakran éppen a legfontosabb dolgokról nem állnak rendelkezésünkre tények, és ha túlságosan hangsúlyozzuk a jelentőségüket, úgy járhatunk, mint elveszített kulcsát az utcai lámpa fényében kereső ember, aki, amikor megkérdezték tőle, ott veszítette-e el a kulcsát, azt a választ adta, nem, de csak ott van fény, ezért keresi ott. A mindennapi gyakorlatban és az oktatáspolitikában folyamatosan döntések sokaságát kell meghozni olyan kérdésekben, amelyekkel kapcsolatban nem állnak rendelkezésre tények, és a döntéseket akkor is meg kell hozni, ha tényekhez nem lehet hozzájutni.

Egy másik hasonló dilemma az indikátorokhoz, azaz a társadalmi folyamatokat mérni próbáló mutatókhoz kötődik. A valóságban zajló folyamatokat megragadni kívánó in-

⁷ A program eredményeiről lásd OECD (2007).

dikátorokkal kapcsolatban többek között az a kérdés merül fel, vajon minek van nagyobb jelentősége: annak, hogy ezek az indikátorok pontosabbak, minél megbízhatóbbak legyenek, vagy annak, hogy ezek jól működjenek, mint a szereplők viselkedését befolyásoló eszközök. Az adatokkal – és az ezekre épülő indikátorokkal – történő kormányzás számára az utóbbi éppen olyan fontos, mint az előbbi. A kettő természetesen nem független egymástól: a megbízható adatokra épülő indikátorok nagy valószínűséggel alkalmazabbak a szereplők viselkedésének befolyásolására, mint a megbízhatatlanok. Ugyanakkor, ha az indikátorok politikai vagy kormányzási eszközként kezdenek működni, kikérülnek a tudomány ellenőrzése alól: hiába pontosak és megbízhatóak az adatok, ha nem gyakorolnak hatást a szereplők viselkedésére, és – fordítva – a viselkedés változását a politika nem pontos adatokkal is el tudja érni. A dilemma a politika és a tudomány kapcsolatrendszeréhez kötődik, és súlyos tudományetikai implikációi is vannak. Vajon a tudományos kutatás örömmel fogadja-e a tényeket tükröző adatok politikai vagy kormányzási felhasználását, és a politikai folyamatok racionalizálódásaként ünnepelje ezt, vagy – éppen ellenkezőleg – kétségeket fogalmazzon meg ezzel kapcsolatban, esetleg visszautasítsa az adatok ilyen formában történő politikai felhasználását? Vajon hol húzódnak a határvonalak a tudományos kutatás és a politikai cselekvés között? Az adatokkal történő kormányzás ilyen és ehhez hasonló kérdések sokaságát veti fel, és az elmúlt években ezek a kérdések maguk is egyre inkább a tudományos kutatás tárgyát képezték.⁸

Az indikátorok politikai vagy kormányzási célokra történő felhasználásával kapcsolatban a tudományos kutatás számára a legérdekesebb kérdés talán nem is az, vajon ezek elég pontosak-e és megfelelő módon tükrözik-e a valóságos folyamatokat, mint inkább az, hogy valóban azt a hatást váltják-e ki, amit az alkalmazóik várnak tőlük. Így például, ha az Európai Unió politikai döntéshozóinak az a célja, hogy a versenyképesség érdekében javuljon a felnőtt lakosság képességszintje, akkor ezt vajon a felsőoktatási végzettséggel rendelkezőkre vonatkozó „headline target”⁹ kitűzésével lehet-e elérni, amit az Unió használ, és vajon az ezzel kapcsolatban a tagállamoknak adott értékelő visszajelzés ténylegesen olyan cselekvésre sarkallja-e ezek döntéshozóit, ami végső soron a felnőtt lakosság képességszintjének javulásához vezet? A szakpolitikai koordináció céljaira használt indikátorok legfontosabb célja elsősorban nem az, hogy ezek segítségével a közösség „megtudja”, mi történik a tagállamokban, hanem az, hogy ezekkel alakítsa és harmonizálja a tagállamok szakpolitikai viselkedését. Az indikátorokat egyre gyakrabban alkalmazó nemzetközi szervezeteket és a mögöttük lévő politikai közösségeket mindenekelőtt az foglalkoztatja, hogy ezek segítségével kikényszerítsék tagjaiknak a közösen kitűzött célokkal konform viselkedését. Az indikátorok a nemzetközi kapcsolatokban tehát éppen olyan „kényszerítő eszköznek” tekinthetők, mint amilyen a jog, és voltaképpen ez utóbbi egy alternatíváját nyújtják, ami felvethet számos közjogi és nemzetközi jogi kérdést is (Bogdandy és Goldmann, 2008; Davis et al., 2012).

Az indikátoroknak a szakpolitikai koordináció és befolyásolás eszközeként történő felhasználását többek között az teszi lehetővé, hogy ezeket a közösség a nyomásgyakorlás eszközeként tudja használni. Segítségükkel rá lehet mutatni azokra, akik a közösen kitűzött célokat nem teljesítik, meg lehet őket „szégyeníteni” (Héritier, 2001), folyamato-

⁸ Lásd többek között a 2. lábjegyzetben említett kutatásokat.

⁹ Ez a „headline target”, melyről a döntést az Európai Unió legfelsőbb politikai döntéshozó testülete, az állam és kormányfőkből álló Európai Tanács hozta meg, az, hogy a 30-34 éves korosztály körében 2020-ra 40%-ot kell elérnie a felsőfokú végzettséggel rendelkezőknek.

san „nógatni” lehet őket annak érdekében, hogy változtassanak a viselkedésükön, és a bel-politikai térben „helyzetbe” lehet hozni azokat, akik nyomást tudnak gyakorolni a kormányokra. Emellett az indikátorokról történő megegyezés maga is politikai folyamat, amelynek során az érintett szereplők között érdekevezérelt alkufolyamatok zajlanak. Az indikátorokkal történő kormányzás egyik hatása éppen ennek az alkufolyamatnak racionális keretek között tartása: a számok érvként használhatóak a vitákban, és a számokról vitatkozóknak könnyebben meg tudnak egyezni egymással, mint azok, akik elveken vitatkoznak. A tudományos kutatás számára a legérdekesebb kérdés ezzel kapcsolatban annak empirikus vizsgálata, vajon az indikátorok alkalmazása valóban kiváltja-e az ezekkel szembesülők alkalmazkodását, és ez mennyire függ az indikátorok pontosságától, és mennyire attól, hogy az ezeket alkalmazók milyen ügyesen bánnak velük.

Az adatokkal való kormányzás az Európai Unióban

Az adatokkal való kormányzás az oktatás területén szorosan összefügg az indikátorok oktatáspolitikai alkalmazásával. Ez mindenekelőtt az OECD INES programjának születéséhez, fejlődéséhez és globális hatásának kiszélesedéséhez köthető, aminek kiterjedt nemzetközi és hazai irodalma van (lásd pl. OECD, 1994; Bottani, 1998; Imre, 2001; 2003a, 2003b; Smith és Baker, 2003; Martens és mtsai, 2004; Kádárné, 2007; Sellar és Lingard, 2014). Az INES program a kezdettől fogva egyszerre szolgált tudományos és szakpolitikai célokat, azaz a megismerést és a politikai kommunikációt, illetve a nemzeti szintű szakpolitikák globális szintű befolyásolását, és ugyancsak a kezdetektől fogva a meghatározó nemzetközi partnerekkel, így különösen az UNESCO-val és az Európai Unióval együttműködésben alakult.

Az Európai Unióban az adatokkal vagy indikátorokkal történő kormányzás – amit hivatalosan természetesen nem így neveztek – az 1997-es Amszterdami szerződést követően kapott figyelmet, amikor ez a közös foglalkoztatási és szociális politika megvalósításának egyik legfontosabb eszközévé vált. Az itt kialakuló kormányzási technikát később az uniós belső kormányzási reform részeként, Nyitott Koordinációs Módszer (*Open Method of Coordination* – OMC) néven beemelték az uniós kormányzati eszköztárba, kiegészítve ezzel a már korábban is létező kormányzási eszközrepertoárt (*European Commission*, 2001). Az OMC négy kulcseleme közül (közös célok kitűzése, ezekhez indikátorok és mérföldkövek hozzárendelése, nemzeti akcióprogramok készítése és ezek tartalmának és megvalósításának közös értékelése) az egyik éppen az indikátorokról szól. Ebben a kontextusban az indikátorok, és az ezek segítségével közösen meghatározott, elérendő célértékek egyértelműen a szakpolitikák harmonizálását és koordinálását szolgáló kormányzási eszközként jelenik meg.

Az oktatással kapcsolatos nemzeti szakpolitikák közösségi koordinálásáról hozott lisszaboni döntést követően, az OMC keretei között indult el az a folyamat, amelynek eredményeképpen az Európai Unióban körülbelül egy évtized alatt kiépült az oktatás indikátorokkal történő kormányzásának intézményi mechanizmusa és eszközrendszere. Ennek meghatározó állomása volt az *Oktatás és Képzés 2010* folyamat stratégiai céljainak kijelölése és az ezekhez kapcsolódó indikátorrendszer meghatározása 2002-ben, amiben kulcsszerepe volt az akkor létrehozott, a tagállamok képviselőiből álló oktatási indikátor munkacsoportnak (*Standing Group on Indicators and Benchmarks* – SGIB). Ez utóbbi az egyetlen a nemzeti szakpolitikák koordinálásának támogatására 2002 óta létrejött,

többször újrászervezett munkacsoportok közül, amely mindmáig (2015) létezik, rálátva az összes többi munkájára, és garantálva az indikátorrendszer koherenciáját. Az *Oktatás és Képzés 2010* folyamat stratégiai céljaihoz kapcsolódó munkacsoportok feladatai között a kezdetektől fogva kiemelt módon jelent meg az indikátorokkal történő foglalkozás. Az oktatási miniszterekből álló Tanács többször foglalkozott az indikátorok témájával, és több alkalommal hozott róluk döntést, meghatározva, hogy ezek közül melyek válnának a nemzeti szintű folyamatok követésének az eszközévé, és melyek fejlesztésére fordítsanak erőforrásokat.¹⁰

A 2007-ben hozott, leginkább átfogó tanácsi döntés az indikátorok négy különböző csoportját különböztette meg. Az elsőt azok alkották, amelyek létező adatokra épültek, és amelyek meghatározásával kapcsolatban már nem voltak viták. A második csoportba sorolták azokat, amelyeket létező adatokra lehetett építeni, de amelyek meghatározása még vitatott volt. A harmadik kategóriát alkották azok, amelyek fejlesztése – részben más nemzetközi szervezetekkel együttműködésben – folyamatban volt. Végül a negyedik csoportba kerültek azok az indikátorok, amelyek esetében új, uniós adatfelvétel elindítását tartották szükségesnek. Ezek közül különös figyelmet érdemelnek az utóbbiak, mivel azt feltételezték, hogy az Európai Unió, többek között az OECD PISA vizsgálatra által illusztrált módon, közvetlenül gyűjtson adatokat a tagállamok intézményei vagy polgárai körében. Az ilyen típusú adatgyűjtések lehetősége elsősorban a 2006-ban elfogadott kulcskompetenciákkal, elsősorban a más nemzetközi szervezetek által nem vizsgált transzverzális kompetenciákkal kapcsolatban merült fel.¹¹

Az Európai Unió által létrehozott és rendszeresen közzétett oktatási indikátorok a tagországok „noszogatását” szolgálják. A közösség ezzel az eszközzel próbálja elérni, hogy tagjai „vegyék komolyan” a közösen kitűzött célokat, és tegyenek olyan erőfeszítéseket, melyek nyomán azok a tíz éves stratégiai időszak végére teljesülhetnek. Az adatok nyilvánosságával és egyre intenzívebb kommunikációjukkal az Európai Bizottság tudatosan nyomást gyakorol a tagországokra, „hátszelet” adva azoknak a belső politikai tényezőknek, amelyek ezeket, a kormányok törekvéseit bírálva vagy támogatva, a belső politikai vitákban is felhasználhatják.

Az adatok ereje: új lehetőségek és kockázatok

Az indikátorokkal történő kormányzást illusztráló Nyitott Koordinációs Módszer megteremtői és korai elemzői hangsúlyozták, hogy a módszert „nemcsak annak érdekében hozták létre, hogy új politikai eredményeket produkáljon, hanem azért is, hogy a politikaformálást segítő folyamatként is működjön”(Hodson és Maher, 2001). Az adatok vagy tények felértékelődése a szakpolitikai diskurzusban elősegítheti a politikai folyamatok kiszámíthatóbbá válását és racionalizálódását, könnyítheti a konfliktusok kezelését és elősegíti a felek közötti párbeszéd tartalmasságát, általában javíthatja a kormányozhatóságot. Emellett elősegítheti, hogy ritkábban vagy egyáltalán ne kelljen használni a jo-

¹⁰ Council Conclusions of 5 May 2003 on reference levels of European average performance in education and training (Benchmarks) (2003/C 134/02); Council conclusions of 24 May 2005 on new indicators in education and training (2005/C 141/04); Council Conclusions on a coherent framework of indicators and benchmarks for monitoring progress towards the Lisbon objectives in education and training 2802nd Education, Youth and Culture Council meeting, Brussels, 24-25 May 2007.

¹¹ Minderről jóval részletesebben lásd Halász (2015)

gi kényszerítés eszközét, ami alkalmatlan olyankor, amikor az érintett szereplők eltérő ütemben haladnak valamilyen cél felé, és így nem lehet és nem szabad azonos cselekvést várni tőlük.

Az adatokkal történő kormányzás uniós gyakorlata azonban csupán egyike a lehetséges példáknak. Az Unió voltaképpen követi azt, ami több országban már korábban gyakorlattá vált. Az Egyesült Királyságban például az oktatás világából érkezett Michael Barber a kétezres évek elején a Blair-kormány vezető tisztviselőjeként vezette be a humán ágazatok egészében az adatokkal vagy indikátorokkal történő kormányzásnak azt a sajátos modelljét, amelyet később a „*deliverology*” fogalmával jelölt. Egy erről szóló kiadványban, melynek egyik szerzője volt, úgy fogalmazott, hogy „adatokat gyűjteni nemcsak hatalmat ad, hanem el is bűvöl” és azt tanácsolta a hozzá hasonló gyakorlatot követni szándékozóknak, hogy a „diagnózist a teljesítmény jelzésére leginkább alkalmas adatok azonosításával kezdjék”, majd az „adatokat elemezve próbálják kiolvasni a magas és alacsony teljesítmény mintázatait”. Így folytatva: „az elemzésnek olyan mérőföldkövekre kell épülnie, amely a vezetett rendszer teljesítményét összeveti a korábbi történeti időszakokban elért saját teljesítménnyel, a legmagasabb teljesítményt nyújtók és más rendszerek teljesítményének adataival, akár odahaza, akár külföldön” (Barber et al., 2009; 11).

Barber itt idézett szavai, de még inkább saját irányítói gyakorlata (lásd Barber, 2008) jól mutatják: az adatokkal történő kormányzás különösen hatékony módja a teljesítmény „kikényszerítésének.” Noha ezt sokan a kormányzás „puha eszközének” tekintik (szemben a jog „keménynek” gondolt eszközével), a Barber által vezetett „*Delivery Unit*” különleges eredményeket mutatott fel abban, hogy elérje néhány nehezen teljesíthető, komplex társadalmi és intézményi folyamatokat feltételező politikai ígéret teljesülését. Ez a miniszterelnöknek közvetlenül alárendelt egység úgy működött, mint egyfajta „irányító-pult”, amely lehetővé tette a kormány számára a társadalmi és intézményi folyamatokról beérkező adatok folyamatos figyelését és azt, hogy az érintett döntéshozókat azonnal riaszthassa, amikor az adatok a kitűzött céloktól való eltérést jelezték.

Fontos hangsúlyozni: azoknak az adatoknak a jelentős hányada, amelyekkel Barber és csapata dolgozott, nem makroszintű, hanem olyan mikroszintű adat volt, amely lehetővé tette egy-egy konkrét kórház, iskola vagy egyéb intézmény teljesítményének nyomon követését. A mikroszintű, az egyes iskolákra, sőt az egyes tanulókra is vonatkozó adatok gyűjtését nemcsak az olyan rendszerek teszik lehetővé, mint amilyen például a hazai kompetenciamérés, hanem azoknak a digitális tanulásszervezési eszközöknek a használata is, amelyek képesek folyamatosan és automatikusan adatokat generálni a tanulói viselkedésről. Mindez egyfelől egészen új lehetőségeket tár fel az egyéni tanulási nehézségek azonosítását és az egyénre szabott beavatkozásokat tekintve, másfelől komoly aggályokat is felvet a személyes adatok védelmével kapcsolatban. A kialakult, óriási mennyiségű adat gyűjtését és tárolását lehetővé tevő rendszerekre új adatelemzési módszerek (*data mining, analytics*) épülnek rá, ami magával hozza az ezzel foglalkozó professzionális adatelemző szakértői réteg kialakulását is. Számos jel utal arra, hogy azok a rendszerek, amelyekben a bizalom viszonylag magas szintje miatt kevésbé alakul ki aggodalom a személyes adatok védelmével kapcsolatban, előnyhöz juthatnak annak nyomán, hogy képessé válnak olyan mikroszintű folyamatok elemzésére, amelyek megismerése nagymértékben növeli a cselekvőképességüket. Az ilyen adatok gyűjtésétől és elemzésétől tartózkodó rendszerek kevésbé képesek megérteni a tanulás eredményességét gátló vagy segítő

összefüggéseket, illetve kevésbé tudnak az eredményességet különösen hatékonyan javító, pontosan célzott beavatkozásokat kezdeményezni.

Az adatok „termelése” (és ennek érdekében folyamatosan újabb területekre kiterjedő adatgyűjtések kezdeményezése), a döntéseket és gyakorlatot támogató tényeket feltáró kutatások támogatása, az indikátorok fejlesztése és ezek rendszerbe illesztése, a mérhető értékekkel kifejezett mérföldkövek meghatározása és mindezekről a legmagasabb szinten szakpolitikai döntések meghozatala együtt egy új kormányzási gyakorlat kialakulását jelzik. Ezt leírható úgy is, mint annak a kormányzási gyakorlatnak a megerősödése, amelyet a politikatudományi irodalom egy 2003-ban megjelent nagyhatalús könyv nyomán „noszogatásnak” (angol szóval „*nudging*”) nevez, és amelyet e könyv szerzői a liberális paternalizmus fogalommal is jelölnek (*Thaler és Sunstein*, 2003). Hogy mennyire nem partikuláris dologról van szó, azt legjobban talán az jelzi, hogy 2013-ban az OECD formálisan is javasolta a tagállamoknak, hogy a következő időszak oktatási munkatervében a *nudging* önálló kutatási témaként szerepeljen.¹²

A tényekre alapuló szakpolitika és a *nudging*, azaz a „noszogatással”, ebben az esetben az adatokkal vagy tényekkel történő folyamatos szembesítéssel történő kormányzás szorosan összekapcsolódnak. Nem meglepő, hogy ez kiváltotta a nemzetközi joggal vagy közjoggal és közigazgatási joggal foglalkozó elemzők figyelmét is (lásd pl. *Bogdandy és Goldmann*, 2008 vagy *Quigley és Stokes*, 2015), hiszen a hatalomgyakorlás és a kormányzás valóban új modelljéről van szó. Olyan modellről, amely nem a jogi előírások és szankciók eszközét alkalmazza, hanem a tényekre hivatkozó, azokat folyamatosan előhozó politikai kommunikációra és megegyezés-keresésre épül, és a tényekkel történő szembesítést, az ezzel együtt járó folyamatos „zaklatást”, a nem megfelelő teljesítményt nyújtók vagy a közösen elfogadott normáktól eltérők nyilvános „megszégyenítését” (*blame*) használja az alkalmazkodást kikényszerítő eszközként.

Mindehhez hozzátartozik annak erősödő elfogadása, hogy a makroszintű vagy aggregált adatok kevésbé alkalmasak arra, hogy a komplex, többszintű rendszerek működését, az azokban zajló változásokat leírják, azaz nem teszik lehetővé, hogy ezekből következtessünk a tényekre. Ehhez mikroszintű adatokra van szükség: olyanokra, amelyek az egyes egyének vagy intézmények viselkedéséről adnak jelzéseket, lehetőleg idősoros módon. A közgazdászok „viselkedés-gazdaságtannak” (*behavioural economics*) nevezik azt a megközelítést, amely a makroszintű piaci vagy pénzügyi folyamatok alakulására az egyének és intézmények mikroszintű viselkedéséből próbál következtetni, és a *nudging* filozófiája mögött többek között éppen az áll, hogy az egyének és intézmények sokaságának mikroszintű viselkedését kell befolyásolni ahhoz, hogy bizonyos makrofolyamatok létrejöjjenek, vagy éppen a kialakulásukat el lehessen kerülni. Továbbá ahhoz, hogy egy-egy szakpolitikai beavatkozás hatását értékelni lehessen, azoknak az egyéneknek és intézményeknek a viselkedését és a viselkedésükben bekövetkezett változásokat kell megfigyelni, akikre/amelyekre a beavatkozás irányult. Ehhez a szakpolitika alakítóinak nagymennyiségű adatra van szüksége a mikro-szintű szereplők viselkedéséről, e viselkedés alakulásáról, és az e viselkedést meghatározó tényezőkről.

A mikroszintű, azaz egyéni és intézményi szintű adatok gyűjtésére és elemzésére épülő szakpolitikai gyakorlat számos elméleti és gyakorlati kérdést vet fel, kezdve a személyiségi jogok és a személyes adatok védelmétől az olyan dilemmákig, hogy a mikroszintű

¹² Forrás: a szerző közvetlen személyes információja.

viselkedési változások mögött rejlő oksági összefüggések végtelen komplexitása mennyire ragadható meg olyan adatokon keresztül, amelyek a viselkedésváltozást kiváltó lehetséges okoknak csupán töredékét képesek megragadni. A mikroszintű adatok gyűjtése és elemzése a leggyakrabban ahhoz a következtetéshez vezet, hogy a megfelelő következtetések levonásához még több adatra lenne szükség, miközben az ilyen adatok begyűjtése rendkívül költséges, és az ehhez szükséges erőforrások általában nem állnak rendelkezésre. Részben ezzel függ össze az a paradigmatisztikus váltás, amely többek között a „big data” és a „data mining” szemléletmódban ragadható meg. Ez utóbbiak esetében nem az előzetes adatigény és az erre épülő tudatos adatgyűjtés és elemzés a kiindulópont, hanem éppen ellenkezőleg, a technológiai adottságok nyomán gyakran előzetes szándékok nélkül keletkező és kevésbé strukturált hatalmas adatmennyiség utólagos feltárása, és ennek révén előre nem sejtett összefüggések megtalálása. Az OECD egy kapcsolódó programja (OECD, 2014, 2015) szerint a jövőben ez lehet a gazdasági tevékenységek és társadalmi szolgáltatások hatékonyabbá és eredményesebbé tételének egyik legfontosabb forrása.

A fentiek alapján az adatokra és tényekre épülő szakpolitika és gyakorlat jövőbeni alakulását alakító tényezők közül kettőt érdemes kiemelni. Az egyik azzal függ össze, hogy az érintett szereplők mennyire értelmezik az adatokra vagy tényekre épülő döntéshozatalt kommunikációs, és mennyire technikai vagy tudományos perspektívából, továbbá mennyire képesek a kettő között megfelelő egyensúlyt fenntartani. E tanulmány elején egy olyan helyzetet idéztem fel, ahol sajátos módon a tudományos kutatás képviselői bíráltak egy olyan intézményt, amely különösen magas szintre emelte a tényekre épülő szakpolitika gyakorlatát. A PISA-vizsgálattal kapcsolatos tudományos kritikák¹³ kiterjeszthetők lennének azonban minden hasonló vizsgálatra, és ilyen kritikák megfogalmazhatóak az Európai Bizottság által alkalmazott indikátorokkal és azok használatával kapcsolatban is. A lényeg azonban itt, mint korábban már hangsúlyoztam, kevésbé az indikátorok megbízhatósága és érvényessége, mint inkább az a tény, hogy ezek alkossák a szakpolitikai párbeszéd alapját, és így nagyobb esélye lehessen annak, hogy e párbeszéd intelligens következtetésekhez vezet, mintha vélekedések alkotnák az alapját. Ugyanakkor az indikátorok fejlesztőinek és alkalmazóinak megfelelő szerénységet kell mutatniuk, és nem szabad úgy tenniük, mintha az adataik vitán felül állóak lennének. Ugyanígy óvatosságot kell tanúsítaniuk akkor, amikor az adatokról való kommunikációba bevonják a médiát. A PISA-vizsgálat adatait gyakran túlságosan direkt módon említik a médiakommunikációban, és elmulasztják felhívni a figyelmet azokra a módszertani dilemmákra, amelyek miatt az adatokat óvatosan kell kezelni. Emellett, éppen a médiával történő kommunikációban, gyakran vannak le belőlük olyan oksági következtetéseket, amelyek az adatokból nem következnek.

A másik tényező, amit az adatokra és tényekre épülő szakpolitika és gyakorlat jövőbeni alakulásával kapcsolatban érdemes újra kiemelni, az, hogy e szakpolitika és gyakorlat folyamatos és egyre nehezebben kielégíthető adatéhséget szül. Minél több adat áll a rendelkezésünkre, annál több olyan dolog tárul elénk, amelyet meg kellene értenünk, és így még több adatra keletkezik igény. Ugyanakkor az adatgyűjtés rendkívül költséges, és komoly munkaterheket rak mind az adatszolgáltatókra, mind az adatok közvetlen begyűjtését végzőkre. Ezért egyre nagyobb jelentősége lehet az olyan rendszereknek, amelyek pusztán működésükkel adatokat generálnak, és amelyek az automatikus adatgyűj-

¹³ Ezeket foglalja össze és elemzi például Lannert (2015).

tést olyan tevékenységekre építik, amelyeket az egyének és az intézmények önként végeznek, általában nem az adatszolgáltatás céljával. Ennek egyik figyelemre méltó példája az OECD felnőttek körében végzett képességfelmérésének (PIAAC) az az önkéntes, online változata, amelyet a szervezet az Európai Bizottság támogatásával hozott létre. Ez olyan felület, amely az odalátogatók számára lehetővé teszi, hogy egy interneten keresztül kitöltött teszt segítségével felmérjék saját képességeiket, miközben a rendszer automatikusan olyan adatokat generál, amelyeket elemzésre is fel lehet használni.¹⁴ A jövőben az oktatás területén várhatóan számos olyan informatikai alkalmazás terjed majd el, amelyek képesek folyamatosan és automatikusan adatokat termelni. Ennek nyomán különlegesen nagy adathalmazok keletkeznek, amelyeket már nem lehet a hagyományos statisztikai eljárásokkal feldolgozni: itt léphet be a korábban említett „data mining” és „analytics” megközelítés, ami maga is automatizálhatóvá válhat.

Az adatokra vagy tényekre épülő szakpolitika és döntéshozatal egyik legnagyobb dilemmája, hogy hol vonja meg az adatoknak azt a körét, amelyek begyűjtése lehetséges a rendelkezésre álló erőforrások keretei között és az adatszolgáltatók és közvetlen begyűjtők túlzott terhelése nélkül. Akár az OECD INES programját nézzük, akár az Európai Unióban lezajlott oktatási indikátorfejlesztést, az ebben résztvevők mindkét esetben beleütköztek a finanszírozhatóság és a terhelhetőség korlátjaiba, és nehéz döntéseket kellett meghozniuk arról, hogy mely adatokról mondanak le vagy melyeket vesznek át másoktól még akkor is, ha így azok nem optimális formában jutnak el hozzájuk. Az adatokra épülő kormányzás sok szempontból a kompromisszumok kormányzása: egyfelől a tudományos érvényesség vagy megbízhatóság normái és a szakpolitikai folyamatok vagy a gyakorlat realizálása között, másfelől a mindig növekvő adatigény és a rendelkezésre álló erőforrások szűkösségéből fakadó kényszerek között.

IRODALOM

BARBER, M. (2008): *Instruction to deliver*. Methuen, London.

BARBER, M., MOFFIT, A. & KIHN, P. (2011): *Deliverology 101. A field guide for educational leaders*. Corwin, Thousand Oaks, California.

BERLINER, D. C. (2002): Educational Research: The Hardest Science of All. *Educational Researcher*. Vol. 31, No. 8, pp. 18–20

BOGDANDY, A. von & GOLDMANN, M. (2008): The Exercise of International Public Authority through National Policy Assessment. The OECD's PISA Policy as a Paradigm for a New International Standard Instrument. *International Organizations Law Review* 5. pp. 1-58.

BOTTANI, N. (1998): The OECD Educational Indicators: Purposes, Limits and Production Processes. *Prospects*. 28(1). pp. 61-75

BORMAN, G. D. (2009): The Use of Randomized Trials to Inform Education Policy. In: Sykes, Gary, Schneider, Barbara & Plank, David N. (eds.): *Handbook of Education Policy Research*. American Educational Research Association. Routledge, New York. pp. 129-138

BREAKSPEAR, S. (2012), *The Policy Impact of PISA: An Exploration of the Normative Effects of International Benchmarking in School System Performance*. OECD Education Working Papers, No. 71,

¹⁴ Lásd az OECD Education & Skills Online Assessment című weblapját: <http://www.oecd.org/skills/ESonline-assessment>

- OECD Publishing (online: <http://dx.doi.org/10.1787/5k9fdqffr28-en>)
- DAVIS, K. E., KINGSBURY, B. & MERRY, S. E. (2012): Indicators as a Technology of Global Governance. *Law & Society Review*. Vol. 46, No. 1. pp. 71-104
- European Commission (2001): *European Governance*. A White Paper. Brussels. 25.7.2001. COM(2001) 428 final
- GREK, S. & OZGA, J. (2008): Governing by Numbers? Shaping Education through Data. *University of Edinburgh Centre for Educational Sociology*. No. 44, February 2008 (online: <http://www.leeds.ac.uk/educol/documents/174686.pdf>)
- GREK, S. & OZGA, J. (2010): Governing education through data: Scotland, England and the European education policy space. *British Educational Research Journal* Vol. 36, No. 6. pp. 937–952
- GREK, S. (2009): *The role of the OECD as an agent of Europeanization: problematisation and change in education governance in Europe*. ESRC Research Project on Transnational Policy Learning Paper presented to Edinburgh's Europa Institute seminar: Practicing EU Government: Problematisation. (online: <http://www.ces.ed.ac.uk/PDF%20Files/TPLWP1.pdf>).
- HALÁSZ G. (2012): *Az oktatás az Európai Unióban – tanulás és együttműködés*. Új Mandátum, Budapest.
- HALÁSZ G. (2015): *Evaluation of Education at the European Level*. An expert paper prepared for the European Parliament. Manuscript.
- HÉRITIER, A. (2001): *New Modes of Governance in Europe: Policy-Making without Legislating?* Max Planck Project Group Common Goods: Law, Politics and Economics. Bonn.
- HODSON, D. & MAHER, I. (2001). The Open Method as a New Mode of Governance, *Journal of Common Market Studies*. Vol 39, No. 4 pp 719-746.
- IMRE A. (2001): Nemzetközi indikátorok használata az oktatási folyamatok elemzésében. Az OECD INES Network C tevékenysége. *Új Pedagógiai Szemle*, 51 (7-8). 233-241. o
- IMRE A. (2003a): Az OECD INES tevékenysége és hatása a nemzeti gondolkodásra. In: *Szakmai napok 2002*. Országos Közoktatási Intézet, Budapest. 2003. 104-111. o.
- IMRE A. (2003b): *Jelzések az oktatásról. Oktatásunk helyzete az OECD adatainak tükrében*. Országos Közoktatási Intézet, Budapest.
- KÁDÁRNÉ FÜLÖP J. (2007): *Magyarország jelenlegi és tervezett részvétele nemzetközi indikátorfejlesztő programokban és adatfelvételekben*. Oktatási Kerekasztal. Háttér tanulmányok. (online: http://www.oktatasikerekasztal.hu/hattertanulmanyok/07/kadarne_magyarorszag_jelenlegi.pdf)
- LANNERT J. (2015): A PISA-adatok használata és értelmezése a módszertani kritikák tükrében. Társi Tudok. Kézirat (online: http://www.t-tudok.hu/files/2/educatiopisa2015_hosszuvaltozat.pdf)
- LAWN, M. (2011): Standardizing the European Education Policy Space. *European Educational Research Journal*. Volume 10. Number. p. 259-272
- MARTENS, K., BALZER, C., SACKMANN, R. & WEYMANN, A. (2004): Comparing governance of international organisations: the EU, the OECD and educational policy. *TranState Working Papers* 7. Bremen.
- NÓVOA, A., CARVALHO, L. M. & YANES, C. (2014): La comparaison comme politique: le tournant du début du 21e siècle. *Revue suisse des sciences de l'éducation*, 36 (2). pp. 265-282
- OECD (1994): *Making Education Count: Developing and using international indicators*. Paris.
- OECD (2007): *Evidence in Education: Linking Research and Policy*. Paris
- OECD (2014): *Data-driven Innovation for Growth and Well-being*. Interim Synthesis Report.
- OECD (2015): *Data-Driven Innovation: Big Data for Growth and Well-Being*. OECD Publishing, Paris

QUIGLEY, M. & STOKES, E. (2015): Nudging and evidence-based policy in Europe: problems of normative legitimacy and effectiveness. In: Alemanno, Alberto & Sibony, Anne-Lise (eds): *Nudge and the Law: A European Perspective*. Hart Publishing. pp. 61-81.

SELLAR, S. & LINGARD, B. (2014): The OECD and the expansion of PISA: new global modes of governance in education. *British Educational Research Journal*. 40(6) pp. 917-936

SMITH, Th. M. & BAKER, D. P. (2003): International Education Statistics: The Use of Indicators to Evaluate the Condition of Education Systems. In: *Encyclopedia of Education*. Second Edition, James Guthrie, ed. NY: Macmillan. pp. 1291-1296

THALER, R. H. & SUNSTEIN, C. R. (2003): Libertarian Paternalism. *The American Economic Review*. Vol. 93. No. 2. pp. 175-179

SZÉLL KRISZTIÁN

Az adathozzáférés és felhasználás nemzetközi gyakorlatai

Az (oktatási) adatok mennyiségének napjainkban tapasztalható erőteljes növekedése egyértelműen felveti a kérdést, hogy kik és milyen célból használják ezeket az adatokat, illetve, hogy milyen hozzáférés biztosított a különböző módszerekkel gyűjtött adatokhoz. A téma kapcsán érdemes körbejárni néhány olyan ország gyakorlatát, amely egységes adatgyűjtési és felhasználási alapelvekre épülő, több évtizedre visszatekintő adatkultúrával rendelkezik. Ennek megfelelően a tanulmány egy angolszász (Anglia) és egy skandináv (Svédország) kultúrkörből származó országban tárja fel a közoktatási célú adatgyűjtések alapján előálló adatbázisokhoz való hozzáférés és az adatfelhasználás gyakorlatait.

A megfelelő tudományos eljárásokkal bizonyított adatokon alapuló döntéshozatal utóbbi évtizedekben tapasztalható felértékelődésének egyik legfőbb oka az elszámoltathatóság erősödő igénye; vagyis a modern közigazgatási gyakorlatban alapvető követelménnyé vált minden olyan beavatkozás hatásának tudományos eszközökkel történő vizsgálata, amelyre közpénzeket fordítanak. A tendenciát tovább erősítette az oktatási-nevelési rendszer minőségét, eredményességét és hatékonyságát célzó legjobb fejlesztések megtalálásának és legjobb megvalósításának (implementációjának) igénye, valamint a tudományos módszerek, eszközök folyamatos fejlődése, alkalmazásuk növekvő térnyerése is (Halász, 2009, 2010a, 2010b). Az adatalapú szemlélet és megközelítésmód tehát hangsúlyos szerepet kapott a nemzetközi szakpolitikai diskurzusban, a nemzetközi szervezetek (pl. Európai Bizottság, OECD, ENSZ, Világbank) deklarált céljává vált, hogy a társadalmi jólét növelése érdekében segítse a tényeken, adatokon alapuló döntéshozatali kultúra elterjedését valamennyi szakpolitika esetében (OECD, 2008).

Ennek megfelelően több nemzetközi program is elindult a tényekre, adatokra alapozott oktatáspolitikai döntéshozatal szemléletének – a korlátok ismeretével együtt történő – széles körű elterjesztésére, a döntéshozók, oktatástatistikusok, oktatáskutatók és más szakterületek képviselői közötti tapasztalatcserére és együttműködés erősítésére, valamint az oktatáskutatások és az oktatás területén végzett adatgyűjtések, hatásvizsgálatok minőségének javítására. Ilyen volt például az OECD 2003-ban induló önálló, *Ténye-*

* A kutatás az Oktatási törzsindikátorok és mérési eredmények alapján készülő indikátorok rendszerének továbbfejlesztése című TÁMOP 3.1.8.-09/1-2010-0004 projekt keretein belül zajlott.

ken alapuló oktatáspolitikai-kutatás (*Evidence-based Policy Research in Education*) elnevezésű programja (OECD, 2007), vagy a PISA-vizsgálatok¹ elindulása is, amely előtérbe helyezte a standardizált mérések eredményeire alapozott elemzések alkalmazását és alkalmazhatóságát az oktatáspolitikai döntések során. A téma az Európai Unió oktatáspolitikájának is fontos elemévé vált, sőt a hazai oktatáspolitikai döntéshozásban is elindult egy olyan folyamat, mely során a kutatási, mérési és tudományos adatokra történő hivatkozás (lásd például az Országos kompetenciamérések jogszabályi szinten is garantált alkalmazását) valamiféle normává kezd válni.

Ideális esetben a felmerülő oktatási-, társadalmi-, illetve gazdasági igények, valamint a fejlesztési területek feltérképezésekor, a beavatkozások szükségességének igazolása-kor és végrehajtásakor a döntéshozók adminisztratív és statisztikai adatokra, tudományos kutatási eredményekre (is) támaszkodnak. Természetesen az adatok szakpolitikai célú felhasználása mellett számos más cél is körvonalazódik, az adatok, kutatási eredmények által közelebb juthatunk például az oktatási jelenségek, a teljes oktatási rendszer megismeréséhez, megértéséhez is. Ugyanakkor azt is látni kell, hogy a valóságban számos korlát, akadályozó tényező befolyásolja az adatokhoz való hozzáférést, az adatok későbbi felhasználását, felhasználhatóságát. Az adatgyűjtés céljai, az adat-előállítás módszerei, a kutatói, statisztikai, informatikai és (szak)politikai elképzelések, megközelítések, érdekek eltérései, a gyakorlati megvalósítás nehézségei, az adatbiztonsági és adatvédelmi szempontok, az adatok szerepének és korlátainak nem kellő ismerete mind-mind olyan tényezők, amelyek hatással vannak az adathozzáférésre és az adatok felhasználására is. Egyetértve Veroszta Zsuzsanna (2011) megállapításával, mindebből az is következik, hogy az adatok nem tekinthetők mindentől független, objektív tényeknek, már csak az adatok keletkezését befolyásoló eltérő szakmai, módszertani, szemléletbeli, stb. előfeltételek, megközelítések, valamint a létrejött adatok többféle módon történő interpretálhatósága miatt sem. Hozzáteve, hogy bármilyen érdek- és értékrendszer mentén is valósul meg az adatgyűjtés és az adatok interpretálása, az adatok felhasználása szempontjából a legalapvetőbb kritérium a gyűjtött adatok minősége, hiszen adatokon alapuló bármilyen (kutatói, indikátorfejlesztési, szakpolitikai stb.) döntés megszületéséhez elsősorban megbízható és érvényes adatokra van szükség. Az adatminőségi problémák egyértelműen hibás döntésekhez, hibás következtetések levonásához vezethetnek.

A keletkező oktatási adatok lehetőséget adnak olyan oktatási indikátorok kialakítására, amelyeket az oktatáspolitikai által befolyásolható területekre fejlesztenek ki azzal a céllal, hogy figyelemmel kísérjék egy kiválasztott oktatási jelenség alakulását, egy kitűzött oktatáspolitikai cél elérését, vagy akár a teljes oktatási rendszer működését. Az oktatási indikátorok lehetnek egyszerű statisztikai adatok, vagy több statisztikai adatból előállított jelzőszámok is (Tót, 2004). Az indikátorok segítik az adatok, számok tengerében történő eligazodást, és normatív jelleggel összegzik – s ezáltal egyszerűsítik is – egy jelenség legfontosabb jellemzőit, módot adva a különféle szempontú (térbeli, időbeli, csoportbeli stb.) összehasonlításokra is (Havasi, 2007).

Az indikátorfejlesztési folyamat során a már gyűjtött, illetve a későbbiekben éppen ezért gyűjtött (statisztikai) adatokból választják, illetve alakítják ki azokat a mutatószámokat, amelyek alkalmasak lehetnek egy adott oktatáspolitikai cél elérésének nyomon

¹ *Programme for International Students Assessment*, az OECD tanulói teljesítmények nemzetközi értékelésének programja.

követésére. A szakmailag megalapozott adatgyűjtések, valamint az azokra épülő indikátorfejlesztési és szakpolitika-alkotási folyamat elejét veheti annak, hogy a politikai célok és érdekek alátámasztására a legkönnyebben elérhető, és sokszor szakmai megfontolásoktól mentes adatokat használjanak fel. Ugyanakkor a különböző nemzeti és nemzetközi (gazdasági, politikai, stb.) érdekek bizonyos értelemben és mértékig erősíthetik is ezt a folyamatot, hiszen a különböző érdekek a kitűzött oktatáspolitikai célokat is befolyásolhatják. Az egyes országokban megvalósuló közoktatási adatgyűjtések és a biztosított adathozzáférések megismerése, megértése szempontjából rendkívül fontos, hogy az indikátorok nem pusztán leíró jellegűek, hanem olyan mutatószámok, amelyeket egy meghatározott szempont- és normarendszer szerint választanak ki, illetve értelmeznek. Hozzátevé, hogy a legtöbb esetben a gyűjtött oktatási adatok körét is ezen nemzeti és/vagy nemzetközi szempont-, érték- és érdekrendszer mentén határozzák meg.

A nemzeti adatgyűjtési rendszerekre a legnagyobb hatást három nemzetközi szervezet, az Európai Bizottság, az OECD és az UNESCO, valamint az általuk 1994-től bevezetett egységes – az uniós országok számára kötelező érvényű – éves közös adatgyűjtés, az UOE-adatgyűjtés gyakorolja. Az indikátorfejlesztések terén is jelentős ezen nemzetközi szervezetek hatása. Az OECD az *INES (Indicators of Education Systems)* program keretein belül, a tagállamok szakértőiből álló munkacsoport (*INES Working Party*) és annak két speciális szakértői hálózata (*NESLI* és *LSO*) közreműködésével, meghatározott keretrendszer mentén fejleszt oktatási indikátorokat. Ezek az indikátorfejlesztések és a mögöttük lévő adatgyűjtések (ide értve a PISA, TALIS,² PIAAC³ felméréseket is) jelentősen befolyásolják a tagországok oktatáspolitikai folyamatait (Halász, 2014; Kádárné, 2004; OECD, 2014). Az Európai Unió indikátormodellje – melyet az Európai Bizottság Indikátor Állandó Munkacsoportja (*SGIB: Standing Committee of Indicators and Benchmarks*) koordinál – nagymértékben támaszkodik az OECD adat- és indikátorfejlesztési tevékenységére. Lényeges különbség azonban, hogy míg az OECD tulajdonképpen tanácsadó jelleggel – ugyanakkor rendkívül meghatározó oktatáspolitikai, iránymutatási és az esetek többségében érdekevezérelt szerepkörrel – végzi tevékenységét, addig az Európai Unió komoly és egzakt oktatáspolitikai célértékeket, monitoringrendszert és támogatáspolitikát rendel az oktatási indikátorrendszeréhez. Ezáltal olyan témakörök is bekerülnek a nemzeti oktatáspolitikákba, amelyeket lényegében az Európai Unió tematizál, annak ellenére, hogy az oktatás területe, az oktatási és képzési rendszer jelenleg nemzeti hatáskörbe tartozik az Unióban. Mindez azt is jelenti, hogy ezek a kiemelt témakörök sok esetben addig, amíg az Unióban célkitűzéssé és finanszírozási alappá nem váltak, fel sem merültek és nem is épültek be szervesen bizonyos nemzetek oktatáspolitikájába. Lásd például a korai iskolaelhagyás kérdéskörét, amellyel a magyar oktatáspolitikai legfőképpen az uniós elvárások miatt foglalkozik, és többnyire akkor is csak ad hoc jelleggel és nem is kellőképpen fókuszáltnan, vagyis a célzott finanszírozás következtében más, csak nagyon közvetve kapcsolódó témakörök (pl. iskolák digitalizálása) megoldását is az intézkedéscsomagokba építve.

A tanulmányban arra keressük a választ, hogy a vizsgált nemzeti adatrendszerek (1) milyen célból gyűjtenek adatokat és fejlesztenek indikátorokat, (2) a célok elérése érdeké-

2 *Teaching and Learning International Survey*, az OECD ciklusonkénti nemzetközi tanárvizsgálata.

3 *Programme for the International Assessment of Adult Competencies*, az OECD felnőttek képességeinek és kompetenciáinak felmérésére irányuló programja.

ben milyen szintű és tartalmú adatokat gyűjtenek, valamint (3) milyen hozzáférést biztosítanak az egyes szereplőknek az adatrendszerek háttéradatbázisaihoz. A vizsgálatához hasznos segítséget nyújtottak az OECD indikátorfejlesztéseit felügyelő szakértői munkacsoport (*INES Working Party*) általunk vizsgált tagállamaiból delegált tagjai, valamint az Eurydice hálózat nemzeti irodáinak munkatársai. Természetesen a gyűjtött oktatási adatok köre és az adatok elérhetősége, felhasználása nem választható el a vizsgált országok oktatási rendszerétől, szerkezeti sajátosságaitól, valamint az adatgyűjtés és adatszolgáltatás intézményi hátterétől, ezek ismerete nélkül nehezen értelmezhető és elemezhető az adathozzáférés gyakorlata. Ennek ellenére mindezek részletes ismertetésétől terjedelmi korlátok miatt eltekintünk, azonban alapvető forrásként ajánljuk az Eurydice nemzeti oktatási rendszereket részletesen leíró honlapját.⁴

Anglia

Az angol közoktatási rendszer egyik fő sajátossága, hogy a különböző oktatási szolgáltatások nagymértékben decentralizáltak, vagyis minden iskola nagyfokú önállóságot élvez, saját maga felelős a költségvetéséért és a személyi döntésekért. Az oktatás gyakorlati megvalósítása terén a helyi önkormányzatokra/helyhatóságokra, a civil szervezetekre, az egyházakra és az egyes iskolák irányító testületeire is felelősség hárul. Angliában minden (állami és magán) iskolának regisztrálnia kell az Oktatási Minisztériumnál (*Department for Education*), így minden iskola tevékenységét rendszeresen ellenőrzik, mely ellenőrzéseket független, parlament által felügyelt tanfelügyeleti hivatal (*Office for Standards in Education, Children's Services and Skills*) végzi el. A legtöbb iskola államilag finanszírozott, így az adatgyűjtési gyakorlat feltérképezésekor elsősorban az állami iskolák adat szolgáltatási kötelezettségét vesszük alapul.

Angliában fejlett adatgyűjtési rendszer működik, melynek egyik nagy előnye, hogy szinte minden lényeges információ egyéni szinten áll rendelkezésre a tanulókról. Az oktatási célú adatgyűjtések és adatfejlesztések deklaráltan több célt is szolgálnak, amelyek közül kiemelendő a tanügyigazgatási szempontok érvényesítése és a költségvetés tervezése, ugyanakkor egyre erősebb súllyal jelenik meg a szakpolitika-elemzési, illetve oktatáskutató célú felhasználás, valamint az oktatási rendszer és szereplőinek (tanulók, tanárok, iskolák) értékelése, amely alapot ad a szakpolitikai és iskolai szintű fejlesztésekre is (1. táblázat).

4 Elérhető: <https://webgate.ec.europa.eu/fpfis/mwikis/eurydice/index.php/Countries>

1. táblázat: Az oktatási adatok gyűjtésének és fejlesztésének céljai Angliában

Oktatási rendszer irányítása, szervezése (országos szintű tanügyigazgatás)	✓
Költségvetés tervezése	✓
Szakpolitika-fejlesztés és szakpolitika-elemzés	✓
Iskolafejlesztés	✓
Oktatáskutatás (alapvető problémák megismerése, megértése)	✓
Diákok értékelése	✓
Tanárok értékelése	✓
Iskolák értékelése	✓
Teljes oktatási rendszer értékelése	✓

Angliában a közoktatási célú adatgyűjtéseket és adatfejlesztéseket az Oktatási Minisztérium felügyeli, vagyis a különböző oktatási célú adatbázisok – amelyek alapján a nemzeti és a nemzetközi oktatási célú adatszolgáltatások is történnek – adatgazdája és fejlesztője az oktatásért felelős szakárca. Ugyanakkor a minisztérium tevékenységét különböző ügynökségek, végrehajtó szervezetek és közszolgálati testületek is segítik, mint például az Oktatásfinanszírozási Ügynökség (*Education Funding Agency*), a Standardok és Tesztelések Ügynöksége (*Standards and Testing Agency*), a Tanítás és Vezetés Nemzeti Intézete (*National College for Teaching and Leadership*) vagy a Képzési és Vizsgaszabályozási Hivatal (*Office of Qualifications and Examinations Regulation*). Az államilag fenntartott iskoláknak törvény által előírt éves, iskolai szintű adatszolgáltatási kötelezettsége van (*early years census, school census*), amely során a legfontosabb – például részvételre, továbbhaladásra, végzettségre vonatkozó – adatokat tanulói szintű egyéni adatokból összeítik valamennyi képzési szinten.⁵

A tanulói szintű egyéni adatokat a *Nemzeti Tanulói Adatbázis (NPD: National Pupil Database)*⁶ különböző moduljai tartalmazzák, amely az adatokat négy szinten tárolja: az 1. szint a legérzékenyebb személyes adatokat, a 2. szint az egyéb, kevésbé érzékeny személyes adatokat, a 3. szint az iskolai szintű adatokat, míg a 4. szint a tanulók oktatással kapcsolatos adatait tartalmazza. Az adatbázis tehát a személyes adatok mellett információt nyújt többek között a teljesítményméréseken elért teszteredményekről, előzetes képesítésekről, tanulói pályáról és előrehaladásról is. A hátrányos helyzetre vonatkozóan is tartalmaz adatot az adatbázis, melynek alapját a lakcím szerint kialakított, halmozottan hátrányos helyzetet jelző index (*Indices of Multiple Deprivation*) és a térítésmentes iskolai étkeztetésben való részvétel (*Free School Meal*) adja. Az évfolyamismétlők alakulását viszont nem követik nyomon, habár ehhez az adatbázisban rendelkezésre állnak az adatok. Ennek egyszerű oka van: az évfolyamismétlés nem jellemző az angol oktatási rendszerre, így ezt nem is kísérik figyelemmel. Az NPD nem „csupán” adatokat, hanem indikátorokat is tartalmaz. Például a hiányzások száma mellett a diákok hiányzásának nyo-

⁵ Megjegyezzük, hogy a magániskolák is kötelesek éves iskolai szintű aggregált adatokat szolgáltatni (*school-level annual school census*) a legfontosabb információkról (pl. iskola adatai, tulajdonosi háttér, tanulószaám, tanárlétszám, kurzusok adatai, tandíj).

⁶ Az NPD-ről bővebb információk érhetők el a <https://www.gov.uk/national-pupil-database-apply-for-a-data-extract> weboldalon.

mon követéséhez egy ún. „tartós távollévő mutatót” (*persistent absentee indicator*) is beépítettek a rendszerbe, amely a hiányzások teljes lehetséges óraszámhoz viszonyított arányát jelzi. Ha ez a mutató egy adott diáknál eléri a 15%-ot, akkor azt jelzi a diák iskolája felé.

Az államilag finanszírozott iskolákban dolgozók legfontosabb jellemzőit az éves gyakorlatú *Iskolai Munkaerő-felmérés* (SWS: *School Workforce Survey*)⁷ biztosítja. A felmérés az állami iskolák tanáiról, valamint a pedagógusmunkát segítő alkalmazottairól gyűjt aggregált adatokat, így többek között az iskola létszámáról, megüresedett álláshelyeiről, a betegszabadságon lévők számáról, továbbá az iskolában dolgozók koráról, fizetéséről, képesítéseiről – a középfokon oktatók esetében a tanított tárgyakról, a szakképzettség megfelelőségéről és a tantárgyankénti óraszámokról – nyújt információt.

Angliában nagy hagyománya van a tanulói teljesítményméréseknek is. Jelenleg az érintett korosztályokba tartozó diákok mindegyike éves szintű mérési, értékelési programokban vesz részt, évfolyamonként eltérő mérési területeken. Fontos jellemző, hogy a mérésekből, értékelésekből keresztmetszeti és longitudinális (azaz évenként egyéni szinten összehasonlítható) adatok is előállíthatók. A tesztek a tanulók kompetenciáinak – igaz, csak bizonyos képességterületeken történő – felmérését célozzák, ezért jellemzően egyéb felvételek (pl. tanulói, tanári háttérkérdőívek) nem társulnak a mérésekhez. Lényeges elem, hogy azon iskolák számára, ahol a minimum szint alatt teljesítők aránya egy meghatározott küszöbérték (*„floor standards”*) alá csökken, kiemelt szakmai támogatást biztosítanak, majd folyamatosan nyomon követik a fejlődésüket. Amennyiben ennek ellenére sem javulnak az eredmények, további beavatkozásokra (pl. az iskolavezetés felfüggesztésére, cseréjére, az iskola külső kontrolljának erősítésére) kerül sor.

Az angol adatrendszer egyik fontos jellemzője, hogy a tanulói adatokat tartalmazó oktatási adatbázisok⁸ összehasonlíthatók, mely tanulói szintű adatintegráció lehetőséget ad a teljes tanulói pálya nyomon követésére. Igaz, az iskolarendszerről kikerülőket már kevésbé képes beazonosítani a rendszer, amely elsősorban annak köszönhető, hogy a közoktatási intézményeknek nem kötelességük nyomon követni az iskolából – általában a tan kötelezettségi kor után – kikerülő tanulókat. A 16–18. év közötti korosztályról az iskolai statisztikai jelentésekből, a Munkaerő-felmérésből (*Labour Force Survey*), valamint a NEET-statisztikákból⁹ lehet informálódni. Sőt, egy új, NEET fiatalokról szóló jelentési kötelezettséget (*„NEET scorecards”*) is bevezettek, amely eredményjelentést minden egyes helyhatóságnak közzé kell tennie.

További fontos jellemző, hogy Angliában rendkívül nagy hangsúlyt fektetnek az oktatási szereplők és a nyilvánosság tájékoztatására, az oktatási rendszerrel, iskolákkal kapcsolatos információk, adatok széles körű hozzáférhetőségére, elérhetőségére, így a rendszer szerteágazó adatfelhasználásra ad lehetőséget az oktatásirányítók, a döntés-előkészítők, a szakpolitikai elemzők, a kutatók és a szélesebb nyilvánosság számára egyaránt. Az Oktatási Minisztériumon kívül az iskolák, a helyi hatóságok, az egyéb kormányza-

⁷ Az SWS-ről bővebb információk érhetők el a <https://www.gov.uk/government/collections/statistics-school-workforce> weboldalon.

⁸ Mint például a legfontosabb közoktatási adatokat tartalmazó *NPD: National Pupil Database*, a továbbképzési adatokat tartalmazó *ILR: Individualised Learner Record*, vagy a felsőoktatási adatokat tartalmazó *HESA: Higher Education Statistics Agency*.

⁹ A nem foglalkoztatott, oktatásban vagy képzésben nem részesülő „NEET”-fiatalok (*young people not in employment, education or training*) statisztikáit lásd: <https://www.gov.uk/government/statistics/neet-statistics-quarterly-brief-january-to-march-2014>

ti szervek és a kutatási szervezetek is hozzáférhetnek az adatokhoz, így tulajdonképpen minden szervezet (egyén) önállóan is végezhet szakpolitikai elemzéseket, oktatáskutásokat.

Az angol oktatási rendszerben alapvető elv az iskolák transzparenciája és elszámoltathatósága, ami azt is jelenti, hogy jogszabályi háttérrel minden olyan adatot biztosítanak a nyilvánosság számára, amely útján az adófizetők, az oktatási szolgáltatásokat igénybe vevők megismerhetik az iskolák oktatással kapcsolatos adatait, eredményeit, és kvázi nyomást gyakorolhatnak az iskolákra (lásd például a szülők iskolaválasztási döntéseit). Ezt szolgálja az is, hogy az államilag finanszírozott iskoláknak a honlapjukon kötelezően publikálniuk kell az iskola legfontosabb adatait, elérhetőségeit, felvételi szabályzatát, tanfelügyeleti jelentéseit, tantervét, valamint az egyes támogatásokhoz történő hozzáférés eljárásait. A tanulói teljesítménymérések, vizsgák eredményeiről nemzeti szintű jelentések készülnek, továbbá az eredményekről mind az iskolák, mind a tanulók (és szüleik) visszajelzést, tájékoztatást kapnak, sőt az iskolák kötelesek közzé tenni az iskolai szinten aggregált mérési eredményeiket és a minimum szint alatt teljesítő diákjaik arányát. Amennyiben a szülők igénylik, akkor az iskolák kötelesek mindezeket az információkat nyomtatott formában is a szülők rendelkezésére bocsátani. Ugyanakkor érdemes megemlíteni, hogy általában a középosztálybeli szülők a leginformáltabbak, amely oda vezet, hogy gyermekeiket a legjobb iskolákba íratják be, felerősítve ezzel az iskolák közötti különbségeket.

Angliában az oktatási adatrendszer tehát figyelemreméltóan nyitott, az Oktatási Minisztérium honlapján¹⁰ (illetve a kapcsolódó linkeken) az adatgyűjtésekről, adatbázisokról, a biztosított szolgáltatásokról, kutatásokról, valamint a különböző statisztikákról, statisztikai kiadványokról, gyűjteményekről részletes tájékoztatás, információ áll rendelkezésre, jól strukturált, kezelhető és kereshető formában.¹¹ A statisztikákban lehetőség van előre definiált, illetve az érdeklődő által előállított interaktív adatlekérdezésekre is. Összességében ingyen elérhetőek a legkülönbözőbb oktatási területeket és témákat lefedő nemzeti, helyhatósági és iskolai (képzési) szintű statisztikák. Így például mindenki számára hozzáférhetőek az iskolák teljesítménymutatói (pl. összesített mérési eredményei, oktatási jellemzői, kiemelt tanulói és tanári adatai), a tanfelügyeleti jelentések, vagy az összes iskola elérhetőségét és legfontosabb adatait, jellemzőit tartalmazó *Edubase*¹² adatbázis is. Sőt, az *NPD* vagy a teljesítménymérések anonimizált tanulói adatai kutatási, elemzési, tájékoztatási célból bárki számára elérhetőek, azonban a hozzáférés nem automatikus, az adatokat igényelni kell az Oktatási Minisztériumtól. Az igénylés megkönnyítéséhez az adatbázis tartalmáról és adatforrásairól szóló technikai leírások, metaadatok, valamint mintaadatbázisok nyilvánosan rendelkezésre állnak. Az igénylés teljesítése során tekintettel kell lenni az Adatvédelmi Törvény (*Data Protection Act*) iránymutatásaira. Ezzel kapcsolatban megjegyzendő, hogy az adatvédelmi irányelvek alapvetően nem gátolják az adatok anonimizált és célhoz kötött felhasználását.

¹⁰ Elérhető: <https://www.gov.uk/government/organisations/department-for-education>

¹¹ A rendelkezésre álló statisztikai adatforrásokról, adatbázisokról összefoglaló kiadvány is készült (*Keen & Bolton 2015*). Hasonló kataszterek állnak rendelkezésre a szintén angolszász kultúrkörbe tartozó Írországban is, lásd például a gyermekekre vonatkozó adatforrás-leltárt (*Inventory of Data Sources on Children's Lives*): <http://www.dcy.gov.ie/inventory-of-data-sources-on-childrens-lives/>

¹² Elérhető: <http://www.education.gov.uk/edubase/home.xhtml>

Svédország

Svédország oktatási rendszere nagymértékben decentralizált, az oktatási szolgáltatások biztosítása a svéd önkormányzatok feladata, és az iskolavezetők is magas fokú autonómiával rendelkeznek. Ez azonban nem jelenti azt, hogy ne lennének olyan egységes alapelvek és célrendszerek, amelyek keretet adnak az oktatásnak, és amelyek megvalósulását és működését a svéd kormányzat a jogszabályokban foglaltakkal összhangban ellenőrzi. A svéd közpolitika általában arra törekszik, hogy a különböző intézmények, hivatalkok által összegyűjtött adatok egy helyen fussanak össze. A kiépült rendszert alapvetően meghatározza a svédok azonosítására szolgáló egyetlen személyi azonosító szám, amely az egyéni szintű tanulói adatok alapja is. Svédországban tehát a személyeknek csak egy azonosítójuk van, ellentétben például Magyarországgal, ahol külön személyi szám, személyi igazolványszám, adószám, társadalombiztosítási szám, vagy éppen tanulói azonosító szolgál ugyanazon személy beazonosítására.

A svéd adatrendszer igen fejlettnak mondható, az oktatás területén is rendkívül széles körű az adatgyűjtés, vagyis tanulói és tanári adatok egyéni és aggregált szinten is rendelkezésre állnak valamennyi képzési szinten. Az adatgyűjtések, az adat- és indikátorfejlesztések alapvető célja, hogy a létrejött adatok és indikátorok átfogó képet adjanak az oktatási rendszerről, és helyi/nemzeti szinten értékeljék az oktatási folyamatokat, továbbá segítsék, támogassák az iskolafejlesztéseket. Ezt figyelembe véve az oktatási adatokat helyi szintű tanügyigazgatásra, oktatásirányításra és szervezésre, finanszírozás-tervezésre, szakpolitikai elemzésre és fejlesztésre, iskolafejlesztésre, oktatáskutatásra, valamint a tanulók, a tanárok, az iskolák és a teljes oktatási rendszer értékelésére egyaránt használják (2. táblázat).

2. táblázat: Az oktatási adatok gyűjtésének és fejlesztésének céljai Svédországban

Oktatási rendszer irányítása, szervezése (helyi szintű tanügyigazgatás)	✓
Költségvetés tervezése	✓
Szakpolitika-fejlesztés és szakpolitika-elemzés	✓
Iskolafejlesztés	✓
Oktatáskutatás (alapvető problémák megismerése, megértése)	✓
Diákok értékelése	✓
Tanárok értékelése	✓
Iskolák értékelése	✓
Teljes oktatási rendszer értékelése	✓

A Svédországra jellemző széles körű, országos szintű éves oktatásstatisztikai adatgyűjtések során az iskolák és az önkormányzatok szolgáltatják az egyéni szintből össze-szített adatokat az oktatási rendszer valamennyi szintjén. Az oktatásstatisztikai adatgyűjtések koordinálásáért alapvetően a *Skolverket* (Svéd Nemzeti Oktatási Ügynökség) a felelős, továbbá az adatok elérhetőségét biztosító online adatbázisok kezelése, fejlesztése is az Ügynökség felügyelete alá tartozik. Egyéb kormányzati szervek, ügynökségek is fejlesztenek adat- és indikátorrendszereket a különféle oktatáspolitikai területek monitorozá-

sára, azonban ezek az adat- és indikátorrendszerek is általában a *Skolverket* statisztikáin alapulnak. A *Skolverket* is felhasználja más szervezetek adatait, dokumentumait (lásd például a svéd iskolafelügyelet jelentéseit, vagy az önkormányzatok minőségi riportjait) a saját adatbázisaiban.

Az adatgyűjtések a gyerekek, tanulók oktatási részvételének releváns ismérveire, a nevelő-oktató személyzet legfontosabb jellemzőire, valamint az oktatás költségeire és eredményeire – amely magában foglalja a tanulói teljesítménymérések eredményeit is – irányulnak. Az adatok túlnyomó többsége (pl. évismérlők, hiányzások, lemorzsolódók, hátrányos helyzetű, illetve migráns tanulók) egyéni szintű, azonban ez nem jelenti azt, hogy ezen adatok mindegyikét iskolai szintre is aggregálják.

Az adatgyűjtések eredményeként előállt adatok elérésének biztosítása egyrészt a *Skolverket* statisztikai adatrendszerében és tájékoztató felületén,¹³ másrészt három fő online adatbázis segítségével történik: (1) *önkormányzati/térségi összehasonlító adatbázis*,¹⁴ (2) *SIRIS*¹⁵ és (3) *SALSA*.¹⁶ Az *önkormányzati/térségi összehasonlító adatbázis* lehetőséget teremt az egyes önkormányzatok óvodai, iskolai és felnőttoktatási adatainak egymással történő összevetésére. Az adatbázis tanulókról és tanulócsoportokról, iskolai személyzetről és tanulói/tanulmányi teljesítményekről nyújt iskolai szintre aggregált információkat kereshető formában. A *SIRIS* az oktatás minőségét és az oktatási eredményeket monitorozó online információs rendszer, amely komplex központi statisztikai adatbázisként, valamint a minőségfejlesztés egyik eszközeként funkcionál. Az adatbázis használatát az iskolákban dolgozóknak, az önkormányzati alkalmazottaknak és szakpolitikusoknak, az újságíróknak, a tanulóknak és a szülőknek egyaránt ajánlják. Az adatbázis – kereshető, interaktív felületen – többek között tartalmazza az oktatási alapadatok, a kompetenciamérési eredményeket, a záró vizsgaeredményeket, az iskolafelügyeleti jelentéseket, valamint az elvégzett oktatási tevékenységekről szóló éves önkormányzati minőségbiztosítási riportokat is. A *SALSA* pedig egy olyan, a 9. évfolyamos tanulói teljesítménymérés eredményein alapuló statisztikai modell, amely figyelembe vesz olyan, az iskola által nem befolyásolható háttértényezőket (pl. nem, migráns háttér, szülő iskolai végzettsége), amelyek befolyásolhatják a méréseken elért abszolút pontszámokat. A *SALSA*-t elsősorban igazgatóknak, tanároknak és szakpolitikusoknak ajánlják.

A svéd oktatási rendszerben, az iskolarendszer értékelési folyamatában központi szerepet töltenek be a tanulói teljesítménymérések is, amely méréseken az érintett évfolyamok valamennyi diákja részt vesz. A teljesítménymérési programok keresztmetszeti és longitudinális adatokat is biztosítanak. Minden évben több évfolyamon is zajlik mérés, évfolyamonként különböző mérési területeken (érdekesség, hogy a felnőttoktatásban is végeznek teljesítményméréseket). A teljesítménymérések tartalmi keretei a nemzeti kerettanterv tartalmi követelményeihez kapcsolódnak.

A svéd adatrendszer alapvetően kettős célt szolgál: (1) az iskolák, önkormányzatok tájékoztatása, hogy összevethessék a saját teljesítményeiket a korábbi teljesítményeikkel és mások teljesítményeivel is, valamint, hogy felmérhessék, miként tudnák növelni és fejleszteni a saját teljesítményüket; és (2) a közvélemény, a nyilvánosság tájékoztatása, hogy

¹³ Elérhető: <http://www.skolverket.se/statistik-och-utvardering>

¹⁴ Elérhető: <http://www.jmfal.artisan.se/>

¹⁵ Elérhető: <http://sir.is.skolverket.se/siris/f?p=Siris:1:0>

¹⁶ Elérhető: <http://salsa.artisan.se/>

jobban megértsék és egyben figyelemmel is kísérjék az iskolák társadalmi funkcióit, kapacitásait és azok változásait. Ebből következően alapvető elv, hogy a gyűjtött adatokhoz az érintettek (amely akár lehet a teljes nyilvánosság is) hozzáférjenek. Ennek megfelelően az oktatási rendszerrel kapcsolatos adatok és statisztikák általában nyilvánosak, és lehetőség van előre definiált, vagy saját magunk által konstruált adattáblák böngészésére és letöltésére is. A kulcsindikátorokat szintén el lehet érni interaktív lekérdező felületeken. A mérési eredményekről iskolai és nemzeti szintű riportok készülnek, valamint a tanulók (és szüleik) is hozzáférhetnek az eredményeikhez. Lehetőség van arra is, hogy a záró vizsgaeredményeket a korábbi mérési eredményekhez kapcsolják, amely anonimizált módon kutatási célokra is felhasználható. A gyűjtött adatok különböző képzési- és iskolatípus, valamint regionális elhelyezkedés szerint is csoportosíthatóak.

A kulcsadatokról és a kulcsindikátorokról, azok változásairól rendszeres tematikus elemző kiadványokat, összefoglalókat publikálnak, valamint ezek alapján rövid szakpolitikai elemzések is készülnek a döntéshozók számára. A kevés fővel működtetett svéd Oktatási és Kutatási Minisztérium közoktatásért, iskolákért felelős osztálya mellett számos más szakmai szervezet is végez szakpolitikai jellegű kutatásokat, elemzéseket, mint például a *Skolverket*, a Svéd Iskolafelügyelet (*Skolinspektionen*), az oktatáspolitikai hatásvizsgálatokat és értékeléseket is végző Munkaerő-piaci és Oktatáspolitikai Értékelések Intézete (*IFAU*), vagy az önkormányzatokat képviselő *SALAR*.¹⁷ Ez azt is jelenti, hogy a szakpolitikai elemzések és oktatáskutatások mögött erős szakmai háttérintézményrendszer is meghúzódik.

Az egyéni szintű adatokat az iskolák és az önkormányzatok gyűjtik a személyes adatok kezeléséről és védelméről szóló törvényben (*Personal Data Act*) meghatározottak szerint. Alapjellemző, hogy az oktatási adatok gyűjtésekor is alkalmazott egyetlen személyi azonosító az adatvédelmi irányelvek betartásával rendkívül széles körű adatintegrációra ad lehetőséget. Hozzáteve, hogy a jogszabályi háttér, az adatvédelem irányelvei alapvetően nem akadályozzák az adatok gyakorlati hasznosulását.

A vizsgált két decentralizált oktatási rendszer adathozzáférési és felhasználási gyakorlata alapján több közös következtetés is levonható. Először is az, hogy a fejlett adatrendszerek elsősorban egyéni szintű adatokat gyűjtenek, így az aggregált statisztikai adatok előállítását az egyéni szint összesítésével történik. Másodsor az, hogy az adatgyűjtések és indikátorfejlesztések céljai között a költségvetés-tervezési és tanügyigazgatási szempontok mellett hangsúlyosan megjelennek a fejlesztési, kutatási és értékelési szempontok is. Végül pedig az, hogy a fejlett adatrendszerekben adott a szándék és a lehetőség az adatvédelmi irányelveknek megfelelően anonimizált formában megvalósuló és különböző célokra (pl. adminisztratív terhek csökkentését, oktatáskutatások és fejlesztések támogatását) szolgáló adatintegrációra, hozzáteve, hogy ezek a rendszerek különös figyelmet fordítanak az adatvédelem és az adatfelhasználás közötti egyensúly megteremtésére.

¹⁷ A *SALAR* (Svéd Helyi Önkormányzatok és Régiók Szövetsége) minden évben kiadja a *Nyílt összehasonlítások – általános iskolák* (*Open comparisons – Compulsory school*) elemzését, amely önkormányzati szinten hasonlítja össze az iskolák teljesítmény és erőforrás-mutatóit. Az elemzések egyrészt a *Skolverket* adatain, másrészt egyéb hivatalos statisztikákon, harmadrészt pedig tanulói kérdőíves felvételen alapulnak.

Fontosnak tartjuk azt is kiemelni, hogy a nemzetállamok adatgyűjtési és indikátorfejlesztési rendszerét erősen meghatározza a nemzetközi szervezetek adatgyűjtési és indikátorfejlesztési tevékenysége. Továbbá azt is hangsúlyozzuk, hogy az adatgyűjtési és hozzáférési alapelvek lefektetése és társadalmi elfogadottsága erősíti az adatok célhoz kötött és széles körű felhasználását. Az alapelvek mentén megvalósuló adatkultúra és az adatok szakszerű, célhoz kötött felhasználása segíti az adatgyűjtésekre irányuló társadalmi bizalom erősítését is. Ehhez járulhat hozzá az oktatási szereplők és a teljes (szakmai) nyilvánosság széles körű tájékoztatása.

IRODALOM

- HALÁSZ G. (2009) Tényekre alapozott oktatáspolitikai és oktatásfejlesztés. In: PUSZTAI G. & RÉBAY M. (eds): *Kié az oktatáskutatás. Tanulmányok Kozma Tamás 70. születésnapjára*. Csokonai Könyvkiadó, Debrecen. pp. 187–191.
- HALÁSZ G. (2010a) *A pedagógiai rendszerek általános hatás- és bevételek-vizsgálati rendszere*. Nemzeti Tankönyvkiadó. Kézirat.
- HALÁSZ G. (2010b) *A pedagógiai rendszerek általános hatás- és bevételek-vizsgálati rendszere*. In: *A pedagógiai rendszerek független hatáselemzése/bevételek-vizsgálati rendszereinek kidolgozása*. Összefoglaló tanulmány a TÁMOP3.1.1. 2. pillér 11. alprojekt számára (kutatásvezető: Vass Vilmos). pp. 128–164.
- HALÁSZ G. (2014) Az OECD és az oktatás. In: SZÉLL K. (ed): *Az OECD az oktatásról – adatok, elemzések, értelmezések*. Oktatáskutató és Fejlesztő Intézet, Budapest. pp. 7–15.
- HAVASI É. (2007) Az indikátorok, indikátorrendszerek jellemzői és statisztikai követelményei. *Statisztikai Szemle* 85(8). pp. 677–689.
- KÁDÁRNÉ F. J. (2003) Az OECD oktatási indikátor-programja. In: IMRE A. (ed) *Jelzések az oktatásról. Oktatásunk helyzete az OECD adatainak tükrében*. Országos Közoktatási Intézet, Budapest. pp. 11–26.
- KEEN, R. & BOLTON, P. (2015) *Statistics, Sources of social statistics: Education*. House of Commons Library.
- OECD (2007) *Evidence in Education: Linking Research and Policy, Knowledge management*. OECD Publishing, Paris.
- OECD (2008): *Statistics, Knowledge and Policy 2007 – Measuring and Fostering the Progress of Societies*. OECD Publishing, Paris.
- OECD (2014) *Education at a Glance 2014. OECD indicators*. OECD Publishing, Paris.
- TÓT É. (2004) *Oktatási Indikátorok, az OECD INES Network-B tevékenysége*. Kutatás közben 264, Felsőoktatási Kutatóintézet, Budapest.
- VEROSZTA ZS. (2011) Adatok az oktatáspolitikai és a kutatás közti térben. *Educatio* 20(4). pp. 521–534.

Köznevelés az adatok bővületében – az adattermelés forradalma

Az adattermelés forradalma szinte kivétel nélkül végigsöpör minden tudományágon.

Hamarosan a pedagógia¹ is célkeresztbe kerül. Vajon feltettük azokat a kérdéseket és megadtuk azokat a válaszokat, melyek felkészültségünkről tanúbizonyságot tesznek? Van elképzelésünk arról, hogy mit jelent az adattermelés forradalma a pedagógiában? Esetleg már kész terveink vannak a közeljövő problémáinak és lehetőségeinek a hatékony kezelésére? Ha a kérdésekre igen a válaszunk, megnyugodhatunk. Ha esetleg mégsem, akkor – miként az alábbiakban is – folytassuk a problémák felvetését és a válaszok keresését mindezen területeken, még ha e válaszok olykor vitára is adnak okot.

Az adattermelés forradalma két – látszólag talán nem annyira jelentős – következménnyel jár a pedagógia vonatkozásában: míg a rendelkezésre álló adatok száma egyre nagyobb ütemben növekszik, addig az adatok aggregáltsági szintje folyamatosan csökken. A következmények végiggondolása olyan paradigmaváltást jósol, mely a pedagógia évszázados hagyományokra visszatekintő rendszerét alapjaiban fogja megváltoztatni.

Az adatmennyiség növekedése a gyakorlatban

A rendelkezésre álló adatok számának növekedését jól példázza, hogy míg Galileo Galilei – korának egyik legnagyobb tudósa – 3 bájt/éjszaka sebességgel „termelte” az adatokat,² addig ma becslések szerint 2,2 millió terabájt adat keletkezik naponta.³

¹ A pedagógia fogalmát tanulmányunkban a tanítás – tanulás folyamatával összefüggésben értelmezzük, az adatfelhasználásban rejlő lehetőségeket elsősorban ebben a kontextusban kívánjuk hangsúlyozni. Ugyanakkor a tanulmány több pontján a pedagógiát is magába foglaló köznevelési ágazat egészére vonatkozó megállapításokat teszünk, amennyiben az adatvagyon csak ágazati szinten értelmezhető. Ennélfogva az ágazati szintű adatvagyon egyik felhasználási területe a pedagógia.

² Forrás: <https://indico.kfki.hu/event/103/contribution/0/material/slides/0.pdf>

³ Forrás: http://www.portfolio.hu/vallalatok/it/elkepeszto_mennyi_adat_letezik_mire_lehet_felhasznalni.186053.html

► *Educatio* 2015/3. Salomvári György: *Köznevelés az adatok bővületében – az adattermelés forradalma*, 73–85. pp.

Hasonló jelenség a köznevelésben is tetten érhető: míg az 1997-ben először említett Köznevelés Információs Rendszere (továbbiakban: KIR)⁴ kezdetben 3 adatbázissal indult, addig ma az adatbázisok száma közel 500 db. Természetesen a nyilvántartott adatok száma is folyamatosan bővül: például 2015. szeptember 3-án reggel 7:00 és délután 18:00 óra között 273 950 db elektronikus kérelmet fogadott be és dolgozott fel a KIR, melynek eredményeképpen milliós nagyságrendű adattal bővült a nyilvántartott adatok száma – csak ezen az egy napon. Ma a KIR hozzávetőlegesen 4,1 millió magyar állampolgárról tartalmaz részletes adatokat, és akkor még nem beszéltünk a nem személyekkel kapcsolatban keletkező adatokról.

Az adattermelés során keletkező adatok jellemzői

Mielőtt az aggregátsági szint csökkenésére rátérünk, célszerű a termelt adatok definíció alapú kategorizálásával röviden foglalkozni – elsősorban logikai megközelítést alkalmazva.

Az *elemi adat* fogalmának definiálása tanulmányunk szempontjából rendkívül fontos. Az „elemi adat” valójában egy olyan tulajdonság, mellyel a meghatározott feltételeknek elegendő adatokat jellemezzük. Az elemi adatok az adott adatrendszer keretein belül nem állíthatók elő más adatokból, tehát nem matematikai művelet eredményeképpen jönnek létre, másrészt nem bonthatók tovább kettő vagy több további adatra. Az elemi adat tulajdonság időben dinamikusan változó jellemző: megszűnik, ha adatkör-bővítés során olyan adatok kerülnek be a rendszerbe, melyek biztosítják a korábban elemi adatnak minősített adat műveletekkel történő előállíthatóságát.

A keletkezési helyükön létrejövő adatokat az egyszerűség kedvéért nevezzük *forrás-adatoknak*. A forrásadatok tipikusan elemi adatnak minősülnek, vagyis műveletek segítségével nem állíthatók elő más adatokból. Például egy szóbeli felelet során szerzett érdemjegy elemi adatnak minősülő, intézményen belüli forrás adat.

Ezzel szemben a *számított adat* egyszerű matematikai művelet segítségével előállított, új adat. A matematikai művelet vonatkozhat elemi adatok vagy számított adatok tetszőleges kombinációjára. Például egy írásbeli dolgozatra adott érdemjegy adott esetben már nem minősül elemi adatnak, miután az egyes feladatokra adott részpontszámok alapján számítással hozható létre.

A *statisztikai adat* általában műveletek sorozatával jön létre. Az adatforrások „környékén” tipikusan még számított adat, azonban az adatfeldolgozás eredményeképpen adatrendszerekbe bekerülve már nagyon gyakran minősül elemi adatnak, vagyis adott adatrendszeren belül nem bontható fel további elemi adatokra. (A definíciók értelmezését a következő, aggregátsági szinttel foglalkozó rész nagymértékben segíti.)

Az adattermelés során létrejövő adatok egy része – különösen a pedagógiában – konkrét személyekkel kapcsolatban keletkezik, az adott személyeket jellemző információ. Ha személyek azonosítására – azaz az adott személy és a vele kapcsolatban keletkező infor-

⁴ A „közzoktatási információs rendszer” kifejezés először egy 1997-ben, az akkori Művelődési és Közzoktatási Minisztérium által kiírt pályázatban jelent meg. A KIR 1999-től jogszabályi felhatalmazás alapján működik, jelenleg a 2011. évi CXCV. törvény a nemzeti köznevelésről, valamint a 229/2012. (VIII. 28.) Korm. rendelet a nemzeti köznevelésről szóló törvény végrehajtásáról szabályozza működését. Továbbá fontos megjegyezni, hogy a KIR jogilag nem tartalmazza az összes köznevelési adatot kezelő rendszert, ugyanakkor – az egyszerűség kedvéért – tanulmányunkban ezeket is a KIR fogalma alatt tárgyaljuk.

máció egyértelmű összekapcsolására – alkalmas és szükséges természetes adatok rendelkezésre állnak, *személyazonosító adatokhoz kötött egyéni szintű adatokról* beszélhetünk. (A természetes személyazonosító adatok a következők: családi és utónév, születési családi és utónév, születési hely, születési idő és anya születési családi és utóneve. Amennyiben a feldolgozott adatok csak részben állnak rendelkezésre, a személy azonosítása nem végezhető el 100%-os biztonsággal.)

Pedagógiai kutatási szempontból különösen fontos megérteni, hogy a személyazonosításra alkalmas adatok és az adott személlyel kapcsolatban keletkező adatok közötti kapcsolat megszüntethető – ebben az esetben beszélhetünk *egyéni szintű, de személyazonosításra nem alkalmas adatokról*. Másképpen abból, hogy egyéni szintű adatok birtokában vagyunk, még nem következik a konkrét személy beazonosíthatóságának lehetősége – amennyiben az egyéni szintű adatok nem tartalmazzák a személy azonosítására alkalmas adatokat vagy a személy azonosítására alkalmas adatokhoz vezető kódokat.

Mit értünk az aggregáltsági szint csökkenésén?

A korábban ismertetett KIR adatok egyértelműen megmutatták számunkra, hogy az adattermelés a köznevelés egész rendszerében jelen van, ezen belül a pedagógia is aktív részese a folyamatnak. Az előre jelzett paradigmaváltásnak ugyanakkor szükséges, de nem elégséges feltétele az adatok mennyiségének növekedése, az aggregáltsági szint csökkenése ugyanúgy nélkülözhetetlen.

Ezek után az aggregáltsági szint csökkenése nem más, mint az adatrendszerekben kezelt elemi adatok „átváltozása” statisztikai vagy számított adatból forrásadattá. Vagyis a magas aggregáltsági szintet képviselő statisztikai adatok helyett a lehető legalacsonyabb aggregáltsági szintet képviselő forrásadatok kerülnek előtérbe. Természetesen az adatmennyiség és az aggregáltsági szint szoros összefüggést mutat: amennyiben nem statisztikai adatokat gyűjtünk, hanem az összes olyan forrásadatot, ami lehetővé teszi különböző számítások elvégzését, lényegesen több adat kezelésére van szükség. (A jelenség nyilvánvalóan technológiafüggő, amennyiben a számítási, adattárolási és adatátviteli kapacitások jelentős bővülése peremfeltételnek minősül.)

A példánknál maradva egy korszerű, pedagógiai célú adatrendszernek nem az osztályátlagra van szüksége, de még csak nem is a tanulónkénti érdemjegyekre, hanem tanulónként a szóbeli feleletek érdemjegyeire és az írásbeli dolgozatok feladatonkénti pontszámára. Vegyük észre, hogy mindkét esetben elemi adatokra „vadászunk”, melyek forrásadatként már nem bonthatók szét további adatokra. Másképpen fogalmazva, míg korábban az adatrendszerek statisztikai adatokat tartalmaztak elemi adatként, addig ma – az aggregáltsági szint csökkenésének következtében – az adatrendszerek elemi adatai leginkább forrásadatok.

Az aggregáltsági szint csökkenése a köznevelési adatrendszerekben is megfigyelhető: például a köznevelés statisztikai adatgyűjtés 1999-ben még papír alapú adatlap kitöltésével történt – ezáltal kizárólag magas aggregáltsági szintet képviselő adatbázis jött létre, addig ma már részben a gyakorlatban is megvalósuló erős törekvés, hogy a statisztikai adatok jelentős részét egyéni szintű elemi adatokból generáljuk.

Az adattermelés forradalmának következményei a pedagógiában

A leírtak alapján nem lehet kétségünk afelől, hogy az adattermelés forradalma a köznevelési ágazatban, ezen belül a pedagógiában is jelen van. Mind az adatok mennyiségének növekedése, mind pedig az aggregáltsági szint csökkenése egyértelműen erre utal. Ugyanakkor még mindig jogosan tehetjük fel a kérdést: miért jelent változást a pedagógiában ez a jelenség?

„Olyan mennyiségű és olyan minőségű adat, amivel ma az egyéni tanulói nyomkövetést lehetővé tevő rendszerek rendelkeznek, soha nem állt a rendelkezésünkre. Ezeknek az adatoknak köszönhetően az iskolai folyamatok „mélyrétegeibe” tudunk belelátni, az oktatási rendszerek immár szinte semmit nem tudnak „elrejtetni”. Látni lehet minden egyes iskolát, minden egyes tanulót, és bizonyos rendszerekben kisebb-nagyobb pontossággal minden egyes tanárt és minden egyes kurzust. Látjuk ezek teljesítményének, eredményességének időbeli mozgását.” (Halász, 2010)

Tehát az adattermelés eredményeképpen létrejövő adatvagyon kifejezetten alkalmas közvetlen pedagógiai célú alkalmazásra, miután a statisztikai megközelítést felváltja az egyéni szintű elemi adatok világa, ami alapvető és egyben sok mindent felforgató újdonságnak tűnik – így méltán beszélhetünk paradigmaváltásról.

A továbbiakban túllépünk az adattermelés következményeinek a feltárásán, és a létrejövő adatvagyon pedagógiai célú alkalmazási lehetőségeit és az alkalmazással kapcsolatos anomáliákat próbáljuk körüljárni.

Az adatvagyon-hasznosítás új lehetőségei a pedagógiában

Az adatvagyon olyan rendezett adathalmaz, melyet az adattermelést követő adatfeldolgozás hoz létre. Az adatvagyon – akár pénzben is kifejezhető – értéket képvisel, mely a felhasználása során jelenik meg és válik mérhetővé. Az adatvagyon létrehozásáról és jellemzőiről később szó lesz, egyelőre nézzük meg, hogy az adatvagyon felhasználása milyen előnyökkel kecsegtet a pedagógiában, milyen konkrét pedagógiai tevékenységek véghezvitők el adatok alkalmazásával – a teljesség igénye nélkül, felsorolásszerűen:

- állapotok jellemzése
- időbeli változások követése, trendelemzés
- változások irányának és nagyságának felmérése, trend-előrejelzés
- előzetes várakozások és/vagy célkitűzések megfogalmazása
- folyamatok modellezése, várható hatások előrejelzése
- folyamatok induló és záró állapotának összehasonlítása
- összehasonlítások széles köre
- indikátorok előállítása
- hatékonyság, eredményesség és méltányosság kifejezése (Lannert, 2004)
- kompetenciák mérése
- gyors és egyszerű kommunikáció pl. a szülők felé
- döntések előkészítése, hatásai feltérképezése.

A felsorolt lehetőségeket alaposan szemügyre véve könnyen juthatunk arra a következtetésre, hogy valójában az adatvagyon-hasznosítás semmi olyan újdonságot nem hoz be a pedagógia rendszerébe, amit egyébként nem végzünk hagyományosnak mondott eszközökkel. Ha ez valóban így van, akkor mégis miért beszélünk paradigmaváltásról?

A probléma feloldásához az erőforrás-hatékonyság környékén szükséges kutakodnunk. Ahhoz, hogy a felsorolt tevékenységeket hagyományos pedagógiai módszerekkel színvonalasan elvégezzük, meglátásunk szerint minden egyes gyermek és tanuló körül kiválóan képzett, széles körű módszertani repertoárral felvértezett pedagógusoknak és egyéb szakemberek hadának kellene tevékenykedni – ráadásul sok esetben az egyes személyekkel külön-külön foglalkozva. (Lehet, hogy ezt hívjuk ideális iskolának?) Ezzel szemben az adatvagyon alkalmazása biztosan nem nyújtja a személyes kapcsolatokra épülő pedagógia minőségét – ellenben kiegészítve azt hatékony erőforráskezelést tesz lehetővé úgy, hogy közben egyetlen gyermeknek vagy tanulóknak sem kell kimaradni az adatvagyon felhasználására is építő pedagógiai rendszerből. (Halász, 2007)⁵

A megfontolásaink alapján úgy tűnik, hogy az adatok felhasználása nem abban jelent újdonságot, hogy különleges és újszerű pedagógiai tevékenységeket hoz létre (bár ez sem zárható ki), hanem sokkal inkább abban, hogy a korábban is végzett tevékenységek tömeges, lényegében minden egyes tanulóra kiterjedő alkalmazhatóságát biztosítja elfogadható színvonalon. (Az „elfogadható színvonal” elsősorban társadalmi közmegegyezés tárgya, de természetesen szorosan összefügg az adatvagyon formájában rendelkezésre álló adatok mennyiségével és aggregáltsági szintjével is.)

Az adatok pedagógiai célú felhasználásának globális kockázata

Mielőtt önfeledten hátradölnénk és megelégedve szemlélnék a kibontakozó adatalapú „szép új világot” – ami ráadásul meglehetősen költséghatékonyak is tűnik –, azonnal ejtsünk szót a legnagyobb globális kockázatról, mely álláspontunk szerint ugyan nem kérdőjelezi meg az adatok széles körű felhasználhatóságát, viszont igencsak óvatossá kell, hogy tegyen bennünket. A legnagyobb globális kockázat felismeréséhez a következő természettudományos szemléletű idézet pedagógiára vonatkozó értelmezését javasoljuk:

„A feltételezés az, hogy ha egyszer megértjük a részeket, akkor könnyű lesz az egészet felfognunk. [...] Közel vagyunk ahhoz, hogy majdnem mindent tudjunk, amit a részekről tudni lehet. A természet egészének megértésétől azonban ugyanolyan messze vagyunk, mint bármikor korábban. [...] Az ok egyszerű. Az egyszerűsítést erőltetve beleütközünk a komplexitás kemény falába. Megtanultuk, hogy a természet nem egy jól megtervezett összerakós játék, amely csak egyféleképpen rakható össze. [...] Manapság egyre inkább felismerjük, hogy semmi nem történik elszigetelten. A legtöbb esemény és jelenség része egy komplex, univerzális kirakójátéknak, amelynek sok-sok darabja egymással kapcsolatban és kölcsönhatásban áll, egymást befolyásolja.” (Barabási, 2013)

A pedagógia ilyen értelemben szintén egy kirakójáték, ahol semmi nem történik elszigetelten, az események és jelenségek egy komplex rendszer részei, melyek sok-sok darab-ból állnak össze, ráadásul az egyes darabok kapcsolatban és kölcsönhatásban állnak, egymást befolyásolják.

Az adatok használatának legnagyobb kockázata tehát a komplexitás elvesztése: az egyes részeket le tudjuk írni adatokkal, de közben az egészet – vagyis a gyermeket és tanulót, mint komplex rendszert – szem elől tévesztjük. A kockázat valós, minden egyes

⁵ Az Egyesült Államokbeli No Child Left Behind átfogó oktatásfejlesztési program – mely korszakos jelentőségű – legfontosabb üzenete éppen az, hogy „kivétel nélkül minden egyes gyermek tanulása sikeressé válhasson”. Ld. bővebben: HALÁSZ G.: Tényekre alapozott oktatáspolitikai, http://halaszg.ofi.hu/download/Evidence_based_study.pdf

adatfelhasználónak minden egyes alkalommal tisztában kell lennie az adatfelhasználás korlátaival. Az adatvagyon-alkalmazás feltételeinek megteremtése nem lehet cél, csak és kizárólag eszköz.

Az adatvagyon hasznosításának sajátosságai a pedagógiában

A meteorológia az adatfelhasználásra erősen építő tudományág. Minél nagyobb számú elemi adat áll rendelkezésre, és minél pontosabbak a folyamatok leírására alkalmazott matematikai modellek, annál pontosabban és részletesebben leírhatók és előre jelezhetők az időjárásban bekövetkező események.

Az egészségügy szintén épít az adatfelhasználásra: ha egy egészséges ember állapotát leíró adatokkal tisztában vagyunk, az állapot leírására használt adatokban bekövetkező változások azonnal jelzésértékké válnak.

A pedagógia esetében hasonló eljárást kellene követni: egyéni szintű elemi adatok segítségével részletesen leírni az adott személy aktuális állapotát, majd megfogalmazni egy jövőben célként elérendő állapotot. Ha ez megvan, a pedagógiai tevékenység a kezdeti állapotból kiindulva a célállapot elérésére irányul. A pedagógiai tevékenység megtervezése és végrehajtása épülhet adatvagyon felhasználásra: ha a tudásbázis tartalmaz olyan korábbi – elemi szintű adatokkal részletesen dokumentált – esetet, amikor hasonló kezdeti állapotból különböző személyek már „eljutottak” a megfogalmazott célállapotba, nincs más dolgunk, mint a „célravezető” tevékenységet újra lefolytatni. Minél nagyobb és részletesebb a rendelkezésre álló adatvagyon, annál nagyobb valószínűséggel választunk korábban már bevált pedagógiai tevékenységet – legalábbis az adatvagyon-hasznosítás általános módszertana szerint.

Végtelenül leegyszerűsítve (és ezzel egyben a szakszerűséget is kockáztatva) a betegellátás sok esetben hasonló logika szerint működik: ha a testhőmérsékletre vonatkozó adat nem megfelelő, beavatkozás kezdődik a kiinduló (láz) állapot megváltoztatására, a kívánt célállapot elérésére. A szükséges beavatkozás kiválasztása korábbi esetek adataira épül, amikor is a kiinduló állapotból sikerült a célállapot elérése.

Ezzel az egyébként rendkívül technokrata megközelítéssel alapvetően két probléma is van a pedagógia vonatkozásában: legyen társadalmunk demokratikus, ami esetünkben azt jelenti, hogy a követendő példák sokféleségével számolunk, és a sokféleséget elismerjük. Még ha egy adott személy aktuális állapotát kellően sokszámú elemi adattal le is tudnánk írni, az elérendő célállapot meghatározása lehetetlenül nagy vállalkozásnak tűnik. Ráadásul további problémát okoz az időbeliség: pedagógiai rendszerekben a beavatkozások hatása sok esetben rendkívül hosszú idő elteltével jelentkezik, vagyis soha nem lehetünk abban biztosak, hogy egy időben lényegesen később megfigyelhető állapot milyen beavatkozások és egyéb hatások eredményeképpen jött létre.

Következtetésünk ezek után az, hogy a pedagógia esetében olyan speciális körülmények is fontos szerepet játszanak az adatvagyon hasznosítása során, melyek feltérképezése rendkívül komoly kihívást jelent a szakterület kutatói és egyéb szakemberei számára. A pedagógia álláspontunk szerint nehezen hasonlítható össze más, adatvagyonra építő szakrendszerekkel az adatvagyon-hasznosítás módszertanát illetően, ugyanis az adatmennyiség további növelése és az aggregáltsági szint csökkenése – ellentétben más rendszerekkel – számos kérdésre elvileg sem adhat kielégítő választ, többek között az itt felvetett problémákra sem.

Az adatvagyon-hasznosításra is támaszkodó pedagógiai rendszerek működtetése

Amikor egy történelemtanár barátomnak megemlítettem, hogy micsoda lehetőségek vannak a pedagógiát támogató adatvagyon tantermi hasznosításában, meglehetősen furcsán nézett rám...

Pedig ezen a területen is alapvető szemléletváltásra van szükség.

Első gondolatunk talán az, hogy az adatok kezelése informatikai szakemberek feladata – holott ez ma már messze nincs így. Először is rengeteg olyan önkiszolgáló rendszer létezik, mely lényegében informatikai tudás nélkül is lehetővé teszi a rendelkezésre álló adatok korlátlan felhasználását. Másrészt az informatikusok két lényeges dologban nem tudnak segíteni: a szakterületi igények megfogalmazása és az adatokból előállítandó információk meghatározása, valamint a kapott eredmények értelmezése és kiértékelése egészen bizonyosan nem az informatika hatáskörébe tartozik, márpedig éppen ezek a kulcsfontosságú momentumok. Meggyőződésünk, hogy informatikai megvalósíthatóság és kiszolgáltatottság ma már nem jelent korlátot az adatvagyon pedagógiai célú hasznosításában.

Az adatvagyon-hasznosításra építő pedagógiai rendszerek szempontjából kulcsfontosságú a kutatással foglalkozó szakemberek tevékeny részvétele a folyamatokban. Miért kellene minden egyes pedagógusnak és a neveléssel-oktatással kapcsolatba hozható szakembernek saját magának kitalálni, hogy adott helyzetben milyen elemi adatokra van szüksége, azokat feldolgozni és a kapott eredményeket módszertani szempontból kiértékelni hogyan célszerű? Persze a kész megoldások konfigurálása és helyi viszonyokhoz illesztése már más lapra tartozik, ehhez az oktatáskutatás inkább csak általános módszertani segítséget tud biztosítani.

Ezek után a feltett kérdésre javasolt válaszunk a következő: az adatvagyon-hasznosításra támaszkodó pedagógiai rendszereket a pedagógiai tevékenységben közvetlenül érintett pedagógusok és egyéb szakemberek működtetik – szoros kapcsolatban a támogatást nyújtó, kutatásokkal rendszer szinten foglalkozó szakemberekkel.

Mint ahogy azt oly sokszor látjuk, a pedagógusok ismét csak kulcsfontosságú szereplői a pedagógiai rendszereknek, az adatvagyon pedagógiai célú hasznosításának.

Kihívások a pedagógusok számára

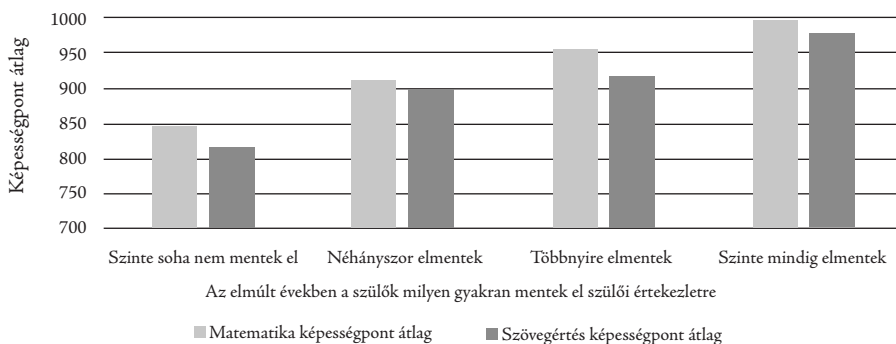
Az adatvagyon hétköznapi felhasználását biztosító információsrendszer-szolgáltatások kezelése valamennyi pedagógus számára könnyedén elsajátítható, ezáltal egyszerűen és gyorsan létrehozhatók a legkülönbébb kimutatások – ez tehát nem jelent kihívást. Ezzel szemben a kapott eredmények értelmezése és pedagógiai célú alkalmazhatóságának megítélése alapos felkészültséget és nagy körütekintést igényel.

Éppen ezért talán az egyik legnagyobb dilemma napjainkban, hogy az oktatáskutatással foglalkozó szakemberek és a pedagógusok között hol húzódjon az adatvagyon-felhasználás határa. Mekkora veszélyt jelent, ha a pedagógusok olyan eszközt kapnak a kezükbe, mely eszköz alkalmazására egyelőre még nincsenek felkészülve?

A probléma érzékeltetésére nézzünk egy példát. A következő táblázat a tesztelési fázisban lévő KIMER⁶ segítségével készült, mely a felhasználók számára sok tekintetben korlátlan adatkonfigurációs lehetőséget biztosít. A használata rendkívül egyszerű, a táblázat elkészítése 1-2 percet vesz igénybe.

	Az elmúlt években a szülők milyen gyakran mentek el szülői értekezletre?			
2014. évi adatok	Szinte soha nem mentek el	Néhányszor elmentek	Többnyire elmentek	Szinte mindig elmentek
Matematika képességpont átlag	846,4	912	956	998
Szövegértés képességpont átlag	815	901	918,8	978

Kompetenciamérés átlageredmények a szülői értekezlet látogatási gyakorisága függvényében



A számadatok és grafikon alapján akár azt a következtetést is levonhatnánk, hogy a kompetenciamérés eredményeit úgy tudjuk javítani, ha növeljük az értekezleteken részt vevő szülők számát – ami nyilvánvalóan téves következtetés még akkor is, ha egyébként a tényadatok ilyen jellegű kombinációja éppenséggel erre utaló eredményt ad.

Vegyük észre, hogy a bemutatott példával ellentétben nem mindig ilyen nyilvánvaló a téves megközelítés, ezért is szeretnénk felhívni a figyelmet az adatértelmezés jelentőségére.

Meggyőződésünk, hogy az adatvagyon szakszerű felhasználásának kialakulása hosszú ideig tartó evolúciós folyamat, mely rendkívüli körülményt igényel – ugyanakkor az is nyilvánvaló, hogy a szűk kutatói kör által alkalmazva az adatvagyon által biztosított lehetőségek kiaknáztatlanok maradnak.

Létezik-e komplex adatvagyon-kezelő rendszer?

Előjáróban: a köznevelés területén belül az adatvagyon pedagógiai célú alkalmazása talán a legkevésbé kidolgozott, egyben a legizgalmasabb feladat is. Ugyanakkor a rendelkezés-

⁶ A Köznevelés Indikátor Menedzsment Rendszere (KIMER) az Oktatási Hivatal által a TÁMOP-3.1.8-09/1-2010-0004 Átfogó minőségfejlesztés a közoktatásban című kiemelt uniós projekt keretében készült. Fontos megjegyezni, hogy az időközben a kimer.oktatas.hu oldalon publikált rendszer végleges változata már nem tartalmazza a bemutatott táblázat előállításának lehetőségét.

re álló adatok között megfigyelhető szoros és szerteágazó összefüggésrendszer nem indokolja, hogy az adatvagyonot részekre szabdalva képzeljük el, és így önálló tanügyigazgatási, kutatási vagy akár önálló pedagógiai adatrendszert hozzunk létre. A továbbiakban az adatvagyonot egységes egészként kezeljük – miközben tisztában vagyunk az alkalmazási lehetőségek és jogosultságok széles spektrumával.

Ezek után nézzük a köznevelési adatvagyon létrehozásának logikai folyamatát, mely álláspontunk szerint a következő lépésekből áll:

1. Koncepció létrehozása és elfogadtatása
2. KIR Adatvagyon-kezelő Központ (továbbiakban: KIRAK) létrehozása
3. Adatvagyon kataszter létrehozása és feltöltése
4. Központi forrás-adatbázis kialakítása
5. Adatellenőrzési rendszer kialakítása
6. Adatfrissítő rendszer létrehozása
7. Adatkapcsolati háló kialakítása
8. Adatpublikációs megoldások kidolgozása
9. Finomhangolás
10. Információbiztonsági Irányítási Rendszer kialakítása, ISO 27001 tanúsítvány megszerzése

Az adatvagyon-létrehozás viszonylag egyszerű koncepcióra épül: a forrásadatok első sorban a keletkezési helyük által támasztott igényekre reflektálnak, ezért logikailag helyhez kötöttek. Ha az egyes alrendszerekből a forrásadatokat „kiszabadítjuk”, lehetővé válik egy komplex adatkapcsolati háló kialakítása. Ha tehát egy alrendszer a benne keletkező forrásadatokat „beadja a közösbe”, adott esetben ugyanazokat az adatokat már egy teljesen más kontextusban kapja vissza – kiegészülve az erőforrásként megjelenő, nem az adott alrendszerben keletkező adatokkal. A koncepció része továbbá, hogy statisztikai jelleg helyett minden esetben elemi szintű forrásadatokra épüljön az adatvagyon.

A koncepció elfogadtatása egyrészt felsővezetői döntést, majd a döntést követően folyamatos támogatást jelent. Másrészt a forrásadatokat „birtokló” alrendszer-gazdák számára világossá kell tenni az együttműködés előnyeit. Olyan adatvagyon-kezelő rendszert kell kiépíteni, ahol a haszon kézzelfoghatóvá válik számukra.

Az előzőekből egyenesen következik, hogy a forrásadatokat birtokló alrendszerek nem képesek és nem is lehet feladatuk az alrendszer független – nem az adatok létrehozására, hanem szerteágazó felhasználásukra koncentrálni – adatvagyon-kezelő rendszer létrehozása és működtetése. Erre a feladatra önálló szervezeti egységet kell létrehozni – lehetőség szerint a KIR adatkezelőjeként működő Oktatási Hivatalon belül. Fontos, hogy tisztán lássuk: a KIRAK (vagyis KIR Adatvagyon-kezelő Központ) nem azért jön létre, hogy az értékes adatokat kivegye az alrendszerek belső szakmai műhelyeinek a kezéből, hanem éppenséggel ennek az ellenkezője igaz: minden elérhető adatot szerezzen meg és tegye ezeket az adatvagyon részévé, ezáltal biztosítva az alrendszerek összes külső adatigényének teljes körű kielégítését garantáltan jó minőségű, széles kapcsolódási lehetőségekkel rendelkező adatokkal.

Az egyébként rendelkezésre álló adatok felhasználhatóságát nagymértékben korlátozza, hogy egyszerűen nem ismerjük az adatokat, és főleg nem a tulajdonságaikat. Az adatvagyon kataszter szerepe éppen az, hogy a köznevelés számára releváns adatokat kata-

lógusszerűen nyilvántartásba vegye, és minden érdeklődő számára kereshető formában megismerhetővé tegye.

A központi forrás-adatbázis adatszinkronizációs és -tárolási feladatokat lát el, az adatok közötti kapcsolati háló kiépítése már nem ezen a szinten történik.

Ha adatvagyon-hasznosításról beszélünk, az adatellenőrzési rendszer kialakítása, működtetése és folyamatos fejlesztése kiemelten fontos terület. Minden egyes tárolt adatról kivétel nélkül meg kell tudni mondani, hogy adatminőség szempontjából hogyan kategorizálható. Ugyanis hiú ábránd, egyben alapvető tervezési hiba lenne, ha feltétel nélkül megbíznánk a központi forrás-adatbázisba beérkező adatok minőségében. Az adatvagyon-kezelő rendszernek saját, önállóan működő hibadetektálási és protokollokra épülő hibakezelési rendszerre van szüksége, melynek működése az adatmező szintű értékészlet-problémák vizsgálatától kezdve egészen a kapcsolati háló segítségével elvégezhető globális anomáliák felfedéséig terjed.

Az adatvagyon természetesen nemcsak tökéletes minőségű adatokból áll, lehetőség szerint minden felhasználásra alkalmas adatot be kell gyűjtenünk – pontosan megadva az adatminőségre vonatkozó jellemzőket. Ezáltal az adatellenőrzési rendszer célja ket-
tős: megkísérelni a hiba kijavítását, szükség esetén az adatforrás bevonásával, illetve pontosan megjelölni a minőségi problémával rendelkező adatokat. További fontos követelmény, hogy a felhasználók gyakorlati tapasztalatait is visszacsatoljuk a rendszerbe, ezáltal minél szélesebb körű a felhasználás, annál nagyobb az esélyünk a hibák felfedésére és kijavítására.

Az adatfrissítő rendszer a központi forrás-adatbázisban tárolt adatok naprakészisé-
gért felel.

Az adatkapcsolati háló az adatvagyon központi eleme, amennyiben a tárolt adatok közötti kapcsolatok kiépítéséért és fenntartásáért felelős. (Az adatok közötti kapcsolatok kialakítási lehetőségeivel külön foglalkozunk.)

Az adatpublikációs rendszer célja, hogy a felhasználók számára a nekik megfelelő formátumban és feldolgozottsági szinten rendelkezésre álljon a szükséges adat. Míg egy kutató tipikusan további feldolgozásra alkalmas adatbázis-részletet tud letölteni, addig az intézményi felhasználók önkiszolgáló webes felületen keresztül jutnak adatokhoz, az alrendszerek felé pedig adatinterfészen keresztül történik az adatáramlás.

A finomhangolás lényegében folyamatosan végzett tevékenység, mely a külső és belső körülmények mindenkorai változásaihoz igazítja a rendszert.

A KIR 2005. és 2008. között rendelkezett ISO-tanúsítvánnyal. A köznevelési adatvagyon jelentőségét és értékét figyelembe véve célszerű feléleszteni ezeket a nemes hagyományokat.

Az adatkapcsolati háló létrehozásának lehetséges megoldásai

Az adatkapcsolati háló az adatvagyon „idegrendszere”, mely az adatvagyon alkalmazhatóságát – az adatok minőségén túlmenően – döntően befolyásolja. Az adatkapcsolati háló olyan logikai és fizikai réteg, mely az adatvagyon minden elemét a gyakorlatilag elérhető összes módon összefüggésbe hozza egymással. Az adatkapcsolati háló fontos tulajdonsága, hogy semmilyen módon nem függ össze az adatok keletkezési helyeként szolgáló alrendszerekkel, azoktól függetlenül létezik – többek között ezáltal is biztosítva az adatok „kiszabadítását”.

A különböző alrendszerek adatvagyon-felhasználással kapcsolatos igényeit végső soron az adatkapcsolati háló „mozgósításával” lehet kiszolgálni.

A folyamatosan változó felhasználási igények okán az adatkapcsolati háló is dinamikusan fejlődik – összhangban az adatok számának és minőségének növekedésével.

Az adatkapcsolati háló értelmezésünk szerint csak komplex adatvagyon-kezelő rendszerekben létezhet – amiről egyelőre még nem beszélhetünk –, ugyanakkor már most is megfigyelhető adatbázisok közvetlen összekapcsolása különböző köznevelési rendszerekben. Az adatbázisok összekapcsolására háromféle, alapvetően különböző módszert azonosíthatunk, melyek a következők:

- 1 A legkézenfekvőbb eset az országos kompetenciamérés, amikor a személyazonosításra alkalmas adatokat, valamint a Hivatalban keletkező és tárolt személyazonosításra nem alkalmas egyéni szintű adatokat a köznevelési intézmény helyben össze tudja kapcsolni. Ekkor jogosultság alapú adatkapcsolatról beszélhetünk, amennyiben a köznevelési intézmény tanulónként nevesítve megismerheti a mérési eredményeket. A Hivatal ezzel szemben nem jogosult az egyébként nála keletkező egyéni szintű mérési eredmények konkrét személyekhez kötésére, ezért ún. mérési azonosító alkalmazásával, személyazonosításra alkalmatlan módon kezeli az adatokat.
- 2 „A tanulók egészségi állapota iránti közös felelősség megélését nagyban támogatni tudja, ha a szülők hozzájárulnak ahhoz, hogy a testnevelők a diákok fitességi eredményeihez közvetlenül hozzákapcsolhassák a tanulók releváns személyes adatait is.”⁷ Miután a személyes adatok között az országos kompetenciamérésben alkalmazott mérési azonosító is megadható, ezért a fitességi adatok és a kompetenciamérés eredményei egyéni szinten, személyazonosításra alkalmas módon összekapcsolhatók. Az adatbázisok összekapcsolása tehát abban az esetben is lehetséges, ha erre az adatok tulajdonosa felhatalmazást ad.
- 3 Az előzőektől eltérő megoldás a KIR és a felsőoktatási információs rendszer, valamint az Országos Nyugdíjfolyósítási Igazgatóság adatbázisainak az összekapcsolása. Ebben az esetben egy harmadik, felhatalmazással rendelkező szervezet (Nemzeti Infokommunikációs Szolgáltató Zrt.) kapcsolati kód segítségével végezte az összekapcsolást, melynek eredményeképpen olyan egyéni szintű adatbázis jött létre, mely nem tartalmaz személyazonosításra alkalmas adatokat, így személyekhez elvileg sem köthető, ugyanakkor a benne szereplő egyéni szintű adatok okán rendkívül jól használhatónak tűnik.

Az adatbázisok összekapcsolása a felsorolt példák alapján többféle módszer szerint is történhet. Az adatkapcsolati háló hasonló funkciókat lát el, de nemcsak két adatbázis között, hanem az adatvagyon egészére vonatkozóan.

Az adatvagyon kezelését biztosító információs rendszer felépítése

Az adatvagyon-kezelés természetsszerűleg kiszolgáló információs rendszer létrehozását és működtetését feltételezi. Az információs rendszer felépítése többféle, egymásra épülő réteggel modellezhető.

⁷ Az idézet a NETFIT® rendszer honlapján (www.netfitweb.hu) elérhető tájékoztatóban szerepel, a 7. Miert jó, ha szülőként hozzájárulok gyermekem adatkezeléséhez? témakörnél.

Az adatgyűjtési réteg biztosítja a forrásadatok folyamatos beáramlását a rendszerbe – az adatforrások felépítésétől és földrajzi elhelyezkedésétől, valamint szervezeti hovatartozásától függetlenül. Miután köznevelési adatvagyonról beszélünk, a legfontosabb adatforrásnak az intézményi adminisztrációs rendszereket tekintjük, melyek főtevékenységként az intézmények helyi igényeinek teljes körű kiszolgálására hivatottak.

Fontos megjegyezni, hogy a forrásadatok sok esetben valamilyen tanügyigazgatási rendszer működése során „melléktermékként” keletkeznek, vagyis kifejezetten nem optimalizáltak univerzális adatvagyon-hasznosításra – ettől függetlenül persze még egy tanügyigazgatási rendszer teljes mértékben képes betölteni jogszabályban előírt feladatát. Ha egy ilyen rendszer forrásadataira szükségünk van – márpedig minden adatra szükségünk van –, akkor kiemelt feladatunk az adatminőség kontrollálása és hibajavítási folyamatok beépítése, mely a rendszer következő rétegének tekinthető.

A harmadik réteg a begyűjtött adatvagyon-elemek fizikai tárolására szolgál.

A következő logikai rétegnek az adatkapcsolati hálót tekinthetjük, melyről korábban már volt szó.

Végül a „legfelső” réteg az adat- és információ-publikációs réteg, mely a felhasználók igényeinek teljes körű kiszolgálására hivatott.

Az adatvagyon-kezelést biztosító információs rendszer leírása alapján azt hihetnénk, hogy ugyan a rendszer létrehozása sok munkával jár, viszont a különböző rétegek kialakítása és egymásra építése mégsem túlzottan bonyolult feladat. Ha pedig ez így van, akkor máris kezdjük neki: minden létező adatot szerezzünk be és hozzuk létre a köznevelési adatvagyon.

Ezzel szemben az az igazság, hogy valójában rendkívül összetett problémahalmazzal állunk szemben, különösen, ha figyelembe vesszük a köznevelés korábban is említett specialitásait.

Meggyőződésünk szerint egyetlen járható út létezik: evolúciós szemlélettel kezeljük a köznevelési adatvagyon létrehozását, vagyis nem a nagy egész gyors létrehozására törekszünk, hanem azokat a folyamatokat és tevékenységrendszereket definiáljuk, majd folyamatosan finomítjuk, melyek alkalmazása – „icipici”, de jól definiált kezdő állapotból kiindulva – elvezet a rendszer teljes körű kiépítéséig. Ellenkező esetben rövid időn belül egy minden adatot tartalmazó, de valójában használhatatlan és fenntarthatatlan rendszerrel fogunk szembesülni.

Az evolúciós szemléletű építkezés további előnye lehet, hogy mivel – valljuk be őszintén – jelenleg nem rendelkezünk a pedagógiai célú alkalmazhatóságához szükséges tudással, az adatvagyon kialakításával és folyamatos fejlesztésével párhuzamosan az alkalmazás lehetőségeit és korlátait is megtanulhatjuk. Törvényszerűnek tűnik, hogy az evolúciós folyamat gyakran fog visszalépésekkel járni, amikor kénytelenek vagyunk felülvizsgálni korábban biztosnak hitt elképzeléseinket, mert az eredmények nem igazolják vissza a fejlesztési irányok kiválasztását.

Az adatvagyon pedagógiai célú felhasználásának jövője

A jövő többféle forgatókönyvet is tartogat számunkra. Nézzük az ideális esetet: a közeljövőben létrejön a köznevelési adatvagyon és ennek a menedzselésére hivatott KIR Adatvagyon-kezelő Központ (KIRAK), mely többek között hatékonyan támogatja az egyéni

adatok pedagógiai célú felhasználását és az ezzel kapcsolatos oktatáskutatói tevékenységeket, de természetesen ezen túlmenően kiszolgálja a tanügyigazgatás teljes adatigényét – a döntés-előkészítéstől kezdve egészen a beválás-vizsgálatokig.

A pedagógia gyermek és tanulóközpontú komplex megközelítést alkalmaz, melynek integráns eleme a pedagógusok adatvagyon-felhasználásra épülő tevékenysége – szakértők és oktatáskutatók által „kívülről” támogatva. Az egyéni szintű idősoros adatok jelentősen megnövelik az egyéni differenciálás lehetőségeit, egyben a pedagógiai tevékenységek hatékonyságát és eredményességét.

Hogyan foglалhatók össze a tanulmányban leírtak?

A Köznevelés az adatok büvöletében elfogult írás.

Az adatfelhasználás pártján áll.

Hogy gyűjtjük, feldolgozzuk, ellenőrizzük, tároljuk, értelmezzük és alkalmazzuk őket.

Tudatosan, tervszerűen, széleskörűen, a korlátokra és kockázatokra figyelve.

IRODALOM

HALÁSZ G. (2010): *Beszámoló az OECD oktatási információs rendszerekről tartott munkaértekezletéről*, 2010. október 14-15, New York. Kézirat

LANNERT J. (2004): Hatékonyság, eredményesség és méltányosság. *Új Pedagógiai Szemle*, 54. évf. 12. sz., 2004. december.

HALÁSZ G. (2007): *Tényekre alapozott oktatáspolitikai. Háttér tanulmány az Oktatási Kerekasztal munkájához*. Kézirat. http://halaszg.ofi.hu/download/Evidence_based_study.pdf

BARABÁSI A. L. (2013): *Behálózva – a hálózatok új tudománya*. Helikon Kiadó Kft., Budapest.

BOGÁROMI ESZTER – MÁTH ANDRÁS

Sokforrású adatbázis-építés – buktatók, nehézségek, megoldási kísérletek

Esettanulmány

Elemezni kell, információkra valamennyi döntéshozónak szüksége van ahhoz, hogy a saját területén érdemi, megalapozott döntéseket hozzon. A kutató hivatása éppen ezen döntések megalapozása, hátterének kialakítása és támogatása. A döntéselőkészítéshez szükséges anyagok, információk gyűjtése és csoportosítása egyre inkább eltolódik a megkérdezéses adatgyűjtés felől a megfigyeléses információszerzés, valamint a sokforrású adatbázisok összekapcsolása felé. Ennek oka nem elsősorban a gyorsabb elemzésben kereshető (gyakran a felmerülő nehézségek akár még több időt is igényelhetnek egy egyszerűbb megkérdezéses vizsgálatnál), hanem az elemzési szint mélysége (túl sok helyen és túl sok embert kellene kérdezni), valamint a felhasznált információk nonkonverzatív jellege miatt. Az adatbázisok forrása, struktúrája, időbeli érvényessége és a lehetséges kapcsolódási pontok egymáshoz rendelése minden esetben eltérő megfontolásokat követelnek meg – éppen ezért ezek ismerete, aktív tudása és érvényesítése a kulcsa egy-egy jól lezárt és értékes kutatásnak.

A kutatás szó használatánál talán érdemes is megállni egy pillanatra. Ennek a kifejezésnek az értelme a társadalom-kutatásban és a piackutatásban egyre inkább megváltozik. Amíg korábban teljesen természetes volt, hogy kutatáson valamilyen adatgyűjtést (értsd: kérdezés, aktív részvétel) és annak feldolgozását, elemzését értettünk, addig 2015-re eljutottunk oda, hogy a gyakorlat egyre kisebb szeletét teszi ki az ilyen jellegű vizsgálati módszerek sora – a megfigyelések, passzív adatgyűjtések segítségével létrejövő adatbázisok és értékelések javára.¹ Ez a folyamat a „kutatás” kifejezés fokozatos megváltozását is magával hozza, hiszen nagyon sokan még manapság is kutatáson az adatfelvé-

¹ Ugyanakkor azt is fontos leszögezni, hogy az elemzési szempontokat nem kezelő adatgyűjtések (és jelenleg még inkább ezek vannak túlsúlyban) értéktelen – bár sokszor irdatlan méretű – adatbázisokat eredményeznek).

► *Educatio* 2015/3. Bogáromi Eszter – Máth András: *Sokforrású adatbázis-építés – buktatók, nehézségek, megoldási kísérletek*, 86–97. pp.

telt érték; holott – főként a fiatalabb nemzedék számára – már inkább az elemzés, értékelés jelentést fed le.

Az alábbiakban összegzett esettanulmány arra szolgál példaként, hogy egy komplex elemzési feladatban milyen akadályok merültek fel, sok, egymástól számottevően eltérő céllal, tartalommal készült adatbázis összeépítésénél, és milyen megoldások vezettek el egy sikeres elemzés kivitelezéséhez – immár egy komplex értékelési környezet felállítását követően. A sikert elsősorban azon mérhettük le, hogy a döntéshozók milyen gyakorisággal térnek vissza az elemzésre, milyen későbbi döntésekben köszönnek vissza az ott elkészített szegmensek, települési értékelések.

Az esettanulmányból bizonyos konkrét információkat kénytelenek voltunk mellőzni vagy megváltoztatni a kutatás egyedisége miatt – de ez a folyamat vagy a felmerült problémák megértése szempontjából, reméljük, nem okoz nehézséget.

A feladat

Egy magyar nagyvállalat termékportfóliója átrendeződésekor lényegesen nagyobb figyelemmel fordult egyes megyék települései felé. Ez az üzleti változás generálta azt az igényt a cég vezetői részéről, hogy szükségük volt a magyar településszerkezet mélyebb ismeretére (ehhez még nem kellett volna adatbázisokat összeépíteni, mert erről vannak kiváló tanulmányok), de amikor a beszélgetésekben az is kiderült, hogy az egyes üzleti aktivitások tervezéséhez településtípológiákat kell majd alkotni, akkorra vált egyértelművé, hogy a lehetőségek szerint legtöbb elérhető adatbázist kell felhasználni a munka során.

Tegyünk itt egy rövid kitérőt az üzleti világba, hogy jobban megérthessük, mire és hogyan volt szükséges a megrendelőnek, vagyis mire akarta használni a végeredményeket.

Az elsődleges cél és elképzelés az volt, hogy a szűkös üzletszerzési (marketing) erőforrásokat olyan területekre érdemes koncentrálni, ahol azok nagyobb valószínűséggel hoznak növekvő eladásokat – vagyis a kutatásnak célpontokat kellett kijelölnie a megrendelő számára. Ehhez természetesen azonosítottunk értékelési paramétereket – mint jövedelmi viszonyszámok, népességarányok és továbbiak. Ez a feladat arra is szolgált, hogy a meglévő értékesítési eredményeket értékelni lehessen – egyfajta elvart/megvalósult mutatószámokat is fel lehessen vázolni – mintegy megmutatva a területen dolgozó munkatársak hatékonyságát.

A másik megfogalmazott igény a hasonlóságok meghatározása volt – vagyis hogy az egyes tesztkörnyezetben végzett aktivitások és az azokra adott reakciók után meg lehessen találni, hogy milyen más településeken várható hasonló reakció a fogyasztók részéről. Könnyű belátni ezen kérdés hasznosságát: számottevően lehet csökkenteni a rosszul vagy feleslegesen elköltött marketingkiadások mértékét egy-egy próba kivitelezésével. Ha valami sikeres egy településen, akkor azt meg lehet kísérelni más helyen is – ha tudjuk, hogy mely települések tekinthetők a szociokulturális, gazdasági és más egyéb szempontok szerint hasonlóknak.

Ezekből a megfontolásokból kiindulva kettős feladatot foglalmaztunk meg a kutatás során: (A) azonosítsuk és „gyűrjük össze” lehetőleg egy mutatóba azokat a mérőszámokat, amelyek képesek jellemezni egy-egy települést az üzleti potenciál szempontjából; (B) alkossunk olyan összetett adatbázist, amely képes arra, hogy egymástól akár földrajzilag jelentősen távol eső településeket is azonos szegmensbe soroljon, a lehető legtöbb ér-

vényes gazdasági, szociokulturális, illetve fogyasztási adat alapján. A továbbiakban lépésként mutatjuk be ennek a feladatnak a megvalósítását.

Az adatbázisok kiválasztása

Amikor sorra vettük a lehetséges dimenziókat, melyek mentén szerettük volna felépíteni az általunk használt modellt, fontosnak tartottuk, hogy ne gondoljunk a hozzáférhetőségre. Tudtuk, hogy vannak adatbázisok, melyek könnyen érhetőek el, másokat létre kell hozni, és megint másokat nem is lehet előállítani. A különböző adatbázisok összekapcsolásának és elemzésének fontos pillanata, amikor még úgy gondoljuk, minden adat elérhető. Ezt a logikát felborítva, jelen tanulmányban nem a valóságban megtörtént adatgyűjtést mutatjuk be, hanem csoportosítva elemezzük a felhasznált forrásokat. Tanulmányunkban a bemutatott adatbázisokat három csoportba soroljuk: ingyenesen hozzáférhető adatbázisok, saját vagy fizetett adatbázisok és az elemzés céljára rögzített adatbázisok.

Az ingyenesen hozzáférhető adatbázisok köre

Az ingyenesen hozzáférhető adatbázisokba azok az adatbázisok tartoznak, melyek már adatbázisba szerkesztve érhetőek el és mindenki számára ingyenesek. Az adatok felhasználásának céljától is függ a termékek árazása. A tudományos céllal felhasznált adatok, egyedi adatszolgáltatások kutatók és egyetemi hallgatók részére gyakran ingyenesen érhetőek el. Elemzésünket piaci felhasználásra készítettük, így ebbe a kategóriába csak azok az adatbázisok tartoznak, melyek mindenki számára ingyenesek. A adatbázisba szerkesztettnek tekintjük azokat az adatforrásokat, melyeket táblázatos formában lehet menteni. Értelmszerűen nem tartoznak ide az interaktív grafikonok és térképek. Utóbbiak az eredmények bemutatásánál lehetnek hasznosak, könnyen értelmezhetővé és látványossá tehetik a prezentálást.

A legszélesebb skálán a *Központi Statisztikai Hivatal* (KSH) mutatói mozognak, melyek a Tájékoztatási Adatbázis Területi statisztikái között találhatóak.² A területi statisztikák között négy forrás mutatói közül válogathatunk. Az *Éves településsziszteztikai adatok a 2013-as településszerkeztben* adatbázis mindig az adott időszakban legfrissebb mutatókat tartalmazza Magyarország minden településére. A nagyobb települések irányítószám szerint nincsenek bontva, aggregált eredmények találhatóak a kimutatásban. Kutatásunk idején a 2012-es adatok voltak elérhetőek. Külön adatbázisban érhetőek el *Budapest kerületeinek adatai*, az egyéb települések adataihoz igazodó frissülésben. Végül a területi statisztikák között találhatóak a *Népszámlálási adatok járárok és Budapest kerületei* szerint. Mivel feladatunk a Budapesten kívüli települések jellemzése volt, a budapesti kerületekre és a járárokra vonatkozó adatbázisokat nem használtuk. A KSH területi statisztikáit tartalmazó online felületen könnyen lehívhatóak az adatok; magunk válogathatjuk össze, hogy milyen dimenziókat szánunk a sorokba, oszlopokba és munkalapokra. Az eredményeket végül menthetjük és exportálhatjuk több formátumban is (grafikon, pdf és excel).

² Forrás: <http://statinfo.ksh.hu/Statinfo/themeSelector.jsp?page=2&szst=T>

A településstatistika tizenhat témakörben kínál több száz mutatót. A témakörök többek között kiterjednek az alap demográfiai kimutatásokra, lakásállományra, oktatásra, turizmusra, közművekre, munkanélküliségre, önkormányzati segélyezésre is. A mutatók éves aggregált adatokat tartalmaznak régiókra, megyékre és településekre.

A *Népszámlálás 2011* adatai a KSH által hozzáférhetőek település szinten is a Népszámlálás saját felületén.³ Megyék szerint külön-külön letölthetőek a népességre vonatkozó, településsoros adatok. A megyéken belül külön kimutatás található az egyes településtípusokra, járásokra, kistérségekre, települések mérete szerinti csoportokra és az egyes településekre.

A mindenki számára hozzáférhető adatsorok közé tartoznak az *Országgyűlési választások* és az *Önkormányzati választások eredményei*. Elemzésünk idején az 2010-es országgyűlési választások eredményei számítottak a legfrissebb eredménynek, ezért a listás szavazati arányokat használtuk fel a települések jellemzésénél.⁴ Az Országgyűlési választások internetes oldalán számos szempont szerinti bontásban érhetőek el az eredmények, így az egyéni választókerületek eredményei, a területi választókerületek eredményei mind listás szavazatok, mind elnyert mandátumok szerint, valamint a pártok eredményei területi listákon. Továbbá egészen a szavazóköri eredményekig bontott eredményeket is megismerhetjük.⁵ Az oldal vélt funkciója egy-egy választókerület vagy település eredményeinek megismerése és nem országos adatok exportálása, így aki arra vállalkozik, hogy minden településre megismerje az eredményeket, annak a kutatás időtervének kialakításánál tervezni kell a hosszabb adatbeszerzési idővel.

A nyilvánosan elérhető adatoknak egyik körét alkotják azok a kimutatások, melyekhez csak igénylés útján lehet hozzájutni. Az igénylésnél a kutató által megadott szempontok jelennek meg elsősorban, de előfordulhat, hogy a kért kimutatások nem állíthatók elő, esetleg nem is rögzítettek. Elemzésünkhöz a bűnelkövetők statisztikáit kívántuk beszerezni, és azon belül is arra voltunk kíváncsiak, hogy a nők ellen elkövetett bűncselekmények száma mennyire jellemző az adott településre. A *bűncselekményekre vonatkozó adatokat* a KSH Népesedési és szociális védelmi statisztikai főosztályáról igényeltük. A főosztálytól megyei bontásban kaptunk adatokat, majd a Belügyminisztériumhoz irányítottak minket, ahonnan a lehető legrészletesebb adatokat lehet igényelni a bűnelkövetés témakörében. Településszinten bontott adatokhoz így a *Belügyminisztériumon* keresztül jutottunk, de további nehézségekbe ütköztünk, melyekről a következő fejezetben írunk. Az adatok igénylésénél a kutatónak pontosan meg kell határoznia, hogy milyen kimutatásban, milyen mutatókra van szüksége, más esetben az igénylési folyamat elhúzódhat. A bűncselekményeken túl más témakörben is igényelhetők adatok az egyes minisztériumoktól és más állami intézményektől.

Nem nyilvános adatsorok

A nyilvánosan nem elérhető adatsorok fontos eleme volt a megbízó (Kiadó) *értékesítési adatai*, melyekkel a települések anyagi helyzetét, kulturális nyitottságát és fogyasztási szokásait jellemeztük. A vállalatok napról napra bővülő, frissülő adatsorok birtokában

³ Forrás: http://www.ksh.hu/nepszamlalas/tablak_teruleti_04

⁴ Forrás: http://www.valasztas.hu/hu/parval2010/354/354_0_index.html

⁵ Forrás: http://www.valasztas.hu/hu/parval2010/354/354_0_index.html

vannak. A különböző szervezeti egységek más és más bontásban, tematikában állítják elő és kéri be az egyes mutatókat, melyek belső elemzéseknél felhasználhatóak. Ezeknek az adatoknak a felhasználása nehezebb, mint a beszerzésük. El kellett döntenünk, hogy milyen időszakra és milyen mutatókra vonatkozó adatsorokat illesztünk az adatbázisunkba; az egyes mutatók átlagait vagy kumulált eredményeit számoljuk az egyes településekre. A választás sokszínűségét mutatja, hogy vannak vállalatok, ahol napi, vagy akár órás bontásban állnak rendelkezésre árushely szintű adatok, minden termékre külön lebontva, illetve vannak vállalatok, ahol csak havi bontásban, nagyobb földrajzi egységekre nézve elemezhetőek az eredmények. Minél részletesebb és kisebb időegységekre vonatkozó adatokkal dolgozunk, annál tudatosabban és módszertanilag alátámasztottan kell megválasztani az elemzési egységeket, mutatókat. Saját kutatásunkban indokolt volt a termékek külön mutatóban való megjelenítése az eltérő olvasóközönség és a termékekhez kapcsolódó eltérő fogyasztási szokások miatt. Az eltérő termékekhez eltérő periodicitás is tartozik, így külön kellett döntést hozni a napilapok bemutatására és a heti- vagy havilapokéra. Átlagos példányszámokkal dolgoztunk, melyeket ezer főre vetítettünk ki. Az árushelyi eladások mellett külön változóként az előfizetéseket is beépítettük az adatbázisunkba.

A nem ingyenesen hozzáférhető adatsorok közé tartozik a *GfK Vásárlóerő kutatása* is, mely számunkra két értékes adatbázissal szolgált, egyfelől az irányítószámokra, másrészt a településekre számolt vásárlóerővel. A két adatbázis két külön adatsorként szolgál. A vásárlóerő-kutatás eredményei évente jelennek meg. A kutatásban öt dimenzió jelenik meg, a népességszám (főben és ezrelékben), a háztartásszám (darabban és ezrelékben), az összes vásárlóerő (millió euróban és ezrelékben), az egy főre jutó vásárlóerő és az egy háztartásra jutó vásárlóerő. A vásárlóerő-dimenziók három mutatót foglalnak magukba, az euró értéket, a magyarországi indexet és az európai indexet. Elemzésünkben az egy háztartásra jutó vásárlóerővel dolgoztunk. Ezzel a mutatóval a települések gazdasági erejét és a lakók gazdasági helyzetét kívántuk ábrázolni.

A települések jellemzésénél a televíziós nézettségi adatok beemelését is fontosnak tartottuk. A *Nielsen Közönségmérés* 1120 háztartásból álló panelen végzi a hivatalos nézettségi mérést. A panel nagysága és reprezentativitása csak a régiós szintű adatok használatát teszi lehetővé, ezért elemzésünkben ezeket a kimutatásokat nem tudtuk használni. Olyan kutatásoknál, melyek nagyobb földrajzi egységekre vagy nagyobb létszámú célcsoportokra vonatkoznak, a Nielsen Közönségmérés megfelelő háttérrel biztosíthat. A felismerés után sem szerettük volna kihagyni az elemzésünkéből a televíziózással töltött időt mint mutatót, ezért más adatforrás felkutatása mellett döntöttünk. Hazánkban egyedül a Nielsen Közönségmérés méri szoftver segítségével a televíziózást, de más elemzésekben – önbevallásos módon – megtalálható a keresett kérdés.

A *Nemzeti Olvasottság Kutatásban* (későbbiekben NOK) is kitérnek arra, hogy egy átlagos hétköznapon és egy átlagos hétvégi napon átlagosan mennyi időt töltenek tévénézéssel. A Nemzeti Olvasottság Kutatást – a *TGI (Target Group Index)* termékkel összekapcsolva – a TNS Hoffmann végzi folyamatos adatfelvétel mellett. Évente négy alkalommal adják ki a legfrissebb olvasottsági adatokat, évente két alkalommal a TGI termékhez tartozó mutatókat. A két kutatásban összesen több száz mutató érhető el, melyek kiterjednek a média területére, az FMCG termékek használatára, vásárlására, különböző attitűd-állításokra, érdeklődési körökre, pénzügyi termékekre és bizonyos életvitelt érintő kérdésekre. Mind a NOK, mind a TGI kutatásban is elérhetőek bizonyos szeg-

mentált változók, így az ESOMAR státusz változó, a TGI fogyasztói státusz, TGI Early Adopter szegmentáció, az életszakaszokat megkülönböztető változó és az informatikai ellátottságra vonatkozó mutató is. Évente a NOK-kutatásban részt vevő válaszolók száma eléri a 25.000 főt, így kisebb célcsoportok is biztosan vizsgálhatók a kutatásban. A TGI adatfelvétele személyes, papír alapú kérdőívvel történik a megkérdezett otthonában, míg a NOK kérdőívet kérdezőbiztos kérdezi le (laptop használatával) a megkérdezett otthonában. A személyek kiválasztásánál valószínűségi véletlen mintát alkalmaznak. Településszintű kimutatások nem érhetőek el a TGI-NOK kutatásban, viszont településméret, megye és régió változó használható az elemzések alkalmával. A TGI és NOK termékek változói közül – a televíziónézésen túl – az elemzésünk során használtuk a különböző lapkategóriák olvasási adatait, az attitűd-állításokat is.

Elemzés céljára rögzített adatbázisok

A hozzáférhető adatbázisok nem feltétlenül tartalmaznak minden változót és adatsort, amit a kutatás során a modellbe be kíván építeni a kutató. A hozzáférhetetlen adatsorokat akár elő is lehet állítani az elemzés céljára. Kutatásunkban két adatsort állítottunk elő, az egyik a közlekedést tartalmazta, a másik a településeken elérhető áruházak listáját.

A megfelelő *közlekedés hiánya* elszeparálhatja a kisebb települések lakóit, illetve a könnyű és dinamikus közlekedés elősegítheti az ingázást a munkahelyekre és serkentheti a szabadidős tevékenységeket is. Az első kérdés, ami felmerült ennél a dimenzióknál, a települések beosztása két csoportba. Az egyik csoportba soroltuk azokat a településeket, ahová el kívánnak jutni más településekről a lakók, a másik csoportban találhatóak azok a települések, ahonnan elutaznak – akár nap nap után – más településekre. Közigazgatási szempontból a kisebb települések a kistérségi központjukhoz tartoznak, itt találhatóak az okmányirodák, ahol a hivatalos iratokat kell kiállítani, az életvitelt tekintve viszont megeshet, hogy a kistérségi központtól eltérő, közelebbi városba járnak be szórakozni, vásárolni, dolgozni a település lakói. Magyarország összes településénél nem lehetett egyesével mérlegelni a lehetséges ingázási helyeket, így a közigazgatást követve, a kistérségi központokat jelöltük meg a 'cél-városoknak', a kistérségben található többi települést pedig 'kiinduló városoknak'. A városok közötti közlekedést több mutató rögzítésével terveztük megragadni, így mind a vonatközlekedést, mind a távolsági busz közlekedést rögzítettük. A közlekedésnél külön változót alakítottunk ki az induló járatok számának és a menetidőnek is, illetve külön változóban rögzítettük a két település közötti távolságot. A képzett változónk, mely a közlekedés könnyedségét mutatta, azt a közlekedési módot tartalmazta, mely a két település között a domináns lehet. Dominánsnak tekintjük az a közlekedési módot, ahol gyakrabban jár a jármű és rövidebb a menetidő. Végül a menetidőt, a járatok számát kombináltuk a két település távolságával, így azt is megtudhattuk, hogy mennyivel nehezebb tömegközlekedéssel eljutni a kistérségi központba, mint saját autóval. A képzett mutató értékei széles skálán mozogtak.

Az *áruházak számára és fajtájára* azért volt szükségünk, mert így a termékekhez való hozzájutást ismerhetjük meg, mely hatással van mind a fogyasztási szokásokra, mind a közlekedésre. Számos áruházláncot vizsgáltunk meg és rögzítettük telephelyeiket, így minden településnél tudtuk, hogy hány áruház található benne és az áruházak mekkora alapterületűek. Az áruházak minden típusa külön változóba került, majd egy képzett mutatót készítettünk, mely a település ellátottságát mutatta.

Buktatók és megoldáskeresés

Az adatbázisok begyűjtése után kezdődött a munka második fázisa, melyben egy adatbázisba kellett importálni a különböző adatsorokat. Az elemzés célja az egyes települések jellemzése volt, így az esetek értelemszerűen csakis a települések lehettek. Az adatsorok összekapcsolása számos nehézség elé állította a kutatókat. A későbbiekben felsorolt problémák közül sok azért nehezíti meg az elemzést, mert megoldásukra nagyon hosszú időt kell fordítani és egyszerű programozással nem oldható meg, a személyes ellenőrzés elkerülhetetlen, ami tovább növeli az elemzésre fordított időkeretet. Fontosnak tartjuk kiemelni, hogy az idő a kutatásnál egy kiemelt fontosságú dimenzió volt, mert az egyes adatbázisok eltérő időpontokra (is) vonatkoztathatnak.

Az első probléma a Nemzeti Olvasottság Kutatás felhasználásánál merült fel, ugyanis ebben az adatbázisban nincsenek a településekre vonatkozó becsült változók. Viszont a minta kellően magas elemszámú, az adatfelvétel háttérben lévő módszertan megbízható és a kutatásban felvonultatott változók egyedülállóak. A hiányosságot a megyék és a településnagyság alkalmazásával hidaltuk át. Egy képzett változót alakítottunk ki, mely a két változó kereszteződéséből született. A képzett változó értékeit az adatbázisunkban is kialakítottuk minden településre, így kapcsoltuk össze a két adatbázist. Több száz változót kapcsoltunk a NOK-kutatásból az adatbázisunkhoz, és az elemzés szakaszára hagytuk a döntést arról, hogy mely változó használható fel és melyik nem.

A második felmerülő kérdés a különböző adatsorok előállításának *időbeli eltérése* volt. Míg a népszámlálást tízévente végzik, addig a terjesztési adatok napról napra, az internetes adatsorok pedig percről percre frissülnek. A területi statisztikák adatfelvételének időpontja az elemzést két évvel előzte meg. Eltérő periodicitás és eltérő adatfelvételi időpont tartozott az egyes adatbázisokhoz. Amennyiben szigorú módszertanhoz tartottuk volna magunkat, nem lett volna két adatbázis, amelyet össze tudtunk volna kapcsolni. Külső adatforrások egyesítésénél a kutatónak meg kell hoznia azt a döntést, hogy milyen időintervallumban fogadja el a különböző adatsorokat. Ez a döntés függ a mutatók érzékenységétől, attól, hogy mennyire változhat az értékük hónapok, évek alatt. Legrégibbi adatok a népszámláláshoz tartoztak. A népszámlálás adatbázisából a vallási megoszlás-mutatókat használtuk. Az egyének felekezeti váltása nem annyira jellemző, inkább a településekről való elvándorlás és az oda való bevándorlás lehet jelentősebb hatással a vallásos lakók arányára. Évente lehet nyomon követni a települések halálozási, születési és vándorlási számait, így tesztelhető, hogy melyik településen volt jelentősebb lakosságcsere. Elemzésünk idején a 2010-es népszámlálás volt a legfrissebb adatfelvétel, a KSH adatsorok adatfelvétele 2012-ben történt. A településeknél jelentősebb népességváltozás nem volt megfigyelhető, ami a vallásra vonatkozó mutatókat befolyásolta volna, ezért elfogadtuk, hogy a Népszámlálás 2010 adatsorait használjuk.

Az Országgyűlési választások eredményei is négyévente érhetőek el, ennél az adatbázisnál is a 2010-es adatsorral dolgoztunk.⁶ A Központi Statisztikai Hivatal jelentései 2012-es adatokra vonatkoztak, így ezeknél az adatsoroknál a 2012-es év eredményeit illesztettük az adatbázisunkba. A saját rögzítésű adatsorainknál nem állt rendelkezésünkre régebbi állapotra vonatkozó adatsor (Volán menetrend, MÁV menetrend, áruhá-

⁶ A kutatás tervezése és előkészítése 2014 tavaszán történt, így nem volt mód a 2014-es eredményeket beépíteni a rendszerbe.

zak listája), így ezeknél az adatsoroknál csak a legfrissebb állapotokat tudtuk megjeleníteni az adatbázisban.

Amennyiben minden adatsort elemzünk, jól látható, hogy az adatfelvétel időpontját tekintve akár három év eltérés is lehet két mutató között. De nem csupán az eltérést kell figyelembe venni, azt is kell mérlegelni, hogy az adott mutató értékeiben milyen mértékű változás állhat be az eltelt idő alatt. Azokat a mutatókat tartottuk meg az adatbázisban, melyeknél az eltérés feltételezett becslése még az elfogadható intervallumon belül volt. Nyilvánvaló, hogy az elfogadottság mértéke minden egyes kutatásnál az elemzési céloktól függ, annak fényében kell egyedi döntéseket hozni.

A következő probléma mellyel szembesültünk, a *települések listája* volt. Az egyes adatbázisokban nem egységes a településlista, eltérő írásmóddal szerepelnek a nevek. A kisebb települések neve változik, több település kötőjellel tartalmaz kiegészítő neveket. Ráadásul vannak adatsorok, melyekben a településnévhez számokat, kódokat is csatolnak, vagy éppen a nemzetközi szoftverek használata miatt ékezetek nélkül írják a neveket.

Az eltérő írásmód mellett a települések listája sem azonos. Alapvetően két megközelítésből lehet listázni a településeket, az egyik az *irányítószám alapú* megközelítés, a második a név szerinti. A két lista nem minden esetben fedi egymást. Az eltérés a legkisebb és legnagyobb települések esetén figyelhető meg. A nagyobb településeknél egy településen belül több irányítószám is megtalálható, míg kisebb településeknél egy irányítószám alá több település is tartozhat. A szállításnál, árusításnál ez nem probléma, viszont egy adatbázis-összevonásnál hibaüzenetként jelenik meg a település. Ezeket a „hibákat” egyesével kell ellenőrizni, besorolni. A besorolás viszont nem minden esetben egyszerű feladat. Mutatóként kell mérlegelni, hogy a két település értékeit hogyan redukáljuk egy esetre – átlagolással, összeadással, ezer főre számolással stb. Nem minden esetben egyértelmű a döntés és nem minden esetben áll rendelkezésre minden adat. Amennyiben ezer főre számoljuk a mutató értékét, akkor lehet, hogy nem áll rendelkezésünkre az egyik település népességének száma, ebben az esetben viszont esetleg olyan adatokat kell használnunk, melyek forrása és adatfelvételi ideje eltérő. Esetünkben nem kellett ilyen kompromisszumokat kötni, de ez csak a saját adatbázis felépítésének volt köszönhető.

A kisebb települések összevonása nemcsak az irányítószámoknál jelenik meg, hanem az élet számos más területén is. Kisebb településeknél előfordul, hogy csak az egyikben van *orvosi rendelés*, és az három település lakóit szolgálja ki. Az egy év alatt összesen el látott járóbetegek száma ebben az esetben a három település összes járóbetegére vonatkozik és nemcsak az adott településére. Természetesen lakóhely szerinti bontásban nem áll rendelkezésünkre az előbb említett mutató; ekkor felmerül a kérdés, hogy felhasználható-e a mutató. Arra biztosan nem használható, hogy a lakosság lélekszámahoz arányosítva tudjuk meg a betegek számát. El kell tehát döntenünk, hogy mire használjuk a mutatót. Amennyiben a település fejlettségét, érzékelt nagyságát kívánjuk megfogni, akkor informatív lehet ez a mutató, melynek értéke nulla, ahol nincs rendelés. Illetve ha a járóbeteg fogasztási szokásokat, forgalmat kívánjuk ábrázolni, elemezni – mint például a gyógyszerforgalom, közlekedés – akkor is informatív lehet a mutató. Az egyes mutatók felhasználásáról minden esetben a kutatónak kell tehát döntenie. Az egyes mutatók értelmezése és felhasználása szerteágazó lehet. A felhasználást ilyen esetekben tudományos értekezésekben pontosan kell definiálni és alátámasztani. Mivel mi az egyes betegségekre voltunk kíváncsiak, így az egyes települések egészségügyi állapotát kívántuk kimutatni, ezt a változót ki kellett emelnünk az adatbázisunkból. A kórházak, kórházi

ágyak számát, szakrendelések számát mutató változókból képzett új változót alkottunk, mely a település fejlettségi szintjét volt hivatott árnyalni.

Az adatbázis rögzítése után, az elemzésnél talákoztunk a következő szembevető problémával, mely a *települések lélekszámából* adódott. Magyarországon pár tíz fős, vagy éppen száz fős települések is vannak, és minél kisebb a település, annál szélsőségesebb értékei lehetnek az egyes mutatóknak. Így találkozhatunk olyan településsel, ahol az átlagos 50 százalékos férfi arányt messze elhagyva 75 százalékos a férfiak aránya. Ezek a jelentős eltérések a klaszteranalízis, főkomponens-analízis és más statisztikai eljárások alkalmazásával borítják fel a modellt. Gyakorlatilag lehetetlenné teszik az érdemi eltérések kimutatását. Ebben az esetben el kell döntenünk, hogy véglegesen kiemeljük a kisebb településeket az adatbázisból, vagy egy külön adatbázisban kezeljük azokat. Bármelyik megoldást is választjuk, mindkét esetben meg kell határozni az lélekszámban mért határt, mely alatt kiemeljük a településeket.

A települések lélekszámahoz kapcsolódik a másik véglet vizsgálata is. A legnagyobb városok városrészei számos adatsorban megjelennek külön értékekkel (ahol irányítószámot is használnak az azonosításnál, vagy kerületeket, városrészeket). Ezt a megkülönböztetést azért is alkalmazzák a hivataloknál, vállalatoknál, mert viselkedésükben, fogyasztási szokásaikban, életvitelükben, szociodemográfiai összetételükben jelentősen eltérőek lehetnek az egyes városrészek. A kutatónak kell döntenie, hogy külön egységeknek tekinti-e az egyes városrészeket, vagy a nagy városokat is egy településnek tekinti. Amennyiben a kerületeket külön eseteknek tekintjük, akkor előállhat az a probléma, hogy bizonyos adatsorokban nincsenek külön vezetve a kerületek adatai. Ekkor vagy nem használjuk ezeket az adatsorokat, vagy az ezer főre vetített mutatókat, illetve a százalékos arányokat mutató változókat megtartjuk és a város értékeit fűzzük minden kerület értékéhez. Ez nem egy az egyben megfeleltetés, mégis lehet olyan hozzáadott értéke az elemzésnél, ami miatt hasznosabb elfogadni ezt a kompromisszumot, mint elvetni az egész adatsort. Illetve egy harmadik választásunk is lehet, ha a nagyobb városokban külön adatbázisban elemezzük és a csak városokra vonatkozó adatsorokat a többi településnél használjuk fel. Jól látható, hogy a választás nem könnyű, és a mérlegelésnél figyelembe kell vennünk mind a módszertani elvárásokat, mind a mutatók tulajdonságait, a településeken belül megfigyelhető heterogenitást. Tanulmányunkban végül csak nyolc megye adatait elemeztük, adatbázisunkban csak 5-7 nagyobb város volt, melyeknél nem tartottuk indokoltnak a kerületek használatát.

Az adatbázisok felhasználása

Az adatbázisokat végül településnév alapján kötöttük össze. Voltak olyan adatbázisok, melyek egységei az irányítószámok voltak. A településneveknél az azonos irányítószám alá tartozó, eltérő településeknél – kisebb települések esetén figyelhető ez meg – minden településnél az adott irányítószám értékeit adtuk meg.

A végleges adatbázisban több, mint ötszáz változó szerepelt, és annyi eset volt, ahány település van Magyarországon.

Az elemzésnél két irányban indultunk el; egyfelől a különböző megyéket kellett jellemeznünk és minél részletesebb leírást kellett adnunk róluk, a lehető legváltozatosabb nézőpontokból. A másik irány egy dimenziócsökkentő eljárás alkalmazása volt, melynek végeredményeként egy mutatóba kellett sűrítanünk a települések változatosságát.

Megyék jellemzése

A megyék és kistérségek jellemzésénél azért volt hasznos az adatbázis, mert gyorsan és hatékonyan tudtuk egy adatbázisból bemutatni az egyes mutatókat, akár a mutatók kombinálásával. A bemutató anyag 25 témakörre terjedt ki, mint:

- Gazdasági helyzet
- Infrastruktúra
- Népesség és települések megyénként
- Nők helyzete
- Fiatalok (15–29 évesek)
- Idősek jellemzése
- Gyerekes családok
- Életmód, életvitel attitűdök
- Oktatás
- Egészségügyi helyzet
- Vallás és pártok
- Településszerkezet, utazás
- Lapolvasási szokások
- Megyei napilap-olvasók
- Sajtóhasználat-attitűd
- Tv lefedettség, napi hatókör
- Kereskedelem
- Vásárlási attitűdök
- Turizmus, nevezetességek
- Munkaügyek
- Online aktivitás
- Saját oldalak átlagos havi látogatószáma
- Átlagos havi affinitás-index
- Megyei napilap online

Az elemzés különlegességét az adta, hogy a megyék jellemzésében kollégáink nem csak az alap mutatókat ismerhették meg. Több, mint 350 változó bemutatása által egy sokkal részletesebb és érdekesebb képet láthattak országunkról. A változók kombinálásával jobban megérthették az ott élők életvitelét, érzéseit, fogyasztási szokásait, értékrendjét. A fogyasztási szokások és az életvitel megismerése pedig hozzájárul a lapok eladásainak megértéséhez.

Példának okáért kevesen látják egyben, hogy míg Jász-Nagykun-Szolnok megyében a települések 45%-áról mondható el, hogy mind busszal, mind vonattal meg lehet közelíteni a kistérségi központot, addig Baranya megyében ez az arány csupán 11%. Baranyában a települések 86%-a csak autóbusz-összeköttetésben van a kistérségi központtal. Továbbá megtudhattuk az elemzésből, hogy Baranya megyében csupán 1524 lakos jut egy háziorvosra, míg Bács-Kiskun megyében 1824 lakos. A gyerekorvosoknál Baranya megyében 807 gyerekre jut egy gyerekorvos, míg Bács-Kiskun megyében 926 gyerekre jut egy háziorvos, Heves megyében pedig 1168 gyerekre. További érdekesség volt, hogy a távfűtésbe bekapcsolt lakások száma Baranya megyében és Komárom-Esztergom megyében volt a legmagasabb, míg Békés megyében a legalacsonyabb, és hogy a játszóterek, tornapályák, pihenőhelyek száma Komárom-Esztergom megyében a legmagasabb. A megyei szerkesz-

tőségekben dolgozók jól ismerhetik a megyéjükben élők életvitelét, azt viszont kevésbé tudhatják, hogy ez mennyire általános vagy különleges az országban. A budapesti kollégák részletes tudása az egyes megyékről úgyszintén hiányos volt. Az elemzésnek köszönhetően egy részletes térképet kaptak, melyből mindenki a munkájához kapcsolódó mutatókat használhatta fel.

Dimenziócsökkentés

Az adatbázis létrehozásának egy másik fontos, ha nem a legfontosabb célja egy besorolás volt, mely számtalan dimenziót figyelembe véve jellemzi az egyes településeket. A jellemzésre több okból is szükségünk volt. Az egyik ok az egyes eladásokra vonatkozó becslő eljárások és más elemzések támogatása volt. A fogyasztási szokásokban bekövetkező változások megértéséhez szükségünk van arra, hogy a célcsoport környezetét is ismerjük, illetve egy-egy termék bevezetésénél többek között tudnunk kell, hogy az adott településen milyen a vásárlóerő, mennyire nyitottak kulturálisan az adott termékekre. A másik ok a kvalitatív kutatások támogatása volt. A kvalitatív kutatások alkalmával a vidéki helyszínek kiválasztásánál a megyeszékhelyeken túli városokat szoktuk választani. Több helyszín megjelölésénél nemcsak az országon belüli régiók vagy megyék szerinti eltérést szeretnénk volna figyelembe venni, hanem a kiválasztásoknál célunk volt, hogy más életvitelű, fogyasztási településeket jelöljünk ki. Feltételezésünk szerint két különböző megyében lehetnek hasonló tulajdonságokkal bíró települések. Ezeket a hasonlóságokat és különbözőségeket szeretnénk volna megragadni.

Különböző adatredukciós módszereket alkalmaztunk, melyek során – utólag már triviálisnak mondható – akadályokba ütköztünk. Az egyik akadály a kis lélekszámú települések szélsőséges értékei voltak. Ahol pár tíz fő él a településen, akár csak a nemüket tekintve is szélsőséges értékeket kaphatunk, melyek redukálják a modellek hatékonyságát. Ilyen esetekben például a települések többsége egy vagy két csoportba került és a szélsőséges csoportok alkottak külön csoportokat. A kutatásunk üzleti alkalmazásra készült, ezért az üzleti célokat figyelembe véve kellett mérlegelnünk. Mivel a nagyon alacsony lélekszámú települések terjesztési szempontból gyakran külön kezelt települések, ezeket ki tudtuk emelni az adatbázisból – és rájuk egy másik elemzésnél külön fókuszáltunk. A lélekszám-határt a terjesztési gyakorlatot figyelembe véve és a módszertani elvárásokat követve határoztuk meg.

A végleges adatredukciónál végül 218 változót használtunk fel, és a városokat 14 csoportba soroltuk be. A 14 csoporton belül lélekszám szerint megkülönböztettünk még négy alcsoportot, így az 5.000 fő alatti települések, az 5.001–10.000 fő közötti települések, a 10.001–20.000 fő közötti települések és a 20.001 fő feletti települések csoportját. A lélekszám szerinti megkülönböztetésre a szervezések alkalmával is szükségünk van, amikor a szűrésnél kiválasztott célcsoportok nagyságával kell kalkulálnunk. A bevont változók között szerepelnek az anyagi helyzetre, a vallási hovatartozásra, a kulturális értékekre, életvitelre, politikai beállítottságra, médiatermékek fogyasztására és az egészségügyi állapotokra vonatkozó változók is.

Végszóként

Tapasztalataink szerint sokforrású adatbázisokat érdemes összeépíteni. Az eredmények bemutatását követően ügyfeleink rendkívül sokat profitáltak az eredményekből, nagyon sok meglepő, hasznos megállapítást sikerült átadnunk számukra. Egy piaci igényeket kiszolgáló kutató számára a legnagyobb siker, ha egy-egy eredményeket bemutató prezentáció általános hivatkozási alap lesz a vállalati munkában. Ennek a kutatásnak a slide-jai a mai napig (2015 ősz) időnként előkerülnek a vezetők asztaláról, újabb esetekben is viszszaemlítve hozzájuk.

Az újabb beszélgetések során viszont egyre gyakrabban felmerül egy nagyon érdekes kérdés, miszerint meddig érvényesek az elemzésben tett megállapítások, besorolások. A válasz nem lehet egységes; minden egyes alkalommal meg kell nézni az eredeti kiinduló adatbázisok érvényességét (és itt nem elsősorban az a kérdés, hogy van-e frissebb, hanem az, hogy egy esetleges új adatfelvétel esetén mennyire lennének mások az eredmények). A mi esetünkben egyes elemek ugyan kis mértékben változtak, de a legfontosabb dimenziók (utazási feltételek, vásárlási szokások, médiahasználati szokások) még nem avultak el. Ha összegezni kellene, akkor éppen a Kiadónak legfontosabb téma, a média-, és ezen belül a technológiai eszközhasználat változik a leggyorsabban. Éppen ezért szinte biztos, hogy ezt a kutatást 2016-2017 körül újra el kell készíteni abban az esetben, ha a megrendelőnek erre szüksége lesz.

IRODALOM

TNS Hoffmann & Millward Brown, Nemzeti Olvasottság Kutatás (NOK), Olvasottsági jelentés 2015/2. negyedév

VALÓSÁG

A rovatunkban megszólaltatott szakértők az adminisztratív adatok felhasználásának két területét mutatják be. Kádárné Fülöp Judittal, az EMMI korábbi OECD szakértőjével az oktatás területén rendelkezésre álló adminisztratív adatállomány kezeléséről és államigazgatási felhasználási lehetőségeiről beszélgettünk. Rudas Tamás egyetemi tanárt (ELTE), a TÁRKI tudományos igazgatóját az adminisztratív adatok társadalomkutatási kezelésének gyakorlatáról kérdeztük.

„Informatikai fejlesztés és információ fejlesztés AVAGY eljutunk-e oda, hogy az adatbázisok mesélni tudnak az oktatási rendszerről?”

Interjú Kádárné Fülöp Judittal

Educatio: Miért fordult a figyelem az adminisztratív oktatási adatbázisok felé nemzetközi és hazai szinten?

Kádárné Fülöp Judit: Nyilvántartás értelemben nagyon régóta van oktatási adatbázis, hiszen a beiratkozási napló és a kartoték is az. A fejlett oktatási informatikát működtető rendszerekben az informatika fejlődésének megfelelő fázisában áttértek a kartoték-rendszerről az elektronikus adminisztrációra. Amikor az adattovábbítási lehetőségek ezt engedték, értelemszerűen létrejöttek központi elektronikus nyilvántartások, amelyekből azután statisztikai célokra használható adatbázisokat is lehetett építeni – az eredeti nyilvántartási cél fenntartása mellett. Ma az Európai Unió tagállamainak kb. a fele ilyen nyilvántartásokból állítja elő az oktatástatisztikai adatok jelentős részét.

Az elektronikus adatgyűjtő rendszer nyilvánvaló előnyei azt a látszatot keltik, hogy ezekből automatikusan elő lehet állítani statisztikai célra is használható adatforrást, adatbázist. Pedig ez tévedés. Az igazgatási célú adatok egy része szöveges adat (pl. név, lakcím), amire igazgatási okokból szükség van, ám egy statisztikai elemzésre szánt adatbázisban kényelmetlenek, mert ezekben a táblázatba rendezhető adatokat lehet jól kezelni.

Az igazgatási adatok elsődleges gyűjtői maguk az intézmények és az intézményfenntartók. (Igazgatási ok például a beiratkozások, bizonyítványok nyilvántartása, az órafelosztás elkészítése, munkaszerződések készítése, a bérelszámolás, stb.). Ezzel szemben az

oktatáspolitikának többnyire összesített statisztikai adatokra és elemzésekre van szüksége. Amiért az oktatáspolitikai is igényli az egyedi adatokat (pl. hogy a tanuló melyik iskolába jár, mikor született, hányadikos, ismételt-e osztályt, stb.) az éppen az elemzési szükségletből fakad. Ugyanis elemi adatokból sokkal többféle összesítést és elemzést lehet készíteni, mint a már eleve összesített intézményi statisztikai adatokból, éppen ezért sokkal kevesebb adatból sokkal több információt lehet nyerni.

Magyarországon a 2000-es években kezdődött igazán nagyvonalú informatikai fejlesztés az oktatásügyben. A 2000-es évek elején az iskolaadminisztrációs programok elterjesztése is megindult, de azután ez hosszú időre leállt. Viszont modernizálták az intézményi alapú statisztikai adatszolgáltatást (az ún. „októberi statisztikát”), ami korábban – egészen 2000-ig – kizárólag papíron történt. Ez megteremtette az igényt az elektronikus adatszolgáltatásra. Az iskolaadminisztrációs programfejlesztések leállása miatt a statisztikai adatszolgáltatás fejlesztése egy másik irányt kezdett követni, t.i. a web-es statisztikai adatszolgáltatás vált uralkodóvá. Ennek két oka volt. Egyrészt igény volt arra, hogy a központi statisztikai adatbázis létrehozásához ne kelljen adatrögzítéshez folyamodni (ami komoly hibalehetőség), ezért mindenképpen arra törekedtek, hogy az intézmények számítógépen keresztül szolgáltatassanak adatot. Másrészt az akkori számítógép ellátottság mellett könnyebb volt megoldani egy web-es felületen történő adatszolgáltatást, mint elérni azt, hogy minden intézmény iskolaadminisztrációs programot használjon.

2008-ra teljeskörűvé vált az intézmények web alapú statisztikai adatszolgáltatása, és ezzel *párhuzamosan* elkezdődött a központi elektronikus nyilvántartások fejlesztése is bizonyos területeken. Az utóbbiak közé tartoznak a személyi nyilvántartások (az intézmény nyilvántartás, a tanuló és a pedagógus nyilvántartás, a középfokú beiskolázást szolgáló adatbázis, az érettségi adatbázis, a felvételi információs rendszer és a felsőoktatási információs rendszer is). Ezek a rendszerek – néhány kivétellel – a mai napig egymással párhuzamosan és egymásról kevésbé tudva működnek.

A megnövekedett figyelem az elektronikus adatbázisok iránt annak is köszönhető, hogy 1990 után alaposan átalakult az oktatási rendszer, ami a szabad iskolaválasztás és a felsőoktatási képzési programok sokasodása révén jóval komplikáltabb lett. 1990-től kezdve az intézmények a fogyó gyermeklétszám miatt kénytelenek voltak különböző speciális, egyedi ajánlattal, tantervvel előállni annak érdekében, hogy minél több diákot vonzzanak. Ugyanez valósult meg a felsőoktatásban is, ahol viszont a csökkenő demográfiai tendencia a bejutási küszöböt csökkentette, ezzel egyidejűleg viszont olyan programok indítására ösztönözte a felsőoktatási intézményeket, amelyek ezeket az újabb hallgatói rétegeket is vonzotta. Egy ilyen sokszínű rendszert persze statisztikailag sokkal nehezebb követni, mint egy olyan rendszert, amely egységes központi tanterveken, szakokon és követelményeken alapszik. Eltekintve attól, hogy sokkal aprólékosabb bontásokra van szükség, ha a változásokat követni akarjuk, egyéb problémák is adódnak, ami a statisztikusoknak fejtörést okoz. Például többek között az, hogy ugyanaz a diák (pl. szakképzésben, felsőoktatásban) párhuzamosan több programban is részt vehet, akár több intézményben is. Így aztán sokkal kifinomultabb, személyes adatokat tartalmazó ún. mikroadatbázisokra van szükség az olyan egyszerű kérdések megválaszolásához is, hogy „hány szereplő van a rendszerben”.

E: Mindez milyen fejlesztéseket és új irányokat hozott előtérbe?

KFJ: A 2000-es évek elején megérett az idő az oktatási azonosító bevezetésére. Ez egy olyan egyedi azonosító szám, amit minden gyermek (most már az óvodába lépéskor) megkap, és ez az azonosító elkíséri egész tanulmányi pályafutása során. Bevezetése persze nem ment zökkenőmentesen. Eleinte előfordult olyan eset, hogy egy tanuló, ha két intézményben tartották számon (pl. egy iskolában és egy szakképző intézményben) két azonosítót is kapott. Problémát jelentett továbbá, hogy majdnem egyszerre kezdtek el bevezetni az oktatási azonosítót a közoktatásban és a felsőoktatásban, így jó ideig átfedések is voltak. Körülbelül 2010/2011-re jutott el a rendszer odáig, hogy az oktatási azonosító nagyjából be tudja tölteni a funkcióját, t.i. azt, hogy a tanuló egész iskolai pályafutása követhető legyen. Ehhez az is kellett, hogy az intézményeket rá lehessen bírni, hogy 5 napon belül jelentsék diákjaik belépését, illetve kilépését nemcsak az intézmény, hanem a program vonatkozásában is.

Sajnos az iskolák informatikai ellátottsága mindmáig nagyon különböző. Van olyan iskola, ahol nincs szélessávú internet vagy lassú és elavult számítógépek vannak. Súlyosabb probléma azonban, hogy az intézmények számára a megnövekedett adatszolgáltatási igény egyre nagyobb terhet jelent, ráadásul minden haszon nélkül, hiszen az összegyűjtött adatokból a központ nem szolgáltat vissza olyan összesítéseket és elemzéseket, amiknek az iskolák hasznát vennék. Ebben egyetlen kivétel az Országos Kompetencia Mérés, ami többek között azzal a céllal készült, hogy az iskolák össze tudják hasonlítani saját diákjaik teljesítményét más hasonló iskolákéival.

E: Milyen a főbb oktatásstatisztikai adatforrások mai állapota?

KFJ: Az „októberi statisztika” máig a legmegbízhatóbb adatforrás az oktatási rendszer állapotáról. Ez részben annak köszönhető, hogy ezeket az adatokat rendszeresen ellenőrzik, mégpedig több szervezet is, mert ebből szolgáltatunk adatokat a nemzetközi szervezeteknek (Eurostat, OECD). A köznevelési intézmények nyilvántartása ma már közhiteles nyilvántartásnak tekinthető, ami nagy előrelépés az 5-10 évvel ezelőtti állapothoz képest, és fontos a statisztikai elemzés szempontjából. A felsőoktatási információs rendszer mint igazgatási adatforrás kitűnő, statisztikai elemzésre kevésbé könnyen kezelhető. Meglehetősen sok fejlesztést igényel még annak megoldása, hogy a személyi nyilvántartásokból lehessen tanulói és pedagógus adatokat szolgáltatni, de nem kétséges, hogy ez a fejlesztés egyedüli útja. Ehhez elsősorban a személyi nyilvántartások statisztikai adatfeldolgozásához szükséges eszközrendszer kiépítésére van szükség. Általában igaz az, hogy az igazgatási nyilvántartások tervezői nemigen vették figyelembe azt az elvárást, amit a Központi Statisztikai Hivatal régóta hangoztat: t.i. hogy mind a pontosság és megbízhatóság érdekében, mind pedig költséghatékonyság okán arra kell törekedni, hogy a nyilvántartásokat tegyék statisztikai elemzésre alkalmassá, ahelyett, hogy újabb és újabb adatgyűjtéseket szorgalmaznának.

E: Miért volt szükség pedagógus-nyilvántartásra is?

KFJ: A finanszírozási rendszer változása és az új pedagógus előmeneteli rendszer bevezetése motiválta a pedagógus nyilvántartás létrehozását. 2013-ig az iskolák önkormányzati fenntartásban működtek, amihez az állam normatív hozzájárulást – fejkvótát – adott a tanulólétszám alapján. Ebben az ún. normatív finanszírozási rendszerben az önkormányzatok maguk döntötték el, hogy hány pedagógust vesznek fel, természetesen a törvényi

keretek között. Amikor a 2011-es köznevelési törvény alapján az állam részben átvette az iskolák közvetlen finanszírozását, a normatív támogatásról áttértek a pedagógusbérek és a közvetlen oktatási költségek finanszírozására. Emiatt a pedagógusbérek költségvetési tervezéséhez a személyes adatok központi nyilvántartása is szükségessé vált (pl. végzettség, életkor, illetve szakmai gyakorlat). A pedagógus életpálya modell bevezetése is megerősítette ezt az igényt, ami korábban fel sem merült.

E: Kialakultak-e az adatok előállításának és kezelésének egységes keretei?

KFJ: A 2011-es köznevelési törvény végrehajtási rendelete (229/2012. kormányrendelet) meghatározza, hogy az intézményeknek milyen adatokat kell nyilvántartaniuk, és az Oktatási Hivatalnak milyen nyilvántartásokat kell vezetniük. A technikai háttér a győri Educatio szervezetében adott volt, ezért az Oktatási Hivatal összes adatgyűjtését erre az elektronikus útra terelték. Az ún. Köznevelési Információs Rendszer keretében működő adatgyűjtések egyúttal kommunikációs felületek is, amelyeken keresztül az iskola bejelenti az újonnan érkező diákot, megadja a felvételi jelentkezéshez szükséges adatait, illetve elküldi az érettségi jelentkezéseket, ami szerint az OH megszervezi a vizsgákat. Az Országos Kompetenciamérések szervezésének nagy része is ezen a kommunikációs felületen történik.

E: Hogyan lehet ezeket az adatbázisokat a jogosultságok kezelésén és a szervezésén túl elemzési célra is használni?

KFJ: A közoktatási információs rendszerben rengeteg adat van. Ezek egy része igazgatási vagy szervezési célra készült és ennek általában meg is felel. Ugyanezeket az adatbázisokat azonban többnyire komoly fejlesztés után lehetne rendszeresen elemzési célokra használni. Például a tanulói nyilvántartásban van kb. tíz millió rekord: adatok, amiket különböző időpontokban és intézményekben vettek fel ugyanarról a tanulóról. Az oktatási azonosító alapján a tanuló persze azonosítható, és az is megtudható, hogy az adott időpontban miért jelentették be. De ez más kérdés, mint az, hogy egy adott időpontban hány tanuló van a rendszerben. Vagy, hogy az összes közül mennyi a lány, hányan vannak a nyolcadik osztályosok, hányan ismételnék osztályt, stb. Ahhoz, hogy az efféle lekérdéseket rendszeressé lehessen tenni, az adatbázist kifejezetten erre a célra kivonatolni kellene, és évente egyszer egy adott ún. referencia időpontban „kimerevítve”, tisztítva és megfelelő dokumentációval ellátva szerkesztett állapotban kellene a statisztikusok rendelkezésére bocsátani. És ez nagyjából így van a többi nyilvántartással is. Ezzel a fejlesztéssel még adós az Oktatási Hivatal.

Jó jel, hogy az elemzés igényeit felismerve az Oktatási Hivatal megvizsgáltatta a KIR adatbázisokat a statisztikai elemzés lehetősége szempontjából, és javaslatokat kért egyes adatbázisok továbbfejlesztésére.

Az oktatási adatbázisokból származó adatok elemzése statisztikai és oktatáspolitikai szakértelmet igénylő munka. Jól érzékelteti ezt, hogy arra a gyakran feltett egyszerű kérdésre, hogy „Hány pedagógus van?”, az elemző általában zavarba ejtő viszontkérdéssel válaszol, t.i. hogy „Milyen célra kell az adat?”. Pedig nem bosszantani akarja a kérdés feltevőjét, hanem csak azt szeretné tudni, hogy hogyan számoljon. Csak a főfoglalkozású pedagógusokat kérdezik? Vagy a főfoglalkozású pedagógusok statisztikai létszámát (amiben a részfoglalkozásúak az óraszámoknak megfelelő súlyukkal szerepelnek)? Esetleg az összes pedagógus számát, aki megfordul az osztályokban? Vagy éppen azok létszámát,

akiket pedagógus státuszban foglalkoztatnak, noha beosztásuk szerint nem végeznek tanítási munkát? Mindegyik esetben más lesz a számítás eredménye. Más is kell kérdezni, ha az egy pedagógusra eső gyermeklétszámot kérdezzük, illetve ha a pedagógusbérek-re szánt összeget kell tervezni.

E: Ki fér hozzá az adatbázisokhoz és mire használják az adatokat?

KFJ: Sajnos jelenleg ez az egyik legégetőbb probléma. Az Oktatási Hivatal egyre kevesebb oktatásstatisztikai adatot tesz közzé, aminek többféle oka is lehet. Az egyik ok nyilván az a félelem, hogy az adatok minősége nem elég jó. Ami részben igaz is, csak éppen a megoldás nem az információ visszatartása, hanem ellenkezőleg: a nagyobb nyilvánosság és az adatok rendszeres használata, publikálása, ami kikényszeríti az adatok minőségének javítását. Ugyanis az adatszolgáltatók aszerint veszik komolyan az adatszolgáltatást, hogy mennyire van tétje a „lazaságnak” az adatszolgáltatásban. Két példát mondok erre. Az iskolai könyvtárak a statisztika minden évben elkéri. Tudjuk, hogy nem egy iskola ugyanazt az adatsort adja minden évben. De ebből még sohasem lett baj és haszon sem származott belőle. Semmit sem publikáltak az adatokból, és sohasem kaptak az iskolai könyvtárak fejlesztési forrásokat könyvekre, attól, hogy nem gyarapodott a könyvvállományuk egyik évről a másikra. Az ellenkezőjére is van példa. Abban az időben, amikor az önkormányzatok normatív támogatást kaptak a tanulók után, szeptemberben az önkormányzatoknak le kellett adniuk a tanulólétszámot, mert ennek alapján kapták a támogatást. Ha ez a létszám több mint 5%-kal eltért az októberi statisztikától, a többlet támogatást vissza kellett fizetniük. Nos, nem mondanám, hogy nem igyekeztek pontosak lenni.

E: Van-e szervezett, egységes stratégia az oktatási adatvagyon felhasználására, menedzselésére, fejlesztésére, esetleg adat-összekapcsolások végzésére?

KFJ: Szerintem nincs. Erre abból következtetek, hogy az elemző kapacitás fejlesztése rendre elmarad, sőt inkább leépül az is, ami van. A háttérintézmények összevonása és a létszámcsokkentések azt a benyomást keltik, hogy éppen az elemzés az, amiből az oktatáspolitikai nem kér. Pedig a rendszer statisztikai elemzése égetően fontos. Az oktatáspolitikai intézkedések hatását és mellékhatását csak ezen az úton lehet vizsgálni. Ha egy diák nem tanul meg olvasni nyolc év alatt, az elég nagy baj a diáknak és esetleg a pedagógus is kudarcként éli meg. De ha az analfabetizmus az országban minden tízedik gyereket érint, az már nagyon nagy baj. Arra való a statisztika, hogy megmutassa a probléma nagyságrendjét, hiszen ettől függ, hogy mikor, milyen módon kell beavatkozni, hogy a kedvezőtlen tendenciákat megállítsuk. Ugyanakkor az intézkedések hatását is figyelni kell. És ez is csak statisztikai elemzés segítségével lehetséges – még pedig hosszabb időszakon át összehasonlítható adatokkal.

E: Összességében milyennek látja az adatalapú oktatáspolitikai döntéshozás hazai helyzetét?

KFJ: Ma egy kicsit úgy néz ki, mintha a politika árnyékában az oktatásstatisztika és a statisztikai elemzés csak pár ember hobbjá lenne. Emellett van egy erős informatikai lobby, melynek elsődleges érdeke az informatikai termékek és fejlesztések eladása. Ezért a fejlesztési források rendre újabb és újabb „fejlesztések” felé terelődnek, miközben kidobjuk azt is, ami jó volt, de legalábbis nincs pénz annak a fenntartására. De azért remény-

kedem, mert látok az oktatáspolitikai területén néhány olyan embert (nem sokat), aki szeretne és tud is tenni azért, hogy az oktatási információs rendszer jelentősen javuljon. Emellett hiszek abban, hogy az oktatáspolitikai igazságtámasza a nyilvánosság. Az oktatás végül is közügy, nemcsak a politikusok és az újságírók – vagy éppen a pedagógusok ügye. És a hiteles oktatási adatok olyan semleges tények, amelyeket – harag és részrehajlás nélkül (sine ira et studio) – megvitathatunk. No nem egymás legyőzése, hanem a legjobb közös megoldás keresése érdekében. Valahogy úgy, ahogy a finnek teszik.

Az interjút Hideg Gabriella készítette

„Az adatok nem önmagukért beszélnek, mi nézünk valahogy rájuk” Interjú Rudas Tamással

Educatio: Elsőként a TÁRKI-ban zajlott adminisztratív adat alapú kutatások áttekintésére kérném.

Rudas Tamás: Tudtommal a TÁRKI-ban az első ilyen jellegű nagy kutatás a '90-es évek közepén volt, és azt hiszem mind szakpolitikai döntés-előkészítő jellegét, mind tudományos tartalmát tekintve messze megelőzte a korát. E kutatás révén a kormányzat az adórendszer különféle elemein bevezetett változtatások hatásait kívánta modellezni. Ebbe a munkába részben adminisztratív adatállományokat is bevontunk – ezek legfőképpen maguk az adóbevallások voltak. Emellett több survey adatállománnyal dolgoztunk: részben a KSH háztartási költségvetés-felvételét, részben TÁRKI-s monitorfelvételek állományait kapcsoltuk össze többszörös imputációs eljárással. E köré a virtuális adatállomány köré építettünk egy valamelyest interaktív lekérdező felületet. Mikroszimulációnak neveztük ezt az elemzési módszert, melynek lényege, hogy az adott, meglévő adatokat tekintjük a valóság leírásának, és azt vizsgáljuk, hogy különböző policy-változatok milyen hatással vannak ezek alakulására. Ezt a modellt a kormányzat frissített adattartalommal még hosszú ideig – talán a 2000-es évek közepéig – használta. Maga az adatösszekapcsolási gyakorlat ekkor a világon már bevett eljárás volt ugyan, ám ez elsősorban record linkage-t jelentett. Ennek során az egyazon személyekről rendelkezésre álló rekordokat, feljegyzéseket azonosítjuk és kapcsoljuk össze. A mi eljárásunk virtuálisabb volt, hiszen az egyes adatbázisokban beazonosíthatóan nem ugyanazok a személyek szerepeltek, hasonlóság alapján kapcsoltunk tehát. Lett is valami sikere ennek a módszernek, de amennyire én tudom, a magyar közigazgatási tevékenységből kikopott. Fennállásáig, a Költségvetési Tanács volt az utolsó megrendelője a munka frissítésének.

E: Hogyan fejlődött ezek után a TÁRKI adatbankja?

R.T.: Az előbbi volt az első olyan munka a TÁRKI-ban, amely már ekkor tartalmazta ennek a most zajló nagy adatrobbanásnak minden elemét: nagy adatállományon zajlott, adminisztratív rekordokat használt, összekapcsoláson alapult, és szakpolitikai döntéselőkészítést szolgált. Azóta – és főleg jelenleg – a TÁRKI jóval kisebb adatgazda, mint azt kívülről gondolni lehetne. A TÁRKI adatbankja viszonylag hosszú ideig az összes, vagy majdnem az összes lényeges társadalomtudományi kutatás adatait megkapta különböző konstrukcióban. Ezek bárki számára ingyenesen elérhetőek és továbbra is rendelkezésre állnak, de az újabb adatállományokat már a TÁRKI saját adatfelvételei je-

lentik. Az adatszolgáltatási megállapodások megszűntek, a partnerek túlnyomó többsége nem kívánta újrakötni ezeket. Emögött két fő tendenciát látok magyarázatként. Az egyik az, hogy technikailag sokkal egyszerűbbé vált az adatok tárolása és közzététele. A másik pedig, hogy az adatok politikai és gazdasági értéke vált sokkal világosabbá. Azt gondolom, e két tényező vezetett oda, hogy pillanatnyilag Magyarországon – tudomásom szerint – nem működik olyan központi társadalomtudományi adatbank, mint amilyen hosszú ideig a TÁRKI adatbank volt.

E: Milyen trendeket lát jelenleg az adatkapcsolásban?

R. T.: Az adminisztratív adatállományok összekapcsolására vonatkozó legnagyobb kísérlet pillanatnyilag az egész világon a register based census. Ez azt jelenti, hogy a népszámlálást már nem azzal a hagyományos módszerrel végzik, amely igyekezett kapcsolatba lépni az országot minden egyes lakójával, akár személyesen, akár elektronikusan. Ehelyett a már létező regiszterek alapján próbálják az ország lakóit összeszámolni és bizonyos tulajdonságaikat megállapítani. Magyarországon ettől végtelenül messze vagyunk, de Ausztriában az utolsó népszámlálás már regiszter alapon zajlott. Ezek az adminisztratív adatállományok a táppénztől a közúti szabálysértésig tartalmazznak egymással összekapcsolható egyéni adatokat. Erre a lehetőségre is utaltam akkor, amikor a nagy adatállományok kezeléséhez, tárolásához és közzétételéhez kapcsolódó politikai és gazdasági érdekek láthatóvá válásáról beszéltem. Például egy regiszter alapú népszámlálás nagyon sok pénz megtakarítását tenné lehetővé egy ország számára. A módszernek azonban nem ez a legnagyobb előnye a mostani hagyományos népszámlálási adatokkal szemben. A lényeges elem, hogy kivész bizonyos politikai döntési szabadságot a népszámlálásból. Ha ugyanis a népszámlálást regiszter alapon végezzük, akkor a regiszterek körött adattartalma használható fel, míg ha megkereséses alapon, akkor a gyűjtött információk köre politikai döntéstől is függ. Ebben az értelemben a big data nem csak nagy lehetőség, hanem bizonyos értelemben véve korlát is az államigazgatás működésében. Persze ez az eljárás olyan regiszterek működését kívánná meg, melyekben érvényes és megbízható, friss adatok vannak.

E: Az adminisztratív adatok eltérő kezelése eltérő felhasználásukat is jelenti?

R.T.: Az adminisztratív adatok összekapcsolása a központi döntések előkészítése érdekében nem kizárólag technikai feladat. Az összekapcsolt adatállományokból kinyerhető információk nyilvánvalóan valamilyen politikai interpretáción mennek keresztül, mielőtt egy politikai döntés alapjait képeznék. Példa erre a felsőoktatásban az alapszakok számának adminisztratív adatösszekapcsolás eredményeire hivatkozó, közelmúltban zajlott módosítása. Tulajdonképpen ha egy más ország, más közigazgatási rendszerből hallok azt az érvelést, hogy az alapszakok listáján tervezett változtatások, a fenntartásukra avagy megszüntetésükre vonatkozó döntés adminisztratív adatok alapján történt, akkor mindez nagyon megnyugtatónak hallatszott volna. Viszont ebben az esetben úgy vélem, hogy nagyjából tudom, milyen adatok állnak rendelkezésre, és azok hogyan keletkeztek, és arról is van legalábbis véleményem, hogy ezeknek mennyire van köztük az egyes szakok értelmes vagy értelmetlen voltához. Mindezen ismereteim alapján valahogy nehéz elképzelnem, hogy az adatokból erős politikai interpretáció nélkül is következne a meghozott döntés. Egy olyan közegben, amikor nemegyszer úgy vélhetjük, hogy az adat inkább csak

hivatkozási alapként szolgál, nagyon fontos erősíteni a hivatalos adminisztratív állományok megfelelő minőségű összekapcsolását és felhasználását.

E: A szabályozási keretek ehhez adottak?

R.T.: Magyarország a szabályozást tekintve egy viszonylag extrém helyzetben van. Ez alatt – nem minősítő, hanem leíró jelleggel – az adatvédelmi törvények igen erős voltát értem. Az erős szabályozó környezethez társul (személyes tapasztalataim alapján) a lakosság viszonylag alacsony szintű érzékenysége vagy érdeklődése az információ felhasználása iránt. Mondhatni, hogy bizonyos értelemben egy, a mostaninál liberálisabb politikai rendszer a lakosságot próbálta védeni a szigorú adatvédelmi törvények a bevezetésével. Ezeket a kutatók jelentős része támadja, hiszen lehetetlenné teszik a különböző adatállományok összekapcsolását. Nagyon nehéz az én pozícióból megítélni, hogy a szigorú adatvédelmi törvények inkább hasznosak vagy inkább korlátoznak egyébként hasznos tevékenységeket. Gyengének látom azt az intézményrendszert, amely ennek a betartásán örködi. A magyar viszonyok kapcsán azt kell mondanom, hogy az egy dolog, amit a törvény leír és egy másik az, hogy ezt milyen mértékben és milyen pontossággal, precizitással tartják be. Zárójelben jegyezném meg ezzel kapcsolatban, hogy éppen TÁRKI-s tevékenységek kapcsán találkoztam olyan helyzettel, amelyben a megrendelő adminisztratív adatállományok összekapcsolását kívánta volna tőlünk, amelyhez mi a nemzetközi szakirodalomban ismert algoritmusokat kívántuk alkalmazni. A meglévő szabályozási keretek azonban a megvalósítást jogi kérdéssé minősítették és a jogász kolléga, akinek a törvénnyel való egybevághóságot kellett megítélnie, nem tudott egyértelműen állást foglalni a kérdésben, az adott keretek között. A projekt így nem tudott megvalósulni. Azóta az adatvédelmi törvényeket lazították, és létezik egyfajta speciális mód az adatösszekapcsolásra, de a magyar törvények szigorúak e tekintetben.

E: Eszerint a jogi környezet nem kedvez a technikailag egyébként megvalósítható felhasználásnak?

R.T.: Mikor egy ilyen eljárás jogi kérdéssé transzformálódik, onnantól nem lesz sikeres. Korábban is volt egy államigazgatásból érkezett gazdasági vonatkozású adatkérésünk, amely során a minisztériumi támogatás ellenére döntöttek úgy az adatgazda jogászai, hogy nem látják biztosítva a megfelelő szintű adatvédelmet. Érthető részükről, hiszen az eljáráshoz nem jogi védelem, hanem nemzetközi szokásos algoritmikus védelem állt rendelkezésre. Hasonlóan például a bankautomatába beütött PIN kódhoz, az algoritmikus védelem technikailag akadályozza meg (vagy teszi hatékonytalanná) a kódok visszafejtését. Mindez bizonyos matematikai és statisztikai eljárásokkal biztosítható, mint ahogy az egész világon komoly pénzügyi tranzakciók zajlanak ezen algoritmikus védelmek ernyője alatt. De ez nem jogi garancia.

E: Az adatok felhasználása a technikai és jogi keretek mellett az államigazgatás részéről bizalmat is feltételez. Megvan ennek a bázisa?

R.T.: Az adatfelhasználást valóban nagyon nagy mértékben beárnyékolja az a bizalomhiány, ami az egész ország, az egész társadalom működésének sok területén látszik. Az alap az, hogy sokan bizonyos naivitással adatokat szolgáltatnak. Azok a hozzáértő szervek és emberek pedig, akiknek az ezeket védő törvény betartása felett kellene örködniük, inkább bizalmatlansággal viseltetnek a lehetséges felhasználók iránt, legyen az ku-

tató vagy gazdasági vállalkozás. Tudomásom szerint nagyon kis mértékben valósul meg az, hogy gazdasági vállalkozások által gyűjtött óriási és szerintem fantasztikus adattartalmú állományokhoz pusztán kutatási célra hozzá lehessen férni, leszámítva persze az adott vállalkozás érdekében történő hozzáférést, hiszen ez az adatgyűjtésnek célja. Tényleg azt hiszem, hogy a bizalmatlanság írja le leginkább a szereplők hozzáállását. Ezt a bizalmatlanságot azonban nem érzem indokolatlannak. Nyilván adott egy államigazgatás, amely nem feltétlenül készült fel szakmailag az adatrobbanás megértésére. Az adatokat termelő procedúrák elmúlt évtizedekben bekövetkezett gyors változására és ennek társadalmi hatásaira senki sem volt – én sem – felkészülve. Érthető tehát az államigazgatás bizonyos fajta óvatossága.

E: Ugyanezt a tendenciát érzékeli a lakosság körében?

R.T.: Maga a bizalmatlanság társadalmi szinten is érzékelhető. A big data valamilyen szinten az információs monopólium csökkentésére is képes, ami még a hozzáértő felhasználókban is nagyon erős és jogos bizalmatlanságot szülhet. Egy pár perccel ezelőtti jó példát is mondhatok erre. Két számítógépet lát az asztalon. Az egyik gépen a booking.com-on néztem egy szálloda adatait. Nem volt nyitva semmilyen ablak, amiben e-mail lett volna és nem jelentkeztem be a booking.com rendszerébe. Ennek ellenére, amint ott bezártam az ablakot, egy perc múlva érkezett egy e-mail a másik gépemen nyitva levő, egy másik e-mailemhez csatolt ablakban, amelyben az állt, hogy a booking.com látta, hogy abbahagytam a keresést, és ha gondolom, ide kattintva tudom folytatni. Sejttem persze, hogy valahol az apró betűs részben jóváhagytam az eljárást. A mobiltelefonra letöltött applikációk esetében a helyzet hasonló, amennyiben a működésükhöz engedélyezni kell, hogy bizonyos információkhoz hozzáférjenek. Egy ilyen közegben tényleg azt hiszem, hogy jó lenne, ha a lakosság tudatosabb lenne a saját adatainak megadásával kapcsolatban. Amíg tehát a politikai döntéshozók részéről az információs monopólium elvesztésétől való féltelmet és az ebből fakadó bizalmatlanságát látom, addig a lakosságnál néha meglepő az óvatosság hiánya. Számos olyan törzsvásárlói program létezik például, amely 50 levásárolható pontért elkéri a telefonszámot, e-mail címet, születési időt és egyéb személyes adatokat. Semmi ördögít nem látok ezekben a tevékenységekben, amennyiben deklarált, világos módon és feltételekkel zajlanak. De az azért jó volna, ha mindenki tudatosabban döntene az adatközlésről.

E: Hogyan reagál minderre a társadalomkutatás?

R.T.: Ahogy előbb is említettem, a társadalom egészen máshogy működik ma, mint 10 vagy 20 éve. Ilyen mérvű fejlődést még a társadalomkutatók sem láttak előre. Az elképesztő mennyiségű rendelkezésre álló adat érdemi felhasználása és elemzése azonban – részben az adatgazdák érdekei miatt – csak bizonyos területekre korlátozódva történik meg. Ha például egy szupermarket gyűjti a vásárlások adatait, akkor azt kizárólag saját fejlesztési céljaira használja fel és csak nagyon ritkán válik hozzáférhetővé más típusú elemzések számára. Ez nem egy örömteli dolog. Egy másik, amit tendenciaként látni vélek a világban, az az új eljárások – mint az adatbányászat vagy a machine learning – speciális humán erőforrás igénye. Ezeket a tevékenységeket elsősorban nem társadalomtudósok, hanem mérnökök és fizikusok végzik. Nyilván ezek az emberek nagyon ügyesek az algoritmusok gyártásában, szoftverek írásában, tehát az adatbázis tevékenység, gépi programozás nagyon magas színvonalon tud megvalósulni. Ami viszont ezeknek a szak-

embereknek az inkább természettudományi, mint társadalomtudományi háttéréből adódóan nehezebben valósul meg, az az eredményekhez tartozó helyes interpretáció. Érdekes kérdés persze, hogy mi is a helyes interpretáció. Azt gondolom, ennek legfőbb jegye mindenképpen a fenntartás. Vagyis nem gondolni azt, hogy z mögött a társadalmi vagy gazdasági jelenség mögött, amelyet valamiféle big data forráson éppen vizsgálunk, egyetlen összefüggés, egyetlen igazság van. Vagy akár azt, hogy az általunk azonosított összefüggés bizonyosan helyes. Minden adatállománynak lehetnek alternatív interpretációi, elemzései, még akkor is, ha sok, a survey-t jellemző bizonytalanság az adminisztratív adatállományokat nem jellemzi. Úgy látom, a szokásos értelmezésből ma hiányzik az a nézet, hogy az adatok nem önmagukért beszélnek, hanem mi nézünk valahogy rájuk. Ez a szemléletmód az adminisztratív adatok kutatási felhasználásában változatlanul releváns. Ennek kezelésében maga a világ van lemaradva abból adódóan, hogy a hatalmas és egyre növekvő hozzáférhető adminisztratív adatállományokat jóval kevésbé társadalomtudósok, sokkal inkább mérnökök és fizikusok elemzik. Ezt a szomorú tendenciát látom legerősebbnek.

E: Mindez a társadalomkutatók iránti elvárásokra is kihat?

R.T.: Szerintem nyilvánvaló, hogy itt is fejlődésre, új kompetenciákra van szükség. Már 30-35 évvel ezelőtt elindult egy változás, melynek során a társadalomkutatók akkori új generációja elkezdett a survey követelményeihez igazodó skilleket megszerezni. Ez a tudáskészlet most, az adatbányászati lehetőségek és feladatok következtében frissítésre, bővítésre szorul. A társadalomtudományi felhasználásban a szükséges változást nehezíti, hogy a technikai belépési küszöb nagyon magas, hiszen mind az alkalmazott algoritmusok, mind az ezeket megvalósító IT eszközök rendkívül szofisztikáltak.

Az interjút Galántai Júlia készítette

KUTATÁS KÖZBEN

KIRÁLY GÁBOR, GÉRING ZSUZSANNA, CSILLAG SÁRA, GÁSPÁR TAMÁS,
KÖVES ALEXANDRA, SEBESTYÉN GABRIELLA

Iskola a jövőben? Egy jövőkutatási kísérlet bemutatása

Napjainkban a felsőoktatásnak számos kihívással kell szembenéznie. Az új technológiák és társadalmi gyakorlatok elterjedése, valamint a gazdasági-társadalmi környezet átalakulása megkérdőjelezi az adottnak vett válaszainkat az oktatással és tanulással kapcsolatban.

A *társadalmi tényezők* között említhetjük az oktatás társadalmi és gazdasági szerepének átalakulását, a gazdasági és társadalmi változások gyors és sokszor „hirtelen” jellegét és az ezzel kapcsolatos bizonytalanságot, amely a specializált szakmai tudással szemben az általános képességek felértékelődését vonja maga után. A *technológiai tényezők* közé sorolják a hallgatók megváltozott információfeldolgozási és koncentrációs képességét, az infokommunikációs technológiák „állandó” jelenlétét az oktatási terekben, valamint új, virtuális alapú oktatási formák megjelenését.

Ezeknek a technológiai és társadalmi kihívásoknak az együttes jelenléte megkérdőjelezi a felsőoktatási gyakorlat megszokott formáit. Az ezzel kapcsolatos szerepváltás érzete egy igen változatos és szerteágazó diskurzust indított el a felsőoktatási gyakorlat új formáival kapcsolatban. Az oktatás jövőjével több kutatás és jelentés is foglalkozott a közelmúltban (OECD, 2008, *Redecker et al.*, 2011), de ezek nem reflektáltak külön a kelet-közép európai térség sajátosságaira, nem szűkítették a témát egy-egy részterületre (jelen kutatás csupán a gazdasági képzésre kíván fókuszálni), valamint nem az általunk használt backcasting módszertani technikára épültek.

Az alábbiakban egy – a gazdasági képzés jövőjével foglalkozó – kutatást szeretnénk röviden bemutatni az olvasóknak. Az írás szerkezetét egyfajta szűkülő tematikus struktúra jellemzi. Az első részében felvázoljuk a kutatás felépítését és a különböző pillérekhez kötődő célokat. Az írás második része már kifejezetten csak a backcasting műhelyhez kötődő jövőkutatási módszertant tárgyalja. A harmadik részben pedig bemutatjuk és elemezzük kutatásunk egyik legizgalmasabb eredményét: a bevont szakemberek által kidolgozott, a gazdasági felsőoktatással kapcsolatos jövőképet. Az írás befejező részében visszatérünk a kutatás eredményeinek fő tanulságaira és a továbblépési lehetőségekre.

► *Educatio* 2015/3. Király Gábor, Géring Zsuzsanna, Csillag Sára, Gáspár Tamás, Köves Alexandra, Sebestyén Gabriella: *Iskola a jövőben? Egy jövőkutatási kísérlet bemutatása*, 108–115. pp.

A kutatás felépítése

Kutatásunk során egy komplex stratégiát alkalmaztunk, hiszen maga a téma is szerteágazó és több elméleti „valóságszinten” (mikro/egyéni; mezo/szervezeti; makro/nemzeti; meta/globális) értelmezhető. Azért, hogy a jövővel kapcsolatos gondolkodást és terveket minél több szinten és színtérben meg tudjuk ragadni, egy három pilléren nyugvó kutatási eljárást dolgoztunk ki. A pillérek a következőképpen keretezik a vizsgálatot:

1. pillér: célja, hogy feltérképezze a téma nemzetközi és hazai szakirodalmát és a felsőoktatással kapcsolatos víziókat;
2. pillér: célja, hogy megvizsgálja, hogy a Budapesti Gazdasági Főiskola és más gazdasági képzőhelyek oktatói és hallgatói milyen egyéni véleményekkel rendelkeznek ezekről a kérdésekről;
3. pillér: célja, hogy feltárja, hogy a gazdasági képzésben aktívan részt vevő oktatók és hallgatók hogyan gondolkodnak a gazdasági képzés jövőjéről, milyen jövőképet tartanak ezzel kapcsolatban ideálisnak (‘normatív vízió’) és milyen utakat és lépéseket javasolnak, hogy el tudjunk mozdulni e ‘normatív vízió’ irányába.

A három pillér közül a legnagyobb súllyal a harmadik bír, de a másik két pillér is hozzájárul a kutatáshoz önálló eredményekkel. Ez a rövid írás a 2014. október – 2015. március közötti első kutatási fázis részeredményeit mutatja be, ez idő alatt az első pilléren belül a nemzetközi szakirodalom és víziók feltérképezésének alprojektje kezdődött el; a második pilléren belül BGF-es oktatókkal készültek interjúk (ennek bemutatásától itt most eltekintünk); a harmadik pilléren belül pedig az oktatói backcasting workshop valósult meg. A mindhárom terület esetében 2015-2016 tanév folyamán folytatjuk a vizsgálatokat. Az alábbiakban a 3. pillér módszereit és a backcasting műhelyen kidolgozott ‘oktatói’ víziót mutatjuk be.

A backcasting alprojekt módszerei

A kutatás legfontosabb, és legújyszerűbb pillérét a *backcasting alprojekt* jelenti. A backcasting (magyarul talán visszafejtésnek lehet fordítani) egy olyan jövőkutatási módszer, amely – az eseményekre gyakorolt emberi befolyás jelentőségét és lehetőségét feltételezve – egy ideális jövőből kiindulva vezeti vissza a lehetséges lépések sorát a jelenig. Ez, az előrejelzéshez képest fordított logika lehetségessé teszi azt, hogy a jelen gondolkodásának kereteit elhagyva tárhassunk fel lehetséges cselekvési irányzatokat (Király et al., 2013). A gazdasági felsőoktatás jövőjét feltáró backcasting folyamat a normatív vízió idejét 2050-ben határozta meg.

Érdeemes kiemelni, hogy a kutatás egyben egy módszertani fejlesztési kísérlet is volt. Mivel a backcasting folyamata inkább tekinthető egy módszertani szemléletnek, mint egy pontosan meghatározott és szekvenciálisan leírható metodológiának, ezért nem egyértelmű, hogy milyen módon is érdemes felépíteni egy ilyen folyamatot. Általában három fő elem mindig részét képezi a backcasting műhelyeknek: a keretezés (milyen témákat tárgyaljanak a résztvevők), a vízióalkotás (hogyan nézzen ki a jövő) és a visszafejtés (milyen lépésekkel lehet a jövőképet a jelennek összekötni).

A kutatócsoport módszertani fejlesztési gyakorlata abban állt, hogy eddig együtt nem használt módszereket kötött össze lefedve a backcasting folyamat különböző szakaszeit. Ezek alapján kutatócsoportunk a backcasting kutatás részvételi technikákon alapu-

ló módszerek egy sajátos kombinációját alkalmazta: a *rendszermodellezés* módszerét a téma keretezésére (Király et al., 2014), a *világkávészó* módszert a vízió kialakítására (Gáspár et al., 2014), és a *jövőkerék* (Glenn, 2009) egy módosított változatát (Király et al., 2015) a szakpolitikai lépések visszafejtésére. Maguk a módszerek kölcsönösen kiegészítik egymást, hiszen eltérő 'belső logikával' bírnak:

- A *részvételi rendszermodellezés* folyamata úgy épül fel, hogy a résztvevők először változókat azonosítanak az adott témában, majd meghatározzák az ezek közötti kapcsolatokat; végül pedig – ha lehetséges – az e kapcsolatokból kirajzolódó visszacsatolási hurkokat tárják fel (Király et al., 2014).
- A *világkávészó* egy olyan kreatív folyamat, amely kifejezetten az együttműködésen alapuló beszélgetéseket és a tudásmegosztást hivatott elősegíteni azáltal, hogy az elhangzó gondolatokat egy egymással összefüggő hálózatba szervezi. A meghívottak kis, általában változó csoportokban, kávézóasztalok köré gyűlve vitatnak meg egy előre meghatározott kérdést vagy egy témát több alkalommal (Gáspár et al., 2014). Ebben az esetben a rendszermodellezés eredményeképpen azonosított hat központi téma szolgált az „asztaltársaságok” fókuszaként, és így ezek köré a témák köré szerveződtek a normatív vízió elemei is.
- A vízió kialakítása után, a vízióhoz vezető szakpolitikai beavatkozások azonosításához használtuk a *jövőkerék* módszert (Glenn, 2009). A technika formailag teljes mértékben a gondolatértékepezést követi: a jövőre hatást gyakorló eseményt vagy folyamatot egy lap közepére felírva annak elsődleges hatásait vagy következményeit körbe, kerekyszerűen össze kell gyűjteni. Ebben az esetben a hatások helyett azonban a jövőből visszafejtett lépések jelentek meg a „kerekekben”.

A backcasting alprojekt egyes szakaszai során kikristályosodó fő témákat foglalja össze az 1. ábra.

1. ábra A backcasting alprojekt egyes szakaszai során kialakult fő témák összefoglalása



Terjedelmi okok miatt nem mutatjuk be az egész folyamatot módszerenként. A következőkben csupán az oktatók által kidolgozott normatív víziót ismertetjük, mivel egyrészt véleményünk szerint ez az egyik eddig legérdekesebb eredménye kutatásunknak; másrészt pedig ez vethető, hasonlítható össze a legjobban a nemzetközi szakirodalomban fellelhető jövőképekkel.

Az oktatói panel normatív víziója a gazdasági felsőoktatás jövőjéről

Ahogy azt a fentiekben már röviden bemutattuk, a backcasting folyamat során a résztvevők egy normatív jövőképet dolgoznak ki azzal kapcsolatban, hogy milyennek kellene lennie a jövőnek egy megadott időpontban. Jelen esetben ez azt jelenti, hogy a résztvevők azt próbálták meghatározni, hogy 2050-ben Magyarországon milyen lenne, milyen jellemzőkkel bírna az ideális gazdasági felsőoktatás.

A következőkben a backcasting folyamat során a bevont 12 oktatóval és kutatóval kidolgozott víziót az általunk feldolgozott nemzetközi tanulmányokkal és néhány további – szintén víziókat megfogalmazó – írással (Facer, 2011, Darden, 2009) összevetésben mutatjuk be. A rendszermodellezési szakaszban azonosított hat témán (ld. 1. ábra) belül két összevonást végeztünk, így az alábbiakban a kutatásunkban kialakult vízió sajátosságait a szakirodalom fényében a következő négy téma mentén tárgyaljuk: (1) finanszírozás; (2) a tanítás tartalma és módja, valamint a hallgató-tanár kapcsolat; (3) infrastruktúra és (4) a hallgatók száma, összetétele és a hallgatói és oktatói motivációk. A vízió összefoglalását adja az 1. táblázat:

1. táblázat. A vízió fő témái és eredményei

Fő témák	Kapcsolódó sajátosságok	Általános jellemzők
Finanszírozás	finanszírozás hallgatói szinten, négy fő pillér: állami, személyes, vállalati, közösségi.	felsőoktatás magas autonómiája. Virtuális és személyes képzés párhuzamosan:
Infrastruktúra	Fizikai infrastruktúra: kis termek a tanításhoz, az épületek és nagy terek funkciója tipikusan reprezentatív és társadalmi célú. Virtuális infrastruktúra: jelentős információs technológiai háttér és szerviz-szolgáltat.	tárgyi tudás főleg virtuális kontextusban – szélesebb elérhetőség, személyre szabott oktatás főleg speciális képességek és készségek elsajátítására (pl. kritikai gondolkodás, probléma-megoldás stb.). Erős nemzetközi verseny.
A tanárok és hallgatók szerepe és kapcsolata	Tanárok, mint 'tudás-brókerek' (a hallgatói és vállalati igényekre reagáló tudásátadás). Aktív, öntudatos és célorientált hallgatók. 'Mester-tanítvány' viszony, intenzív személyes tanulási forma kiscsoportokban, kölcsönös bizalomra és visszacsatolásra építve.	A felsőoktatási intézmények 'tudás-értékelő' szerepe felértékelődik. Nagyfokú mobilitás (különösen virtuális környezetben). Vállalatok aktív szerepben: gyakorlati terep, vállalati egyetemek, speciális oktatási igények megfogalmazása.
A hallgatók száma és összetétele	Hallgatók magas száma. Nagyfokú diverzitás az életkor, a nem, a társadalmi státusz és a nemzetiség tekintetében.	Élethosszig tanulás, mint társadalmi jellemző.
A tanítás tartalma és módja	Friss és gyakorlat-orientált tudás az elméleti alapok mellett (választható). Inspiráló és élmény-alapú tanulás és tanítás. Projekt-alapú tanulás és tapasztalati formák.	Rugalmas és moduláris felépítésű felsőoktatás (egy szemeszter csakúgy elérhető, mint teljes képzések).
Oktatói és hallgatói motivációk	Tanári szabadság és autonómia a tanított tartalom és a tanítási módszerek tekintetében. Társadalmilag és anyagilag is megfelelően elismert tanárok. A tanári szerepek (tanár, kutató, gyakorlati szakember stb.) és az értékelési módok rugalmassága. Motivált hallgatók számos, de egyértelmű követeléssel. Aktív hallgatói részvétel projektekben és kutatásokban.	A felsőoktatás társadalmi felelőssége: nyílt és hozzáférhető tudás szolgáltatása, tudatos hallgatók (állampolgárok) képzése.

Finanszírozás

A felsőoktatás finanszírozásának kérdése az egyik leggyakrabban tárgyalt téma a szakirodalomban. A mi backcasting folyamatunk során is felvetették és beszéltek róla a résztvevők, de összességében közel sem játszott olyan központi szerepet, mint az a nemzetközi írásokban tapasztalható. Sokkal inkább úgy jelent meg, mint egyfajta – a minőségi oktatáshoz és kutatáshoz szükséges – háttérfeltétel. Ez a különbség részben azzal magyarázható, hogy a nemzetközi szakirodalomban talált írások legfőképp makro (társadalmi) és meta (globális) szinten vizsgálják a témát, míg ez a részvételi módszer sokkal inkább kiemeli a mikro (egyéni) és mezo (intézményi) szinteket.

A kutatási folyamat során kialakult vízió szerint a felsőoktatási finanszírozás négy pilléren nyugszik majd. Egyrészt az állam megmarad fontos finanszírozási partnernek, de a jelenleginél kisebb mértékben. Másrészt a vállalatok (globálisak és helyiek egyaránt) nagyobb részt vállalnak ezen a területen. Harmadrészt megjelenik a helyi közösség is a finanszírozók között, és negyedik pillérként maguk a hallgatók, azaz a személyi finanszírozás.

A finanszírozás legfontosabb sajátosságaként a résztvevők azt hangsúlyozták a kialakított vízióban, hogy az oktatás pénzügyi kereteinek biztosítása az előző forrásokból nem az intézményi, hanem a hallgatói finanszírozás formáján keresztül fog megvalósulni. Ez azt jelenti, hogy mindenkinek a rendelkezésére áll egy meghatározott keretösszeg ('voucher'), amelyet a maga által kiválasztott képzésekre, kurzusokra költethet el.

Ez az elképzelés egyfelől rezonál néhány általunk vizsgált írásra (Schwartz, 2013, Calderon-Mathies, 2013), amelyek azt hangsúlyozzák, hogy a finanszírozás és a minősítés ne input-kritériumok alapján történjen (pl. jelentkezők száma), hanem az intézmények output indikátorai alapján. Ennek ellenére a nemzetközi szakirodalomban továbbra is a hallgatói tandíj (*student fee*) jelenik meg, mint a felsőoktatás finanszírozásának a fő formája, hiába szembesül sok problémával ez a rendszer (Wenish, 2012).

A tanítás tartalma és módja, az oktatói és hallgatói szerepek és viszony

A vízió szerint az oktatás két fő pilléren nyugszik majd a jövőben: a virtuális (on-line) és a személyes (face-to-face) oktatáson. A tárgyi tudás átadásának terepe főként a virtuális tér, amely nagy tömegekhez képes eljuttatni jelentős mennyiségű információt. A résztvevők szerint ez az egyik legnagyobb kihívás, amivel a felsőoktatásnak számolnia kell, mivel az online oktatás egyaránt átalakíthatja mind a felsőoktatási intézmények szerepét, mind az egész felsőoktatási rendszert. Ez a vélemény a nemzetközi szakirodalomban is erőteljesen megjelenik, szinte mindegyik elemzett cikk foglalkozik a kérdéssel, még ha különböző jövőképeket is festenek ezzel kapcsolatban. A szakirodalomban megjelenik továbbá az a gondolat is, hogy az online oktatás jelensége más-más mértékben érinti a különböző tudományterületeket. Az elemzések szerint ugyanis sokkal mélyrehatóbb változásokat idézhet elő a társadalomtudomány (pedagógia, menedzsment, stb.) területén, mint például az orvosi vagy mérnöki területeken, ahol a laboratóriumi vizsgálatok, és a személyes (úgynevezett 'hands-on') aktivitás elengedhetetlen (Markert et al., 2013, Long nd).

A tudományterületi különbségektől eltekintve az online képzés opciója az oktatási mező minden szereplőjét érinti, hiszen a hozzáférhetősége, a rugalmassága és a növekvő fontossága nyomán alapjaiban változik meg az oktatási rendszer. A vizsgált cikkek

szerzői hangsúlyozzák, hogy online kurzusok és tananyagok szinte mindenki számára elérhetővé válnak, ugyanakkor sokkal személyre szabottabbak és diverzebbek lehetnek (Schwartz, 2013, Long, nd, Markert et al., 2013), mint a tömeges off-line oktatás. Így az online forma egyaránt magával hozhatja az oktatási és a tanulási formák és folyamatok változását is (Dew, 2010, Esson – Ertl, 2013). Néhány írás azt is megemlíti, hogy ez az új szintér megváltoztatja az intézmények közötti verseny feltételeit is. A fennmaradás-hoz a felsőoktatási intézmények közötti szorosabb együttműködés, valamint a vállalatokkal kialakított közös programok is szükségesegek lesznek (Gee, 2014, Schwartz, 2013, Darden, 2009).

Mindezek ellenére az online oktatással párhuzamosan a személyes tanítás sem tűnik el, és ezt a backcasting workshop résztvevői többször is kiemelték. Ahogy a vízió is bemutatta, ez az oktatási forma megmarad ugyan, de jelentősen átalakul az elképzelések szerint 2050-re. Egyfelől a tömeges oktatás helyett a személyre szabott és a speciális képességek és készségek (például kritikai gondolkodás, probléma-megoldás) fejlesztésére fókuszál. Másfelől tartalmát tekintve 'mester-tanítvány' formájú, személyes tudásátadásként fog működni. Ez az új típusú kapcsolat – ahogy azt a résztvevők kiemelték – nem a tanár-diák közötti aszimmetrikus viszonyra alapul, hanem egy kiscsoportos, kölcsönös bizalmon és visszacsatoláson alapuló kapcsolatot feltételez, ami nemcsak a kiválasztott kevesek számára (például aki meg tudja fizetni, vagy a legjobb) érhető el, hanem mindenki számára, aki ezt a fajta aktív tanulást szeretné választani. Bizonyos esetekben pedig ez a forma éppen a sajátos tanulási igényű, vagy speciális oktatási szükségletű hallgatók számára lenne elérhető.

A kidolgozott vízió szerint az oktatók 2050-ben egyfajta 'tudás-bróker' szerepet töltenének be az oktatási rendszerben. Ebben a szerepben feladatuk az lenne, hogy kezeljék a számos, de világosan megfogalmazott hallgatói igényt, valamint ezeket és az oktatási kínálatot összhangba hozzák a vállalati, munkaerő-piaci igényekkel. Ennek megfelelően a tananyagokat és a tanítási módokat is nagyfokú rugalmasság jellemzi, és fontossá válik a projektszemléletű és a tapasztalati tudást átadó képzési forma.

Infrastruktúra

A személyes oktatás formájának átalakulása nyomán (lásd előző pont) az oktatási épületek szerepe és fontossága is nagymértékben megváltozik. Ahogy a vízióban megfogalmazták a résztvevők, bár 2050-ben az oktatási épületek nem tűnnek el, de a személyes oktatási célokat szolgáló kisebb termek mellett a nagy előadók és aulák már csak reprezentatív és társadalmi funkciókat látnak el. A korábbi oktatási terek (iskolaépületek, kollégiumok) nem a tömeges oktatás, hanem sokkal inkább a(z) – oktatói és hallgatói – szocializáció, a közösségi élet terei lesznek az ehhez szükséges kiegészítő funkciókkal felvértezve (például kávézókkkal, sportolási lehetőségekkel, civil szervezetek jelenlétével stb. [ilyen típusú tereket mutat be például Facer, 2011, pp. 109-123]).

A workshopunk résztvevői kiemelték ezeknek a közösségi aktivitásoknak és tereknek a fontosságát, mint amelyek nélkülözhetetlen szerepet töltenek be a személyes élmények és kapcsolatok kialakításában. Ez egybevág Long (nd) megállapításával is, aki szerint: *„semmi sem hasonlítható ahhoz, mint amikor pár hallgatóval kis asztal körül ülve Immanuel Kant 'transzcendentális esztétikájáról' vagy John Rawls 'tudatlanság fátyláról' beszélgetünk. Az alapképzés során megélt négyévnnyi élmény – bár vélhetően átalakul – nem fog eltűnni.”* (Long nd, 60. o.)

A vízió szerint az online oktatás tekintetében pedig az intézmények hatalmas és folyamatosan fejlődő virtuális infrastruktúrával, valamint az ehhez szükséges szolgáltatásokkal egyaránt rendelkeznek.

A hallgatók száma, összetétele és a hallgatói és oktatói motivációk

A fentiekben bemutatott intézményi változásoknak, a nagyfokú mobilitásnak és az online képzések általános hozzáférhetőségének köszönhetően – a vízió szerint – 2050-re a hallgatói létszám növekedni fog, és sokkal vegyesebb nemzetiségi összetételt fog mutatni, mint manapság. Az élethosszig tartó tanulás elterjedése révén a hallgatók korösszetétele is a jelenleginél sokkal sokszínűbb lesz.

Ennek nyomán a vízióban megjelenő gazdasági képzést nagyfokú rugalmasság és moduláris felépítés jellemzi. Ez azt jelenti, hogy többféle mértékben és szintig vehetők igénybe a különböző oktatási szolgáltatások (a paletta egy-egy kurzustól teljes képzésekig terjed).

A vízió szerint 2050-ben a felsőoktatási intézmények magas autonómiával bírnak, ami kiegészül tudás-értékelő és –minősítő funkciókkal is. Ehhez kapcsolódik a felsőoktatás felelősségének kérdése is, ami egyrészt megjelenik nyílt és hozzáférhető tudás szolgáltatásában, másfelől a tudatos hallgatók (és ezáltal tudatos állampolgárok) képzésében.

Az autonómia a résztvevők szerint egyéni szinten is érvényesül a vízióban. Az oktatóknak ugyanis nagy szabadsága van abban is, hogy mit és hogyan tanítanak. Emellett az oktatói pálya társadalmi és anyagi szempontból is elismertnek számít majd a jövőben. Ezek a tényezők pedig mind azt segítik elő, hogy a jövő felsőoktatási intézményeinek oktatói motiváltak.

Ha ezeket a pontokat a nemzetközi szakirodalomra vonatkoztatjuk, látható, hogy az oktatók és a hallgatók motivációjának kérdésköre nem jelenik meg a vizsgált cikkekben, de a mobilitást és a hallgatók nemzetközi összetételének színesedését több helyütt is megemlítik (*Huisman et al.*, 2012, *Dew*, 2010, *Esson – Ertl*, 2013, *Long*, nd, *Calderon – Mathies*, 2013).

Visszatekintés és további tervek

A projektünk szíve az *oktatói backcasting műhely* volt. A műhely két napból állt, amely során a résztvevők megalkották azt a normatív jövőképet, hogy szerintük milyen lenne az ideális gazdasági felsőoktatás 2050-ben. Érdeemes megjegyezni, hogy a nemzetközi szakirodalomban tárgyalt témákhoz képest új szempontok is megjelentek ebben a jövőképben. Ezek leginkább az oktatás „mikro-szintű” jellemzőire vonatkoztak, mint az oktatók és a hallgatók közötti kapcsolat minősége, valamint a motiváció kérdése. Mivel a nemzetközi szakirodalom általunk feldolgozott korpuszában megtalálható jövőképek leginkább makroszinten vázolták felsőoktatás előtt álló jövőbeni utakat, ezek a mindennapi oktatási gyakorlatokat meghatározó (mikro-szintű) tényezők kevésbé jelentek meg bennük.

Kutatócsoportunk jövőbeli terveit tekintve a fent ismertetett három pillér mentén további irányok látszanak, amelyek egybevágna a kutatás céljaival. Egyrészt – azért, hogy pontosabb képet kapjunk a felsőoktatás szereplőinek jövőképéről – a kutatás kezdetektől fogva legalább két csoport bevonását tervezte. A felsőoktatásban dolgozó oktatók, kutatók, kutatásszervezők csoportján túl, a jövőben szeretnénk megismerni a hallgatók gondolkodását is. Ez lehetővé teszi a két csoport vízióinak és szakpolitikai javaslatainak összehasonlítását. Középhosszú távú tervként felmerült továbbá a vállalati szféra, valamint

a most doktori képzésben résztvevők jövőképeinek feltárása is. Ehhez használnánk a másik két pillér további alprojektjeit (hazai szakirodalom feldolgozása, gazdasági képzőhelyek hallgatói véleményének feltárása félig strukturált interjúkkal).

IRODALOM

- CALDERON, A. – MATHIAS, Ch. (2013): Institutional Research in the Future: Challenges within Higher Education and the Need for Excellence in Professional Practice. *New Directions for Institutional Research* no. 157 pp. 90.
- DARDEN, M. L. (2009): *Beyond 2020. Envisioning the Future of Universities in America*. Rowman & Littlefield Publishers, Maryland
- DEW, J. R. (2012): The future of American higher education. *World Future Review* 4(4) pp. 7-13
doi: 10.1177/194675671200400403
- ESSON, J. – ERTL, H. (2013): Securing the Future of Higher Education in England. *The Political Quarterly* 84(3) pp. 404-409
doi: 10.1111/j.1467-923X.2013.12031.x
- FACER, K. (2011): *Learning Futures – Education, technology and social change*. Routledge, New York
- GÁSPÁR T., KIRÁLY G. & CSILLAG S. (2014): Fehér asztal mellett: A világkávészó részvételi technika szemlélete és módszertana. *Kovács XVIII*(1-4.) pp. 11-41.
- GEE, G. (2014): Recalibrating the Future of Higher Education. *Indiana Law Review* 47(1) pp. 272-278.
- GLENN, J. C. (2009): *Futures Wheel, Futures Research Methodology Version 3.0*, The Millennium Project, Washington, D.C.
- HUISMAN, J., DEBOER, H. & BÓTAS, P. Ch. P. (2012): Where Do We Go from Here? The Future of English Higher Education. *Higher Education Quarterly* 66(4) pp. 341-362.
- KIRÁLY G., CSILLAG S., GÁSPÁR T., GÉRING Zs., KÖVES A. & SEBES-TYÉN G. (2015): *Jelentés az 'Iskola a jövőben: a gazdasági felsőoktatás Magyarországon 2050-ben' projekt kutatási eredményeiről*. Kutatási jelentés. BGF. Elérhető: <http://publikaciotar.repozitorium.bgf.hu/622/>
- KIRÁLY G., PATAKI GY., KÖVES A. & BALÁZS B. (2013): Models of (future) society: Bringing social theories back in backcasting. *Futures* 51(July) pp. 19-30.
- KIRÁLY G., KÖVES A., PATAKI GY. & KISS G. (2014): Rendszermodellezés és részvétel: egy magyar kísérlet tanulságai. *Szociológiai Szemle* 24(2) pp. 90-115.
- LONG, Ch. (nd): *The Changing Face of Higher Education: The Future of the Traditional University Experience*. www.harvard-kennedyschoolreview.com
- MARKERT, L. R., MERILL, CH., REEVE, E. & SEYMOUR, R. (2013): The future of technology and engineering higher education. *Technology and engineering teacher* December/January pp. 35-42.
- OECD (2008): Four Future Scenarios for Higher Education. *OECD/France International Conference Higher Education to 2030: What Futures for Quality Access in the Era of Globalisation?*
- REDECKER, C., LEIS, M., LEENDERTSE, M., PUNIE, Y., GIJSBERS, G., KIRSCHNER, P., STOYANOV, S. & HOOGVELD, B. (2011): *The Future of Learning: Preparing for Change*, European Commission, Joint Research Centre, Institute for Prospective Technological Studies, Luxembourg, pp. 97.
- SCHWARTZ, S. (2013): The future of higher education: faster, cheaper, better. *Policy* 29(2) pp. 3-9
- WENISCH, M. (2012): Student Loan Crisis and the Future of Higher Education. *The Catholic Social Science Review* 17 pp. 345-350.

A tanári tevékenység mint a tanár-diák konfliktusok forrása

Kutatásomban a tanár-diák konfliktusokra, s ezen belül azokra az összetűzésekre fókuszáltam, melyeket a tanárok különböző tevékenység- és magatartásformái váltottak ki. Vizsgálatom célja az volt, hogy a tanár-diák konfliktusokra egy, a hagyományostól eltérő megközelítésben világítsak rá: a pedagógusok konfliktust kiváltó viselkedésének szempontjából. Milyen gyakran, hogyan és mivel váltanak ki konfliktust a tanárok saját maguk és tanulóik között?

A kutatás elméleti háttere

A tanárok és a diákok az iskola legfontosabb szereplői. A köztük lévő kapcsolat számos módon meghatározza a tanítási-tanulási folyamatot: hatással van a tanítás hatékonyságára (Handley, é. n.), befolyásolja a tanulók fejlődését és teljesítményét, illetve a diákok iskolához és tantárgyhoz fűződő attitűdjét is (Hamre és Pianta, 2006).

Mint minden emberi kapcsolat, úgy a tanár és diák kapcsolata is konfliktusokkal tarkított. A konfliktus szót különböző érdekek, nézetek ellentétének leírására használjuk (Szekszárdi, 2002). A pedagógus és a tanuló között kialakuló konfliktusok általában személyesek és közvetlenek. Kialakulásuknak számos oka lehet: tanárok és diákok különböző nézetei, a diák személyisége, rossz magatartása, a tanár személyisége, tanári értékbizonytalanságok, stb. (Szekszárdi, 2002; Hunyady, 2006). A tanár-diák konfliktushelyzetekben a tanárok és a legtöbb kutató is a tanulói cselekedeteket (óra megzavarása, hazugság, késés, csavargás) tekintik a leggyakoribb konfliktusforrásoknak, azonban nem szabad elfelejtenünk, hogy nemcsak a diák, hanem a pedagógus tevékenysége is lehet konfliktusforrás, például, ha görcsösen ragaszkodik a hatalomhoz, túlterheli a diákokat, vagy kategorizál bizonyos sztereotípiák mentén.

A pedagógusok gyakran helytelen viselkedésükkel váltanak ki konfliktust, ami olyan tanári hibák formájában mutatkozik meg, melyek negatív hatással lehetnek a tanulók életére mind tanulmányi, mind pedig pszichoszomatikus szempontból (romló tanulmányi eredmények, impulzivitás, iskolafóbia, önbizalomhiány) (Sava, 2002). Az Orange (2008) által összegyűjtött leggyakoribb tanári hibák a helytelen fegyelmezési módszerek használata, a megszegényítés, a kivételezés, a személyeskedés, illetve a kategorizálás.

► Educatio 2015/3. Györfiné Körtvélyesi Franciska: A tanári tevékenység, mint a tanár-diák konfliktusok forrása, 116–120. pp.

Minden konfliktushelyzetben a fő cél a konfliktus kezelése. Alapvetően öt konfliktusmegoldó stratégiát különböztethetünk meg: a versengő, az önalávető, az elkerülő, a kompromisszumkereső és a problémamegoldó módszereket (Szekszárdi, 2002).

Szekszárdi Júlia (2002) szerint a posztmodern kor új kihívásai miatt ma kiemelten fontos lenne, hogy a tanárok jártasak legyenek a konstruktív konfliktuskezelésben. A gyakorlatban azonban a tanári konfliktuskezelés más képet mutat. Horváth Szabó Katalin (1991), illetve Sági Matild és Szemerszki Marianna (2012) kutatása alapján a pedagógusok többsége tekintélyelvű konfliktuskezelő módszert alkalmaz a diákokkal előforduló konfliktusok rendezésére, ami azonban a tanár-diák viszony elhidegülését eredményezi.

A tekintélyelvű módszerrel ellentétben a problémamegoldás esetén nem az a fontos, hogy a pedagógus győzzön, hanem az, hogy a tanulóval együtt oldja meg a problémát. Ezt azonban nagyon nehéz egy alapvetően egyenlőtlen kapcsolatban megvalósítani, melyben a gyengébb félnek kell az erősebbhez alkalmazkodni (Szőke-Milinte, 2006. 138. o.). Ennek alapján egyfajta szemléletváltásra lenne szükség ahhoz, hogy tanárok és tanulók valódi partnerek lehessenek, s konfliktusaikat konstruktívan tudják kezelni. A cél egy olyan partnerkapcsolat kialakítása a tanárok és diákok között, melyben mindkét fél felszólalhat, mindenkit meghallgatnak, s melyben a pedagógus nem a tekintélyét félti és védelmezi, hanem a tanulókkal együttműködve próbál segíteni. Ennek következtében a destruktív tanár-diák konfliktusokat felválthatnák a konstruktív konfliktusok, melyek napvilágra hoznák a nézetkülönbségeket és érdekegyeztetésre adnának lehetőséget.

A kutatásról

A korábbi tanár-diák konfliktusokkal kapcsolatos kutatások inkább a diákok konfliktust kiváltó magatartására fókuszáltak. Tény, hogy gyakrabban alakul ki konfliktus a tanulók fegyelmezetlensége, tiszteletlensége miatt, azonban nem szabad arra a következtetésre jutnunk, hogy a tanárok soha nem kezdeményeznek konfliktusokat. Vizsgálatom azt a megállapítást igyekszik alátámasztani, miszerint az iskolában nem ismeretlenek az olyan tanár-diák összetűzések, melyeket a tanárok helytelen viselkedése váltott ki.

Kutatásomban három fő kérdést fogalmaztam meg:

1. Milyen gyakran fordul elő, hogy a tanár tevékenysége, viselkedése vált ki konfliktust a tanárok és a diákok között?
2. Mik azok a tanári magatartásformák, melyek konfliktusokat eredményeznek tanárok és diákok között?
3. Hogyan kezelik a konfliktusokat a diákok és a pedagógusok a tanári viselkedés által kiváltott konfliktushelyzetekben?

A kutatást egy budapesti művészeti szakközépiskolában végeztem. Fontosnak tartottam a pedagógusok és a tanulók véleményét is megkérdezni, hogy komplex képet kaphassak a tanári tevékenység által kiváltott tanár-diák konfliktusokról. Az iskola pedagógusainak véleményét részben strukturált interjúk segítségével ismerhettem meg, a diákok álláspontjának megismeréséhez két kutatási módszert alkalmaztam. Egyrészt kérdőív segítségével tártam fel a tanulók véleményét az osztályukban előforduló, a pedagógus által kiváltott tanár-diák konfliktusokról, másrészt olyan internetes fórumtémákat elemeztem, melyek a tanári tevékenység által kiváltott tanár-diák konfliktusokról szóltak. Összességében öt tanár és három osztály diákjai vettek részt a kutatásban.

Eredmények

A kérdőívet összesen 74 (10-edikes és 11-edikes) tanuló töltötte ki, 61 lány és 13 fiú. A pedagógusok közül pedig két matematika, egy magyar, egy művészettörténet és egy rajz szakos tanárral készítettem részben strukturált interjút.

A megkérdezett tanulók majdnem háromnegyede (71%) szerint, az osztályukban átlagban kb. hetente egyszer fordul elő olyan tanár-diák konfliktus, amit a pedagógus viselkedése váltott ki.

A diákok összesen 107 konfliktust kiváltó tanári viselkedésformát sorolnak fel, melyeket három kategóriába lehet sorolni: a visszaélés a hatalommal (33%), a kivételezés (30%), és a kategorizálás (37%) osztályaiba. A megkérdezett öt tanár is elismerte, hogy előfordul, hogy az ő magatartásuk vált ki konfliktust: sokszor kivételeznek egy-egy tanulóval (általában pozitív irányba), továbbá gyakori az előítéleteken alapuló kategorizálás és az igazságtalan osztályozás is (bár ez általában nem tudatos a részükről).

Azt, hogy miként gondolkodnak a diákok a tanárok viselkedése által kiváltott konfliktusokról, nagyban meghatározza, hogy hogyan gondolkodnak magáról a tanár-diák viszonyról. A kérdőív eredményei alapján elmondható, hogy a diákok 73%-a szerint a tanár-diák viszony hierarchikus, azaz alá-fölérendelt viszonyról van szó. Ezzel ellentétben a pedagógusok egyértelműen pozitívnak találták a tanár-diák kapcsolatot az iskolában – szerintük nem alá-fölérendelt, hanem partneri kapcsolatban vannak a tanulóikkal.

Megosztó volt a tanulók és pedagógusok véleménye a tanári és tanulói konfliktuskezelésre vonatkozó kérdéseknél. A diákok szerint a leggyakoribb tanulói reakció az ilyen konfliktushelyzetekben az, hogy a tanulók nem vállalják fel a konfliktust, de később társaik előtt szidalmazzák a tanárt, míg a legkevésbé gyakori tanulói reakciónak a bosszúból történő fegyelmezetlenséget jelölték meg. A tanárok véleménye az általuk kezdeményezett konfliktusokra adott tanulói reakciók tekintetében egymásnak ellentmondóak. Három pedagógus szerint a tanulók leggyakrabban „visszaszólnak” a tanárnak, megvédik magukat. Az órai fegyelmezetlenséget két tanár is jellemzőnek tartja. Két másik tanár szerint azonban a diákok többsége nem szól, ha sérelem éri, helyette inkább a társai előtt szidalmazza a tanárt.

A tanulói válaszok alapján a tanárok konfliktuskezelési módszereinek rangsorolásánál a tekintélyelvű módszer bizonyult a leggyakoribbnak, míg a legritkább tanári reakció a bocsánatkérés a tanulók szerint. A tanárok a saját konfliktuskezelésükről alkotott véleménye nagymértékben eltér a diákok véleményétől. Három pedagógus szerint, ha ők okozták a konfliktust, akkor ezt bevallják, és bocsánatot kérnek a diáktól. Szintén hárman a problémamegoldást (is) alkalmazzák az ilyen esetekben. Egy tanár mondta azt, hogy ő gyakran inkább elkerüli a konfliktust, vagy tekintélyelvűen lezárja azt.

Az eredmények értékelése

Az ismertetett eredmények alapján számos fontos következtetés vonhatunk le. A megkérdezett tanulók és pedagógusok véleménye is alátámasztja, hogy a tanár tevékenysége által kiváltott tanár-diák konfliktus létező jelenség az iskolában; a tanulók háromnegyede szerint átlagban legalább hetente egyszer előfordul. A konfliktust kiváltó tanári viselkedésformák három kategóriába sorolhatóak: a visszaélés a hatalommal, a kivételezés és a kategorizálás.

A tanulói és tanári konfliktuskezelésre vonatkozó eredményeket tekintve is érdekes következtetésekre juthatunk. A pedagógus viselkedése által kiváltott konfliktushelyzetekben a diákok szerint a tanár szidása, a tanárok szerint pedig a „visszaszólás” a leggyakoribb tanulói reakció. Ez az ellentét abból következhet, hogy a tanárok leginkább a diákok felszólalásából jönnek rá arra, hogy magatartásukkal konfliktust váltottak ki, míg az, hogy egy diák szidalmazza őket, rejtve marad előttük. Az azonban a tanárok és a diákok szerint is szórványosan fordul elő, hogy a diák egyáltalán nem reagál a konfliktusra. Vagyis a tanulóknak fontos kifejezniük, hogy sérelem érte őket, amit viszont leginkább a társaik felé mutatnak ki, kevésbé gyakori, hogy a tanárnak jeleznék. Ennek hátterében a diákok tanár-diák kapcsolatról alkotott képe állhat: úgy gondolják, hogy ők alárendelt szerepben vannak a tanárokkal szemben, ezért értelmetlen szembeszállni a tanárokkal, hiszen így is, úgy is a pedagógusnak, mint a hatalom képviselőjének lesz igaza.

A tanári reakcióknál a tekintélyelvű megoldás, a diák szidalmazása a leggyakoribb reakció a tanulók véleménye szerint; ritka, hogy a tanárok beismerik helytelen viselkedésüket és bocsánatot kérnek a megbántott diáktól. Ezzel ellentétben, a pedagógusok szerint ők bocsánatot kérnek, ha ők okozták a konfliktust, illetve a probléma megoldására törekednek. A diákok és tanárok véleményének ellentmondására a tanári önreflexió jelenthet megoldást: ha a pedagógusok gyakran reflektálnak a tanulókkal való bánásmódjukra, a közöttük és a diákok között kialakult konfliktusok előzményeire és okaira, akkor gyakrabban fény derülhet arra, hogy valóban az ő magatartásuk eredményezte az adott konfliktust. Ilyen helyzetben már valószínűbb, hogy a pedagógus bevallja, hogy hibázott és ténylegesen bocsánatot kér a megbántott diáktól (ahogy ezt a megkérdezett pedagógusok is állítják).

A kérdés és a tartalomelemzés eredményeinek összevetése

Az internetes fórumok elemzése teljes mértékben alátámasztja a kérdőív és az interjúk eredményeit. Összesen öt fórumtémát elemeztem a gyakorikerdesek.hu oldalon, melyek kivételezésről, kategorizálásról és tanári igazságtalanságokról szólnak.

A fórumokon megemlített tanári magatartásformák is besorolhatók a fent megnevezett három kategóriába, ami még inkább alátámasztja a kérdőívek és az interjúk alapján megfogalmazott következtetést, miszerint a leggyakoribb tanári konfliktust kiváltó viselkedésformák a kategorizálás, a kivételezés és a hatalomhoz való ragaszkodás.

Nagyon érdekes felvetéseket lehet olvasni a tanárok konfliktust kiváltó viselkedésének hátterében álló okokkal kapcsolatosan. Egyes hozzászólók szerint lehetséges, hogy a tanárok a diákok fegyelmezetlensége, tiszteletlensége miatt frusztráltak és ezt a feszültséget a tanulókon vezetik le. Más fórumozók szerint azonban a diákok fegyelmezetlensége, tiszteletlensége pontosan a tanári hatalmi visszaélések következményei.

E két álláspont egy fontos kérdésre hívja fel a figyelmet: Vajon mi minnek a következménye? A tanárok a diákok tiszteletlensége miatt felgyülemelő feszültséget vezetik le a tanulóikon igazságtalanságok formájában, vagy pont a tanári hatalmi visszaélések azok, amik a diákokat tiszteletlenségre ösztönzik? Ez egy ördögi kör, hiszen az egyik kiváltja a másikat, ami további konfliktusokhoz vezet. Ezt az ördögi kört csak az állíthatná meg, ha a tanárok fel- és beismernék, hogy ők is kiváltanak konfliktusokat, s tudatosan törekednének a konfliktusokhoz vezető magatartásformák elkerülésére. A diákok feladata pedig,

hogy merjenek felszólalni, hiszen így tudnak segíteni a tanároknak abban, hogy felismerjék, mikor viselkednek helytelenül.

Záró gondolatok

A kutatás eredményei természetesen csak a kutatás helyszínéül szolgáló szakközépiskola pedagógusaira és diákjaira érvényesek, azonban úgy gondolom, hogy a levont következtetések nem csak ebben az iskolában helytállóak. A kutatás legfontosabb megállapítása az, hogy a pedagógusok viselkedése által kiváltott tanár-diák konfliktus létező jelenség az iskolában. Ennek ellenére mégis kevés hangsúlyt fektetünk erre a problémára. A kutatás eredményei megvilágítják a tanári konfliktust eredményező magatartásformákat: a hatalommal való visszaélést, a kategorizálást és a kivételezést, melyek sokszor nem is tudatosak a pedagógusok részéről. A kutatás arra is rávilágít, hogy a tanulók hierarchikus tanár-diák kapcsolatban gondolkoznak, ami közrejátszik abban, hogy a tanár viselkedése által kiváltott konfliktusok „problematikája” továbbra is tabutéma maradjon. A diákok többsége nem meri a tanároknak jelezni, ha megsértették, helyette a társaiknak panaszkodnak, aminek következtében a tanárok gyakran nem ismerik fel, hogy az ő tevékenységük váltotta ki a konfliktust.

IRODALOM

- HAMRE, B. K. & PIANTA, R. C. (2006): Student-teacher relationships. In: Bear, G. G. és Minke, K. M. (szerk.): *Children's needs III: Development, prevention, and intervention*. Bethesda, MD: National Association of School Psychologists. 59–71. URL: <http://www.pearweb.org/conferences/sixth/pdfs/NAS-CBIII-05-1001-005-hamre%20&%20Pianta%20proof.pdf>. Utolsó letöltés 2014. október 13-án.
- HANDLEY, G. D. (é. n.): *Personality, learning and teaching*. Routledge and Kegan Paul Limited, London.
- HORVÁTH SZ. K. (1991): *Konfliktusok az iskolában*. Edukáció, Budapest.
- HUNYADY GY-NÉ (2006): Konfliktusok az iskolában. In: Golnhofér Erzsébet (szerk.): *A gyakorlati pedagógia néhány alapkérdése*, 6. kötet: *Az iskolák belső világa*. ELTE PPK, Budapest. 98-115. URL: <http://mek.niif.hu/05400/05468/05468.pdf>. Utolsó letöltés 2014. október 13-án.
- ORANGE, C. (2008): *25 biggest mistakes teachers make and how to avoid them*. Corwin Press, Thousand Oaks.
- SAVA, F. A. (2002): Causes and effects of teacher conflict-inducing attitudes towards pupils: a path analysis model. *Teaching and Teacher Education*, 18. 8. sz. 1007-1021. URL: <http://www-psychology.concordia.ca/fac/kline/sem/qicss/sava.pdf>. Utolsó letöltés 2014. október 13-án.
- SÁGI M.& SZEMERSZKI M. (2012): Hogyan hatnak a tanárok a diákok konfliktusmegoldási beállítódására? *Új Pedagógiai Szemle*, 62. 7-8. sz. 22-42.
- SZEKSZÁRDI J. (2002): *Konfliktusok pedagógiája*. Veszprémi Egyetemi Kiadó, Veszprém.
- SZŐKE-MILINTE E. (2006): *Konfliktuskezelés és pedagógusmesterség. A pedagógusok kríziskommunikációjának fejlesztése*. Országos Pedagógiai Könyvtár és Múzeum, Budapest.

Kisgyermekkori nevelés Magyarországon

A kezdetektől az I. világháború végéig (1828-1918)

A mai értelemben vett óvodák hosszú történelmi fejlődés eredményei. Az ipari fejlődés útjára lépő nyugat-európai országokban hamar felmerült a kisgyermekek valamilyen intézményes gondozásának, nevelésének igénye, ennek oka elsősorban a szegénységben élő nők tömeges kényszer-munkavállalása volt. Ezzel ellentétben, Magyarországon az óvodák megjelenését elsősorban nem az anyák munkába állásának szükségessége katalizálta, hanem az, hogy a reformkor haladó politikusai nemzeti ügynek, a hazafias nevelés színterének tekintették a kisdédóvó intézeteket.

A magyarországi óvodatörténet, az európai mintához hasonlóan, Samuel Wilderspin *Infant Education* (London, 1825) című művének német nyelvű kiadásával kezdődött. (Hornyák, 2004). Wilderspin elévülhetetlen érdemeket szerzett az óvoda eszméjének európai terjesztésében, annak ellenére, hogy az iskolában alkalmazott nevelési tartalmakat és módszereket vezette be az óvodába is.

Magyarországon Brunszvik Teréz alapításában 1828. június 1-jén Budán nyílt meg az első óvoda. Közép-Európában ez volt az első ilyen típusú intézmény, amelyet hamarosan számos intézet követett az ország területén.

Brunszvik idejében még nem beszélünk a mai értelemben vett óvodákról. A kor viszonyainak ismeretében korszerű intézménytípusok voltak a kisgyermekiskolák, ahol a rájuk jellemző módszertani sajátosságokkal, didaktikai alapelvek alapján folyt a széles körű elemi oktatás. A tantervi anyagban megjelentek az elemi iskola követelményei, a gyermekek a nap jelentős részét tanulással töltötték, és az adott kornak megfelelően magyar és német nyelven folytak a foglalkozások. Brunszvik elsősorban a rászoruló, ártatlan gyermekek megóvását, védelmét tűzte ki célul, amely a korabeli intézmények elnevezésében is megjelent: kisdédóvó, illetve Kleinkinder – Asyle, azaz kisgyermekmenhely.

Az óvodai tanítók továbbképzéséről az 1837-ben Tolnán létrejött képzőintézet gondoskodott. Az intézetben az elméleti oktatást követően a növendékek a „példány óvóintézetben” (Kövér, 1987, 19.) gyakoroltak (ez már a mai gyakorló óvoda megfelelője). Az első képzőintézet igazgatója Wargha István akadémikus volt. Az előtanulmány-

okkal rendelkező férfiak jelentkezése tette lehetővé, hogy képzésük csak egy évig tartson. Wargha pedagógiájának alapja, hogy: „Az egész gyermek úgy vétezik, amint van, nem mint lennie kell.” (Wargha István, idézi Kövér, 1987, 20.) A kisdédóvókban továbbra is jelen volt az elemi ismeretek tanítása, amelynek gyakorlati oka az volt, hogy a gyermek rendelkezzen a számára szükséges elemi ismeretekkel, ha a népiskolát nem kezdhetné el. Az ismeretátadás a gyermeki érdeklődésre, cselekvésre, szemléltetésre épített.

Az 1867-es kiegyezést követő gazdasági változások és a polgári fejlődés velejárójaként felértékelődött az intézményes nevelés jelentősége. Nőtt az óvodák száma és egyre nagyobb lett az igény az itt folyó tevékenység egységesítésére, amit az óvodákat korábban működtető egyesületek már nem tudtak biztosítani. Mindeközben Friedrich Fröbel pedagógiája is egyre ismertebbé vált hazánkban. Fröbel pedagógiai filozófiájának alapja az volt, hogy a gyermeket az általa kifejlesztett foglalkoztató eszközök segítségével a gyermeki tevékenységen keresztül kell fejleszteni. Fröbel hatására kezdik el kindergartennek,¹ gyermekkertnek hívni a kisgyermekkel foglalkozó intézményeket, ahol eleinte a foglalkozások is német nyelven folytak.

A népiskolák állami szabályozásának rendezése után az óvodák működését is szabályozták. 1891-ben jelent meg a kisdédóváról szóló törvény (1891: XV. t.c. Kisdédóvási törvény), amelyben rögzítették az óvoda feladatát, a kisdédóvás célját. A törvény szabályozta az óvodák belső életét, és egyértelművé tette, hogy az óvodákban nincs helye az iskolai típusú oktatásnak. Az óvodák fenntartását az állam felügyeli, és kötelezi a szülőket a 3-6 éves gyermekek óvodába járatására ott, ahol az elérhető. A törvény hatására tovább fejlődött az óvodai nevelés ügye és számos óvoda jött létre. A törvény további eredménye, hogy gondoskodott a képzőintézetek elterjedéséről is. A törvény életbelépése után a hazai óvodapedagógiára a nemzeti szellemben folyó nevelés volt a jellemző. A nemzeti irányzat sajátossága a népies jegyek hangsúlyozásában, a hazánkban természetesen fellelhető anyagok használatában nyilvánult meg, amelynek képviselője Peres Sándor volt. Óvoda-pedagógiával foglalkozó elődeihez képest kiemelten foglalkozott a játék fontosságával, a szemléltetéssel, az óvodai napirenddel, módszertani javaslatokkal.

Óvoda a két világháború között (1918-1945)

A magyar óvodatörténetben ezt az intervallumot további két, egymástól jól elkülöníthető szakaszra lehet szétválasztani:

1. 1920-1934/36 – az óvoda relatív függetlenségének korszaka,
2. 1934/36 – 1945 – az állami befolyás növekedésének korszaka.

Az első szakaszra a reformpedagógia (elsősorban a Montessori-pedagógia) megjelenése, illetve e szemlélet elterjedése nyomja rá a bélyegét, ugyanakkor az óvodafelfogás magában megőrizte a magyarországi kialakuláskor jelenlévő romantikus jellemzőket is. Ez a

¹ Ettől kezdve a korábban használatos „*infant school*” magyar megfelelőjét német mintára a „*kindergarten*” váltja fel. Az itt dolgozó pedagógusokat gyermekkertésznőként illették, hiszen ekkor már többségében nő pedagógusok foglalkoztak a gyermekekkel. Az angol nyelvben használatos óvoda megfelelője, a magyar óvodai szemléletet kifejezve, a „*kindergarten*,” az itt dolgozó pedagógus a „*kindergarten pedagogue*,” vagy „*early childhood educator*,” ami mindkét nem esetében használatos. A magyar óvodapedagógiai szemlélet szerint a magyar nyelvben nem helyes az angol „*nursery school*” vagy „*preschool*” kifejezés, mert nem tükrözi a tudományterület hazai sajátosságait, mint ahogyan az itt dolgozó pedagógus sem „*nursery school* vagy *preschool teacher*.”

két hatás egyszerre érvényesülve tudott sokáig (1936) ellenállni egyrészt a politikai, másrészt az 1940-es évektől érvényesülő iskolásító törekvéseknek.

A Montessori-hatás

Már az I. világháború előtt megjelenik az első olyan ismertetés a Kisdednevelés szakmai folyóiratban Várnai Sándortól, amely a Montessori-pedagógiáról értekezik (Ő. *Várnai Sándor*, 1912. 63–69. idézi Kövér 1987, 66.), de az elterjedésére az 1920-as évek végéig várni kellett. A Montessori-pedagógia legfőbb hazai népszerűsítője, Bélaváry-Burchard Erzsébet 1923-ban részt vett Hollandiában a Montessori nemzetközi tanfolyamon; miután hazatért, elkezdte szervezni saját intézményét és elkezdte Montessori írásait magyar nyelvre fordítani. A Montessori-pedagógia alkalmazásának elterjedéséhez nagymértékben hozzájárult Bardócz Pál 1924-ben megjelenő könyve a „Montessori nevelési rendszer és módszere” (Kövér, 1987, 67.). Ezt követte 1930-ban Montessori alapművének (*Mód-szerem*) magyar fordítása.

Alapirodalomnak számított a korszakban Bardócz Pál óvónőképzős hallgatók számára készült nevelési könyve „A magyar kisdednevelés vezérkönyve” is. Az 1928-ban megjelent munka szerzői között olyan meghatározó személyiségeket találunk, mint Imre Sándort, vagy Kacsóh Pongrácz zeneszerzőt, aki az ének-zenei fejezetet jegyzi (Kövér, 1987, 73.). Újnak számított több területen is a tankönyv, hiszen egyértelműen meghatározta az óvodai foglalkozások területeit (beszélgetés, mese, vers, ének-zene, játék, munka, testgyakorlás, rajz.), illetve Kacsóh Pongrácz által jelennek meg elsőként a gyermekhangszerek is az óvodai nevelésben (dob, facsörgő, triangulum stb.).

A magyar óvodai nevelés érdekes sajátossága azonban, hogy a reformpedagógia romantikus gyermekközpontúsága a század közepére fokozatosan háttérbe szorult az ugyancsak a reformpedagógiából származó másik meghatározó elv, a cselekvő-alkotó-aktív gyermekfelfogás mögött, amellyel „Az óvoda (...) a század folyamán egyre inkább iskola-előkészítő, tágabb értelemben: életre felkészítő szerepre tett szert.” (Pukánszky, 2005, 712.).

Az óvodapedagógusok képzésében 1926-ban történt alapvető változás: a kétéves időtartam négyévéssé bővült. A képzőbe való bejutás feltételévé vált a jó vagy jeles eredmény, a zenei és testi alkalmassági vizsgálat is. A negyedik év a gyakorlati képzést jelentette.

A népi-nemzeti irányzat megerősödése

1936-ban az óvodai nevelés felügyelete a vallási- és közoktatási miniszter kezéből a belügyminiszter hatáskörébe került. Ezzel párhuzamosan a szakpolitika felől egyre erősebb nacionalista szólamok határozták meg az intézmények működését. Ennek a német befolyás elől a szelet kifogó politikának létezett a '30-as évek végétől egy olyan sajátosan magyar alternatívája, mely a népi hagyományokat helyezte középpontba a nevelés-oktatás terén is. Az irányzat vezéregyéniségei voltak Imre Sándor, Kodály Zoltán, vagy a magyar népmese irodalmat az óvodai nevelésbe beemelő Illyés Gyula is (*Magyar parasztmesék, Székely népmesék* kiadványai). A korszak szimbolikus nyitánya Kodály „Zene az óvodában” c. előadása volt 1941-ben (Sztrinkóné, 2005, 138.). A népi-nemzetiek kísérlete a háborús években már nem vezethetett igazi sikerre, s a német megszállás (1944–45), vala-

mint az 1947-től megerősödő „szovjetizálás” olyan utakra vitte a magyar köznevelést, így az óvodákat is, amely hosszú évtizedekre meghatározta azok működését.

A szovjet-orosz óvodapedagógia két évtizede Magyarországon (1948-1971)

1948-ban Magyarországon az óvodákat és az iskolákat államosították. A következő évtizedekben az óvodák száma rohamosan nőtt. Ez egyrészt a növekvő gyermeklétszámnak volt köszönhető, másrészt a nők tömeges munkába állásának, amelyet a sztálinista erőltetett iparosítás és a „szocialista munkaerőkölcs” követelt meg. Harmadrészt a hatalom felismerte a családon keresztüli befolyásolás intézményi lehetőségeit, illetve a jövő „szocialista állampolgárainak” nevelésében az óvoda szerepét.

A pedagógia környezet azonban gyökeresen megváltozott a háború előttihez képest. 1949-től elkezdtek a szovjet mintát másolni, az lett a követelmény, s a megjelenő szakirodalmat „elárasztotta a szovjet pedagógia szelleme, illetve maga szovjet pedagógia.” „Az oktatásügy is a pártpolitika „szolgálóleányává” vált.” (Pukánszky B.-Németh A. 1996, 658.).

1950-ben fordították le és adták az akkori óvónők kezébe Bleher *Szervezett foglalkozások az óvodában* című művét, amely néhány évig meghatározta a magyar óvodák szemléletmódját is. Ez a könyv a szovjet óvónőknek készült, de fel sem merült, hogy a magyar viszonyok alkalmasak-e ennek az alkalmazására.

A kiadvány magyar bevezetője hosszútávra fogalmazta meg a magyar óvodák nevelési szellemiségét: „Az óvoda nemcsak általánosságban veti meg a sokoldalúan művelt szocialista ember nevelésének alapját, hanem az iskolai oktatásnak is tervszerű előkészítést ad.” (Bleher F. H., 1950, 3.). Ekkor kezdődött az óvoda közoktatási szerepének fundamentális lekötése az iskolára való felkészítésben, amivel majd a későbbi dokumentumokban sorra fogunk szembesülni.

A magyar óvodai nevelés történetében több szempontból is fontos az 1953-as év, mert törvény született az óvodai nevelésről, és a háború után újjászervezett óvodák számára megjelent egy a közös gondolkodást segítő országos dokumentum. A törvény hozzárendeli az óvodát az iskolához, mint ami előkészíti az iskolai oktatást. Ezzel az óvodák a korábbi helyzettől eltérően bekerülnek a közoktatás intézményei közé.

Az 1953-ban kiadott *Módszertani Levelek* ezzel a törvényi háttérrel megerősítve próbálja meg az óvodák egységes szemléletét kialakítani. A *Módszertani Levelek* alapvetően a kötelező foglalkozásokról szól, ami az oktatás, a gyermekek tanításának elsődleges eszköze volt. Ezek valójában az iskolai órarend óvodai adaptációi. Ezen keresztül valósultak meg a nevelés általános céljai, akárcsak egy iskolában.

Az általános témakörök mellett természetesen az aktuálpolitika óvodás elvárásaival is foglalkozniuk kellett az óvónőknek, ügyelve a fokozatosságra. Az óvoda első évfolyama kiscsoportosainak már fel kellett ismerniük képről Lenint, Sztálint, és Rákosit, de az akkor nagycsoportosoknak nevezett gyermekeknek már azt is tudniuk kellett, hogy kik is ők („ők a mi vezetőink, akik nagyon szeretik a gyerekeket”). A családdal kapcsolatos együttműködés fel sem merült a dokumentumot kiadó Oktatási Minisztérium illetékeseiben.

A magyar óvodapedagógiai gyakorlatot meghatározó következő dokumentum 1957-ben vezették be *Nevelőmunka az óvodában* címen, amellyel egy időben az 1953-s *Módszertani Levelek* hatályukat veszítették. Ez már könyv terjedelmű dokumentum volt és magán

viselte az elmúlt 10 év szocialista pedagógiai praxisának tapasztalatait és a hrucsovi idők enyhébb légkörét is. A miniszterhelyettesi bevezetőjében a magyar óvodát az 1828 óta folyó nevelés örökösének és a még friss szocialista pedagógia letéteményesének nevezi, rögzítve azt is, hogy az óvoda a köznevelési rendszer legalsó fokú intézménye.

Új és előremutató eleme volt a dokumentumnak, hogy elvárja az óvónőktől, hogy az óvodai élet „tudatos irányításakor” a 3-6 éves kor sajátosságait és szükségleteit vegye figyelembe, az eszközök és eljárások biztosítsák a gyerekek sokoldalú fejlődését. Ez azért volt haladó gondolat, mert az 1950-es években a lélektan Magyarországon a letűnt polgári társadalom maradványának számított, így ezek pontos definíció nem is léteztek.

Felismerték a „játék különös jelentőségét” is, igaz az oktatás szerepének erősítése érdekében. „Abban az óvodában, ahol sokat játszanak nyugodt, fegyelmezett légkör uralkodik, a kötelező foglalkozásokon a gyermekek figyelme hatványozottabb” (Nevelőmunka az óvodában, 1957, 58.). A kötelező foglalkozások tervezése viszont azokat a korábbi mechanizmusokat vitte tovább, amit 1953-ban vezettek be.

A Kézikönyv tanterv-jellegű volt: pontosan megadta azokat a tartalmakat, amelyeket minden óvodában fel kellett dolgozni. Precíz paramétereket határozott meg a kisgyermekek neveléséhez, amit csak „be kellett tartani” és nem kellett a nevelés kérdésein gondolkodni. Így inkarnálódott a herbarti pedagógia újra az iskola után az óvodában is Magyarországon az 1950-es évek vége felé. A *Nevelőmunka az óvodában* tanulásszemléletét továbbra is az iskolai tanulás sajátosságainak óvodai adaptációja hatotta át. Az óvodák iskolaelőkészítő funkciója megkérdőjelezhetetlen volt, évtizedekig meghatározva az óvodai nevelőmunka szemléletmódját.

Ez a szemléletmód ugyanakkor elősegítette az óvodapedagógia és az óvodapedagógusi pálya presztízsének erősödését, amely 1959-ben oda vezetett, hogy az óvodapedagógus képzés bekerült a felsőfokú képzések körébe, bár a főiskolai szint elérésére még bő két évtizedet (1985) várni kellett.

Az ideológiát felváltja a szakmaiság (1971-1989)

A magyar óvodapedagógiát szabályozó következő dokumentum, az *Óvodai Nevelés Programja* volt; célképzése egyértelmű változást mutat: „Az óvodai nevelés célja az egységes szocialista nevelési rendszer szerves része. Célja a 3-6 éves gyermekek sokoldalú, harmonikus fejlődésének elősegítése.” (Az Óvodai Nevelés Programja 1971, 9). A kötelező „szocialista” jelző ellenére, ebben a dokumentumban már a gyermek személyiségfejlődésén an a hangsúly.

Az összes előző dokumentumtól eltérő módon a családot is azonnal megemlíti a program a közös felelősség megfogalmazásával. Elfogadja ez a koncepció, hogy egyszerre neveli ugyanazt a gyermeket a két legnagyobb hatású színtér, a család és az óvoda, de a „hatások egysége érdekében az óvoda a kezdeményező” (Az Óvodai Nevelés Programja 1971, 11). „A játék az óvodáskorú gyermekek elsődleges tevékenysége.” (Az Óvodai Nevelés Programja 1971, 77.) – jelenti ki a program, amely a gyermekközpontú személet felé tett komoly szakmai lépésnek mondható. Fontos megemlíteni, hogy az új dokumentum a játékhoz szükséges egyéni és közös tapasztalatszerzést kiterjeszti az óvodán kívülre is, miközben növeli a gyermeki szabadságot és választhatóságot a játéktevékenységekben. Az életkori sajátosságok felismerésének nagyhatású következménye, hogy a kötelező foglalkozások mellett bevezettek egy új tanulási formát, a kötetlen foglalkozásokat, kez-

deményezések rendszerét. A '71-s program mérföldkőnek tekinthető, mert megnyitotta az utat egy olyan szakmai gondolkodás előtt, amelyben egyre inkább a gyermek áll a középpontban.

A rendszerváltástól napjainkig (1989-2014)

A rendszerváltás utáni társadalmi, politikai és szellemi változások az óvodákat sem hagyták érintetlenül. A következő 1989-ben megjelent dokumentum nem új óvodai program, mert az előző továbbfejlesztésének eredményeképpen jött létre.

A *továbbfejlesztett Óvodai Program* (TONEP), a megelőző dokumentumokhoz hasonlóan még mindig egy egységes, minden óvoda, óvodapedagógus számára kötelező programot jelentett. A hangsúlyok azonban már módosultak, amelyet a célképzésben is érzékelhetünk: „Célja, a 3-6-7 éves gyerekek, sokoldalú, harmonikus fejlődésének elősegítése.” Ebben azonnal látható az életkor kiszélesítése, mert időközben (1985-86) törvényesítették a „rugalmas beiskolázást,” a fejlettségnek és nem az életkornak megfelelő iskolakezdést.

Az új programban tovább folytatódik az ONEP-pel kezdett lélektani nyitás is: az óvodától már kifejezetten elvárja a program, hogy „érzelmi (emocionális) burokrendszert” biztosítson a gyermek fejlődéséhez. Az általános elvek között elsőkségre került a gyermeki méltóság, a gyermek jogainak biztosítása. Ez utóbbi annyira új volt, hogy még tartalmi elemei sem voltak akkor. A pedagógiai etika is ekkor fogalmazódik meg először, ami a szülőktől származó információk kezelésére utalt.

A program három fő tevékenységet különböztet meg egymástól: játék, munka, tanulás.

Az abszolút prioritást azonban már a játék kapja: „Közülük elsődleges és alapvető a játék, mely áthatja a gyerek egyéb tevékenységeit.” (Az Óvodai Nevelés Programja 1989, 9.) A tanulást sem csak szűken egy tevékenységre használja, mert a lélektan tágra értelmezett fogalmát is használja, a gyermeki fejlődés, a változás lényegének leírásához. „A gyerekek fejlődésében beálló változás a tágra értelmezett érés és tanulás együttes következménye.” Tehát most már nemcsak a szervezett tanulásra kell gondolni, mint a korábbi évtizedekben, amit oktatásnak neveztek, hanem a spontán helyzetek tanulási értékét is felfedezik.

A program az óvoda-iskola viszonyban nem csak a defenzívát tartja elfogadhatónak az iskolai doktrínával szemben, hanem a szemlélet elfogadtatása érdekében a proaktív megoldásokat is. „Az óvoda nevelőtestülete tekintse feladatának, hogy ha szükséges, az óvoda gyerek szempontú szemlélete, módszerei hassanak a gyerekeket befogadó iskolára.” (Az Óvodai Nevelés Programja 1989, 279.) Noha ez utópisztikus elképzelésnek bizonyult, a program lerakta a gyermekközpontú, és a gyermekek öntevékenységre épülő, demokratikus szellemiségű nevelőintézmény alapjait.

Az 1990-es évek Magyarországon az új társadalmi- politikai intézményrendszer törvényi megalapozása folyik, így a teljes közoktatás számára is új törvény volt szükséges: olyan törvény, amely legitimálja a többféle fenntartású közoktatási intézmények működését, kiiktatja az állami intézményekből az ideológiai egyoldalúságot és lehetővé teszi a pedagógiai pluralizmust. 1993-ban fogadta el az országgyűlés a LXXIX. törvényt a közoktatásról, benne az óvodai nevelést meghatározó néhány passzust.

Az óvoda továbbra is hároméves kortól fogadta a gyerekeket, de az európai tankötelezettségi kort átvéve, ötéves kortól minden gyermek számára kötelező az óvodai nevelés. „Az óvoda az Országos óvodai nevelési program alapján pedagógiai programot készít, vagy az Országos óvodai nevelési program alapján készített pedagógiai programokat tesz közzé.” (A közoktatásról szóló a Magyar Köztársaság 1993. évi LXXIX. törvénye.45.§ 1.) Ez gyökeres változást jelentett az összes óvoda számára, mert a helyi, saját nevelési program írásán semmilyen előzménye nem volt a magyar óvodák múltjában, s komoly felelősséget, szakmai erőfeszítést igényelt minden óvodapedagógustól. A két szintű tartalmi szabályozás bevezetése a lokalitás elvének érvényesülését biztosította.

Az óvodák számára egy új, központi program kidolgozása is elkezdődött, amely szellemiségében és funkciójában is szakított az államszocialista éra szabályozó dokumentumaival. Az *Óvodai Nevelés Országos Alapprogramja* (röviden Alapprogram) 1996-ban született meg, amely egy rövid, általános érvényű elveket, szempontokat meghatározó keretjellegű dokumentum. Fontos jellegzetessége, hogy nincs „tantervi” tartalomelvárás benne, tehát valóban az óvodák állíthatták össze az adott intézményben fontosnak vélt tevékenységek tartalmát a saját, helyi nevelési programjukban. Inkluzív szemléletmódjában az Alapprogram a gyermeki sajátosságokat tudatosan előtérbe állító, a korábbi szegregáló megoldásokat (pl. életkori elkülönítés elvárása) mellőző, erősen *gyermekközpontú, játék-meghatározottságú és tevékenységalapú* dokumentum.

Az Alapprogram, gyermekközpontúságából és keretjellegéből fakadóan megengedte a különböző reformpedagógiai és alternatív pedagógiai koncepciók elterjedését: Magyarországon a Freinet-, Waldorf-, Montessori-féle pedagógiák vannak elsősorban jelen, illetve az amerikai eredetű Lépésről Lépésre Program. Mindemellett sok óvoda készített modern szemléletű saját programot is, vagy vett át, esetleg adaptált meglévő óvodai programot. Az átvétel folyamatát segíti, hogy létezik az ún. „minősített (szakmailag ellenőrzött, szűrt) programok” csoportja, amelyekhez az óvodák szabadon hozzájuthatnak.

2011-ben új köznevelési törvény született Magyarországon, amely több szempontból újra tervezte az intézmények irányítását, ellenőrzését, a pedagógusok értékelését. Ez a törvény ugyanakkor nem érinti az óvodák szakmai és fenntartói pluralizmusát. Új elemként azonban megjelenik a törvényben a 3 éves kortól történő kötelező óvodáztatás, amelytől a kormányzat az esélyegyenlőség javulását, az Alapprogramban is megfogalmazott hátránycsökkentést várja. Mindez azonban aligha befolyásolja a magyar óvodai nevelés modern, gyermekközpontú, szakmailag sokszínű voltát. Az új törvény után 2012-ben finomított Alapprogram tartósította a (oktatási intézmények számára is mintául szolgáló) gyermekközpontú szemléletet. Így jövőbe mutató kihívások az óvodapedagógia számára a nélkülözhetetlen játék túlsúlyának kialakítása, és a gyermeki szükségletekre épülő tevékenységpedagógia realizálása.

IRODALOM

- Az óvodai nevelés programja (1971) Tankönyvkiadó, Budapest.
- Az óvodai nevelés programja (1989) Országos Pedagógiai Intézet, Budapest.
- BARDÓCZ P. (1928) *A magyar kisdiednevelés vezérkönyve*. Budapest: Székesfőváros Házinyomdája.
- BLEHER, F. H. (1950): Szervezett foglalkozások az óvodában. Közoktatásügyi Kiadováállalat, Budapest.
- GOLNHOFER E. (2004) *Hazai pedagógiai nézetek 1945 – 1949*. Pécs: Iskolakultúra.
- KÖVÉR S. (1987) *Az értelmi nevelés története óvodáinkban 1828-1975 között*, Budapest: Közgazdasági és Jogi Könyvkiadó.
- KÖVÉR S. (1996): *Óvoképzésünk múltjának néhány mának szóló tanulsága*. In: Varga Gyula (szerk): „...Amit a tükör önmaga terem” HWIPF, Hajdúböszörmény.4-16.
- Módszertani levelek – óvodai foglalkozások (1953) Tankönyvkiadó, Budapest.
- Nevelőmunka az óvodában: útmutatás óvónók számára (1957) Tankönyvkiadó, Budapest.
- PÁLFI S. (2010): *A magyar óvodai projektpedagógia elméleti alapjai*. Debrecen University Press, Debrecen pp.48-95.
- PUKÁNSZKY B.& NÉMETH A. (1996): *Neveléstörténet*. Nemzeti Tankönyvkiadó, Budapest.
- PUKÁNSZKY B. (2005): *A gyermekről alkotott kép változásai az óvoda történetében*. *Educatio*, 4: 700-711.
- SZTRINKÓNÉ N. I. (2009): *Óvó-iskola, gyermekkert, óvoda: a kisgyermekkor neveléstörténete*. Debrecen: Didakt.
- VÁG O. (1979): *Óvoda és óvodapedagógia*. Tankönyvkiadó, Budapest.
1993. évi LXXIX. törvény a közoktatásról. 137/1996. (VIII. 28.) Kormányrendelet az Óvodai Nevelés Országos Alapprogramjának kiadásáról.
- A nemzeti köznevelésről szóló 2011. évi CXC. törvény
1891. XV. Kisdiedovási törvény www.1000ev.hu/index.php?a=3¶m=6422
- (Utolsó látogatás időpontja: 2014. május 05.)

SZEMLE

EGY FONTOS KÖNYV A STATISZTIKAI SZEMLÉLET TÖRTÉNETÉRŐL

A mai világban élők megszokták, hogy az újságokban és az elektronikus médiában folyton-folyvást megjelennek adatok a gazdaság teljesítőképességéről, a népességszám változásairól, vagy újabban a tanulók tudásszintjének alakulásáról stb. Olyannyira így van ez, hogy az emberek többnyire természetesnek, magától értetődőnek tekintik azt, hogy a gazdasági vagy társadalmi élet ilyen-olyan aspektusait számszerűsített formában jelenítik meg. Ha egyáltalán kérdéseik vetődnek fel bennük a számszerűsítéssel kapcsolatban, leginkább a közölt adatok hitelességét kérdőjelezzik meg, gyanúperrel élve azzal kapcsolatban, hogy vajon az adatokat előállító szervezetek, pl. a statisztikai hivatalok, nem hamisítják-e meg az inflációval vagy éppen a munkanélküliséggel kapcsolatos adatokat annak érdekében, hogy jobb színben tünthessen fel az éppen regnáló kormány. Az efféle kérdésfeltevésben már benne rejlik egyfajta társadalmi tudás vagy tapasztalat azt illetően, hogy az adatokat létrehozók érdekeltségi viszonyai torzításokat okozhatnak (például a statisztikai intézményeknek az állami szervektől való függősége okán, vagy az egyes közvélemény-kutató szervezeteknek a politikai pártokhoz fűződő kapcsolata miatt).

Ez utóbbi beállítódásokhoz képest merőben újszerű szemléletet képviselt Alain Desrosières, aki hazáján túl számos más országban is nagy ismertségre tett szert *A nagy számok politikája – A statisztikai gondolkodásmód története* című francia nyelvű könyvével, illetve e mű más világnyelveken (pl. angolul, németül, spanyolul) megjelent változataival. A 2013-ban elhunyt szerző pályája során egy-

felől mindvégig a sok tekintetben a KSH-hoz hasonlítható francia intézménynek, a *Statistiques et Économiques* (INSÉE – Institut National des Études Statistiques et Économiques) az alkalmazásában állt, másfelől pedig másodállásban vezető franciaországi társadalomtudományi kutatóhelyeken is dolgozott.

Statisztika-történeti feldolgozások korábban is léteztek, ám ezekhez képest merőben új, konstruktivista szemlélet jelenik meg Desrosières említett munkájában és számos egyéb művében. A szerző a statisztikai adatokat magától értetődő, objektív létezőként kezeli, szemléletével szakítva munkáiban abból az egyszerű, ám mély értelmű „alapállásból” indul ki, hogy „nem tekinthetjük adotttnak az adatok létezését.”¹

A könyv több évszázadon átívelő áttekintést ad arról, hogy miként jelentek meg, bontakoztak ki és alakultak át a statisztikai tevékenységek négy országban, nevezetesen Németországban, Franciaországban, Nagy-Britanniában és az Egyesült Államokban. A *nagy számok politikája* nem eredeti források tömkelegét mozgósít

¹ Avec Alain Desrosières entre pénombriens. *Supplément à la Lettre d'information de Pénombre (La lettre grise)*, 2013, Vol. 9, Nr. 12 (7. o.) http://www.pénombre.org/IMG/File/Lettre_grise_12.pdf/.

E kijelentés eredeti változata pompás francia szójáték: „Les données ne sont pas données.” Az imént megadott magyar jelentésen túl ez azt is jelenti, hogy „költéses dolog adatokhoz jutni.” Mindkét francia jelentés figyelmet érdemel, lévén, hogy Desrosières munkái beilleszkednek a statisztikai tevékenységek szociológiájának megalkotását célzó tűző programba, melynek gyökerei – mint ezt a recenzió egy későbbi pontján érzékeltetni fogom – Pierre Bourdieu munkásságára vezethetők vissza.

tó, a történeti folyamatokat aprólékosan bemutató munka, hanem inkább lényeges összefüggéseket felmutató könyv, amely sikeresen épít az egyes nemzeti mintákkal kapcsolatos eddigi kutatások eredményeire. Miként a mű címe is jelzi, a szerző elsősorban a statisztikai tevékenységek mögött meghúzódó szemléletmódokra helyezi a hangsúlyt, azt vizsgálva, hogy milyen kontextusban jelentek meg a statisztika című alatt számot tartott első munkák, majd utóbb hogyan változott a statisztikát művelők szemlélete, hogyan alakultak ki új statisztikai megközelítések, eljárások.

Bár a birtokolt javakról, jószágokról számot adó összeírások, leltárak az ókori birodalmak vagy egyes középkori királyságok, kereskedőtársaságok, szerzetesrendek stb. égisze alatt is készültek, a statisztikai tevékenységek alapító atyáiként XVII–XVIII. századi német fejedelemségek egyes egyetemi professzorait szokták megjelölni, akik arra törekedtek, hogy hallgatóik, illetve a fejedelmek és ezek tisztviselői számára adjanak módszeres áttekintést az egyes államok jellemzőiről. Érdeklődésük államközpontságára már abban is megnyilvánult, hogy tevékenységük megnevezésére használt kifejezésben az állapotot jelentő latin *status* szóból képzett *stato* (állam), illetve *statista* (államférfi) olasz szavak bujkálnak. Az első „statisticusok” azonos szempontokat mozgósítva igyekeztek bemutatni az országokat – eleinte megannyi irodalmias megoldással is élve –, miközben az országbemutatókhoz a skolasztikus filozófusok egyes Arisztotelésztől átvett, az okok és okozatok közötti kapcsolatot előtérbe állító kategóriái szolgáltak vezérfonalul.

Paul Lazarsfeld egy Magyarországon még nem publikált tanulmányára² támaszkodva

² Lazarsfeld, Paul (1961): Notes on the history of quantification in sociology: Trends, sources and problems. *Isis*, Vol. 52. pp. 277–333. Ezt a szöveget Desrosières a tanulmány francia fordítására hivatkozva idézi, amely a Lazarsfeld írásait közlő „A társadalomtudományok filozófiája” (*Philosophie des sciences sociales*) című kötetben jelent meg a Gallimard kiadónál 1970-ben. 73–162. o.

Desrosières jelzi munkájában, hogy a statisztikai szemlélet megjelenésére kedvezően hatott a harmincéves háború utáni helyzet, mikor is „közel 300 szegény és egymással rivalizáló államocskából állt a Német-római Birodalom, ahol a jogokkal és kötelezettségekkel kapcsolatos kérdések újradefiniálása meghatározó jelentőségű volt. (...) Egy ilyen helyzet tekintéllyel és presztízzsel ruházta fel azokat, akik nem új dolgok létrehozására, hanem a szisztematikus katalogizálásra voltak hajlamosak (...)”

A göttingeni statisztikai iskola XIX. század elején élt követői a földrajzi adottságokkal, természeti kincsekkel és a népesség jellemzőivel kapcsolatos ismeretek bemutatásakor előszeretettel folyamodtak a táblázatos bemutatáshoz és egyúttal a jelenségek számszerűsítéséhez. A mennyiségi összefüggéseket bemutató kereszttáblák megjelenése kapcsán Desrosières egyfelől arra hívja fel a figyelmet, hogy „létrejött egy olyan nézőpont, melynek révén egyidejűleg és azonos fogalmi rács révén lehet egymügyre venni különböző országokat.” (33. o.) Másfelől azt is kiemeli, hogy a számszerűsítés egyúttal elvonatkoztatással is jár, azaz megannyi egyébként fontos részlet, összefüggés háttérbe szorításával.³

Az elvonatkoztatás, absztrakció mozzanatának kiemelése szervesen következik Desrosières művének alapkoncepciójából, nevezetesen abból, hogy munkája során a szerző a konstruktivista nézőpont következetes alkalmazására tesz kísérletet. Vagyis annak elemzésére vállalkozott, hogy az elmúlt két évszáz

³ Ennek kapcsán a szerző Paul Lazarsfeld említett tanulmányának francia kiadására támaszkodva azt is jelzi, hogy az elvonatkoztatás, általánosítás veszélyeivel a számszerűsítést ellenzők korai statisztikusok is tisztában voltak: „Ellenfelek különbséget tettek a »kifinomult, disztigvált« és a »vulgáris« statisztika között. Ez utóbbi szerintük: »[...] a magas szintű művészetet ostoba penzumma degradálta. Ezek a szegény ostobák azt a badar gondolatot terjesztik, hogy egy állam hatalmának megismeréséhez elegendő megtudnunk mekkora a területe, a lakossága, a nemzeti jövedelme és az ott legelésző államok száma.«” (99. o.)

zadban miként hoztak létre, fogadtak el, majd kérdőjeleztek meg újabb és újabb olyan adatokat, fogalmakat, elemzési módszereket, melyek alkalmasnak tűnhettek egy-egy állam, netán több társadalom állapotainak, dinamizmusainak tanulmányozására.

A Desrosières munkáját meghatározó szemlélet kialakításához Pierre Bourdieu gondolatvilágának egyik idehaza kevéssé ismert ága, a társadalmi kategóriák társadalmi megformáltságát előtérbe állító megközelítés adta az első és talán legfontosabb ösztönzést.

Itt érdemes röviden kitérni arra, hogy a társadalmi fogalmak társadalmi megformáltságát előtérbe állító vizsgálódások kialakulásában fontos szerepet játszottak az eredetileg filozófusnak készült Bourdieu által jól ismert francia ismeretelméleti kutatások,⁴ melyek egyebek között Koyré, Bachelard és Canguilhem nevéhez köthetők, akik különböző természettudományos területeken bekövetkezett változásokat elemezve visszatérően a tudósok fejében lévő megközelítések, fogalmi keretek fontosságát, illetve a vizsgálódásokat vezérlő szellemi konstrukciók esetenkénti radikális átalakulásának jelentőségét emelték ki.

Nemrégiben egy kutató, Jean-Francois Braunstein a francia ismeretelméleti vizsgálódások egyik fő jellemzőjeként emelte ki a történeti megközelítés, illetve a megszakítások (Louis Althusser kedvenc szavával élve: cezúrák) előtérbe állítását. Mint írta: „a folyamatossgot hangsúlyozó történetiséggel szemben a tudománytörténetet francia módra művelők

a tudománytörténetet a diszkontinuitás, illetve a »szakítások« történeteként fogják fel.”⁵ (Itt érdemes megemlíteni a tudományos forradalmak szerkezetét elemző, a paradigmaváltások fogalmát előtérbe állító világhírű amerikai társadalomtörténész, Thomas Kuhn fő művének 1. számú lábjegyzetében a munkásságára különösen nagy hatást gyakorolt szerzők között első helyen említette a francia tudománytörténészek fenti felsorolás elején említett Koyré Galilei-tanulmányait.)⁶

Bourdieu konstruktivista gondolataival még akkor ismerkedett meg Desrosières, amikor egy statisztikusokat és gazdaságelemzőket képző francia elitképző intézményben (École Nationale Supérieure de Statistique et d'Économie Appliquée) hallgatta az Algériából nemrégiben hazatért fiatal szociológus társadalmi tagoltsággal kapcsolatos kurzusát: „Bourdieu akkoriban hívta fel a figyelmünket arra a gondolatra, hogy «társadalomtörténeti szemszögből kell szemügyre venni a statisztikákat és ezen belül is a foglalkozási statisztika nomenklatúra-rendszerét. (...)» Másként mondván, beavatott minket abba a gondolatba, hogy nem előzmények nélkül hullanak ölünkbe a nomenklatúrák, hanem történetük van, és hogy országokként különböznek egymástól.”⁷

Visszakanyarodva a Desrosières által nyújtott elemzéshez, érdemes felvillantani, hogy a statisztikai tevékenységek története tekintetében az egyes országok tradíciói, fejlődési útjai között milyen különbségek rajzolódnak ki. A szerző szerint a Napóleontól elszenvedett vereség nyomán újjászerveződő német államokban, mindenekelőtt Poroszországban a XIX. század első felében kiépített statisztikai hivatalok mű-

⁴ A *Le métier de sociologue, Préables épistémologiques* című, Pierre Bourdieu, Jean-Claude Chamboredon és Jean-Claude Passeron által a hatvanas években összeállított, részint a szerzőhármas által írt eredeti szövegekből, részint nagyszámú szemelvényből álló tankönyvből jól kitérni, hogy Bourdieu és közvetlen munkatársai milyen nagy jelentőséget tulajdonítottak a francia tudománytörténeti, illetve ismeretelméleti iskolának, hiszen a szövegválogatásba beemelt francia szerzők között Durkheim neve mellett Bachelard és Canguilhem nevével találkozhatunk a leggyakrabban.

⁵ Braunstein, Jean-Francois (2002): *Bachelard, Canguilhem, Foucault. Le „style français” en épistémologie*. In: Wagner, P. (ed.): *Les philosophes et la science*. Paris: Gallimard Folio. 940. o.

⁶ Thomas Kuhn: *A tudományos forradalmak szerkezete*, Gondolat, Budapest, 1984 (8. és 277. o.)

⁷ Bouchard, Julie (2008): *Communiquer, gouverner, et être gouvernés par les chiffres. Questions à Jacques Durand et Alain Desrosières. Médiation et communication*, No. 28. 9. o.

kódésében „háromféle hagyomány ötvöződött: a politikai, történeti és földrajzi leírásokat alkotó egyetemi oktatók által képviselt tradíció, a hivatalnokok adminisztratív nyilvántartásainak hagyománya és az amatőr tudósok gyűjtőmunkája révén előállt táblázatos adatgyűjtemények tradíciója.” (220. o.)

Desrosières rávilágít arra is, hogy a franciaországi és az angliai statisztikai gondolkodás korai korszakára a német államokétól merőben eltérő, ugyanakkor egy szempontból felettébb hasonló út volt a jellemző: a valószínűség-számításon alapuló szemléletmód korai térnyerése. Már a XVIII. század végén kialakult annak gyakorlata, hogy a francia abszolutisztikus monarchia különböző hivatalai – elsősorban gyakorlatias célok jegyében – különféle adatgyűjtéseket szerveztek, például az állami adószedés szempontjából meghatározó fontosságú volt, hogy minél pontosabb népesség-nyilvántartások szülessenek. Utóbb fontossá vált annak feltárása is, miként befolyásolják a járványok vagy éhínségek az alattvalók létfeltételeit. Az így született leírásokra jellemző volt, hogy „függetlenül attól, hogy az uralkodónak vagy az általa irányított közigazgatási apparátusnak címezték őket, királyi privilégiumként kezelt titkos dokumentumok voltak. (...) Ugyanakkor az 1750-es évektől a mind szabadabban szerveződő közvélemény kialakulása elősegítette a kormányzattól független tudásformák létrejöttét. (...) A közigazgatás egésze alatt zajló vizsgálatokat övező titok annak kedvezett, hogy (...) a tudósok részleges információkra támaszkodva, becsléseken alapuló munkákat alkossanak, melyek minták alkalmazásán és számítások elvégzésén alapulnak...” (41. o.)

Míg a mintavételen, illetve becsléseken alapuló módszerek előtérbe kerülése mögött Franciaország esetében az erősen központosított állam nyilvántartásainak titkossága húzódott meg, addig Angliában „a közvetett módszerekhez való folyamodás (...) a liberális államfelfogásra és az állami hatáskörök korlátozására vezethető vissza, ami megakadályozza olyan nagy adatgyűjtések megszervezését, amelyekre

a kontinentális országok, pl. Franciaország már sort kerítettek. Így például 1753-ban egy népszámlás megtartását élesen ellenezte a Whig párt arra hivatkozva, hogy ez «az angol nép még meglévő utolsó szabadságjogainak felszámolásához vezetne.»” (36. o.)

Desrosières könyve jól érzékelteti, hogy a becsléses, illetve mintavételes eljárások előtérbe kerülése nyomán alakult ki a XIX. század végi Angliában az a Galton és Pearson nevéhez köthető, többek között a regresszió- és korrelációs számítással fémjelzett módszertani arzenál, amelyet napjainkban matematikai statisztika névvel illetnek. Az ismertelméleti megközelítésnek nagy jelentőséget tulajdonító Desrosières kiemeli, hogy Pearson a XX. század elején működött osztrák fizikus és filozófus, Mach szemléletmódját átvéve tudatosan szakított az ok-okozatiságot előtérbe állító korábbi megközelítésekkel: „Az okság metafizikus fogalma helyére lép a korreláció, amely a megfigyelt jelenségek közötti, várhatóan a jövőben is reprodukálódó és a korreláció szóban forgó fogalmával mérhető kölcsönös összefüggéseket jelöli.” (134–135. o.)

Helyhiány miatt csak futólag jelezném, hogy Desrosières könyvében még számos igen érdekes, a tudományszociológiai munkákban externalista magyarázatokként számon tartott ilyesfajta összefüggésről olvashatunk. Egyetlen további példaként említem meg az Egyesült Államok esetét, ahol ellentétben számos európai országgal, amelyekben még a XIX. század során kiépültek jelentős statisztikai apparátusok, egészen a XX. század elejéig kellett várni erre a fejleményre. Ez utóbbi kapcsán Desrosières munkájából az is kiderül, hogy a statisztikai adatgyűjtés és -feldolgozás intézményesülését, majd az intézményrendszer későbbi átalakulását nem csupán a tágabb társadalmi kontextus befolyásolhatja, hanem egyes konkrét személyek, illetve csoportok szerepvállalása is. Így például Francis Walker, az 1870-es és 1880-as amerikai népszámlálások szuperintendása az adatgyűjtés megbízhatóságát növelendő erőteljes lobbitevékenységet végzett annak ér-

dekében, hogy több személyt alkalmazzanak az egyes censzusok lebonyolításához. Ugyanő 1888-ban már azért harcolt, hogy a népszámlálások megszervezése a nem sokkal korábban felállított, jelentős apparátusra támaszkodó Munkaügyi Hivatalhoz kerüljön. Erőfeszítései azonban hiábavalóknak bizonyultak, mivel az akkoriban megtartott választás győztese inkább abban voltak érdekeltek, hogy saját híveik kapjanak megbízást a következő népszámlálás megszervezésére. A statisztikai tevékenységek az Egyesült Államokban csak a XX. század elején intézményesülnek, amikor is a Kongresszus megszavazta egy *Census Bureau* nével illetett állandó hivatal felállítását, amely – nevével ellentétben – már nem csak a népszámlálások megszervezésével foglalkozott, hanem többek között a gazdasági területeken is hozzálatott az adatgyűjtéshez. E fejlemény szorosán összefüggött az állam gazdaságszervező, illetve -szabályozó szerepének növekedésével, ami fontossá tette az árakkal és az üzleti tevékenységekkel kapcsolatos információk összegyűjtését. A XX. századi amerikai fejleményekről adott áttekintésből is kiderül, hogy az USA-béli statisztikai tevékenységek bővülésében fontos szerepet játszottak az egyetemi berkekben működő kutatók, statisztikával foglalkozó szakemberek, illetve az ezeket tömörítő olyan szakmai egyesületek, mint az *American Statistical Association*, illetve az *American Economic Association*.

Itt jegyezném meg, hogy igen jó lenne, ha Magyarországot illetően is képet tudnánk alkotni arról, hogy egyes korszakokban egyes tudósok, lobbicsoportok milyen szerepet játszottak az állam égisze alatt végzett adatgyűjtések megszervezésében. Feltétlenül említést érdemel Schwarcz Gyula, aki sokoldalú tudományos és politikai tevékenysége mellett a dualizmus létrejötté előtt és után visszatérően szorgalmazta a hazai oktatásügyi statisztikai adatgyűjtés elindítását.

Mint recenzióm elején már jeleztem, Desrosières munkájában meghatározó szerepet játszik a konstruktivista szemlélet. E meg-

közelítés jól tetten érhető a könyv azon részeiben, amelyek a vizsgált országok példája alapján érzékeltetik, hogy az egyes bevettnek számító, illetve magától értetődőnek tartott statisztikai kategóriák, klasszifikációs rendszerek valójában történelmi képződmények. A napjainkban is használt franciaországi osztályozási rendszer gyökerei Desrosières elemzése szerint a Nagy Francia Forradalom körüli időkre vezethetők vissza: „E nomenklatúra hosszú története három szakaszban foglalható össze, amelyek mindegyike magyarázza e rendszer mai vonásait. Az első időszak a régi értelemben vett szakmák nyomjára béklyűket. A második időszakban, az 1850-es évektől kezdve érzékelhető a bérből élők és a nem bérből élők közötti megkülönböztetés fontossá válása. A harmadik, az 1930-as évektől kezdődő időszakban kerül előtérbe a bérből élők csoportján belüli hierarchikus tagoltság, amelyet a kollektív szerződések rögzítettek a képzettségi rendszert figyelembe véve.” (323–324. o.)

Itt érdemes megemlíteni, hogy a harmadik periódushoz köthető *cadres* fogalmának kialakulása, amely a vállalati közép- és felsővezetőket – más országok statisztikai gyakorlatától eltérően – egy csoportként kezeli. Ez utóbbi kategória létrejöttének, elterjedésének, illetve elfogadottá válásának társadalomtörténetét egyébként egy másik Bourdieu-tanítvány, Luc Boltanski írta meg egy nagyhatású könyvben.⁸ Művéből kiderül, hogy a fogalom a második világháború előtti időkben jelent meg, amikor a munkásszakszervezetek szervezeti modelljét átvevő mozgalmak a vállalati közép- és felsővezetők összefogását tűzték célul. Nem lényegtelen mozzanat, hogy Boltanski könyvének megírásakor egyebek között Desrosières egy korai tanulmányára támaszkodott.⁹ Ez az írás egy

8 Boltanski, Luc (1982): *Les cadres*. Paris: Les Editions de Minuit.

9 Desrosières, Alain (1977): *Éléments pour l'histoire des nomenclatures socioprofessionnelles*. In: (INSEE): *Pour une histoire de la statistique*, tome 1, *Contributions*. Paris: INSEE-Economica. 155–231. o.

olyan kötetben jelent meg, amely egy, a francia statisztikai-történeti kutatásoknak döntő lökést adó 1976-os konferencia előadásait adta közre. (A rendezvény létrejöttét Desrosières és másik, szintén Bourdieu-nél tanult statisztikus kezdeményezte.)

Egy mégoly terjedelmes recenzió kerekéi között is lehetetlenség lenne bemutatni Desrosières statisztikai-történeti vizsgálódásainak minden főbb mozzanatát. Ismertetésem inkább annak jelzésével fejezem be – mintegy kedvcsináló jelleggel –, hogy a szerző más neves kortárs francia társadalomkutatók (Bruno Latour, Laurent Thévenot) munkáira is támaszkodva kutatásai során arra a meggyőződésre jutott, hogy a statisztikai fogalmak, illetve adatok olyan gyakorlatias célok megoldására is használt, viszonylagos elfogadottsággal rendelkező konvenciók, melyek konstruált jellegével mindenképp érdemes tisztában lenni. A konstruált jellegből ugyanakkor nem következik, hogy a statisztikai adatgyűjtésekben, elemzésekben foglaltak teljesen fiktívek lennének és semmilyen köszönő viszonyban ne lennének a leírások, elemzések tárgyául szolgáló különféle társadalmi jelenségekkel, folyamatokkal.

(Alain Desrosières: *La politique des grands nombres. Histoire de la raison statistique. La Découverte*, 2000 [1993], Paris. 460 p. /Második, a szerző eredeti utószavával bővített kiadás./)

Bajomi Iván

ADAT, MÉRÉS, FEJLESZTÉS – NEMZETKÖZI PÉLDÁK AZ OKTATÁSI ADATOK FELHASZNÁLÁSÁRA

A világ minden táján számos kérdéssel kerülnek szembe a pedagógusok nap mint nap: hogyan javítsunk bizonyos tanulók eredményein, milyen oktatási eszközöket válasszunk a hatékonyság növelése érdekében vagy épp milyen jellegű szakmai továbbképzésen vegyünk részt. Megfelelően átgondolt döntésekre azonban

gyakran túl kevés idő jut az iskolai mindennapok forgatagában, sokszor nem állnak rendelkezésre a szükséges adatok, vagy az információkat nem tudják megfelelően értelmezni az iskolai dolgozók; ráadásul a pedagógusok egy része saját szakmai tapasztalataira hivatkozva mellőzi az adatok használatát. Pedig az oktatással kapcsolatos tények felhasználása nem csupán az elszámoltathatóságot segíti elő, de hozzáértő kezében az oktatás minőségének látványos javulását is, már az individuális szinttől kezdődően.

Ma, az információrobbanás korában, mindenhol adatok vesznek körül bennünket, így az oktatásban is. Az oktatási adatokra alapozott döntéshozatal egyre fokozottabb elvárás a kormányzatok részéről. Az oktatásért felelős központok úgy vélik, ha a tanárok elemzik a teljesítményinformációkat, akkor ennek fényében képesek lesznek hatékonyabban tanítani. Az egyes nemzetek különféle oktatási kultúrái egymástól eltérő, ám a saját rendszerüknek leginkább megfelelő megoldásokat dolgoztak ki; ezeket mutatja be a Kim Schildkamp, Mei Kuin Lai és Lorna Earl által szerkesztett, *Data-based Decision Making in Education* című tanulmánykötet. A szerzők hét ország – Új-Zéland, Hollandia, Anglia, Dél-Afrika Köztársaság, Belgium, Egyesült Államok és Kanada – adatfelhasználással kapcsolatos módszereit és nehézségeit ismertetik különböző aspektusokból.

A kötet első részében a szerzők több fejezetet szentelnek a fogalmak tisztázására, illetve egy elméleti modell bemutatására. Ismertetik a legfontosabb fogalmakat, azok típusait és alrendszeit, és az adathasználat folyamatait. Ezekből a hasznos részekből megtudhatja az olvasó például azt, hogy oktatási adatnak nevezünk „minden olyan szisztematikusan gyűjtött információt a tanulókról, iskolákról, iskolavezetőkről és pedagógusokról, melyekhez kvalitatív (pl. osztálytermi megfigyelés) és kvantitatív (pl. értékelési eredmények) módszereken keresztül jutunk” (177. o.). Az adatvezérelt döntéshozatal (data-based decision making) mindig az iskoláztatás, az iskolarendszer javítását szolgálja. Azonban nem minden adathaszná-

lat szükséges és megfelelő: a szerzők megkülönböztetik az adatok kívánatos (konceptuális és instrumentális adathasználat) és kevésbé kívánatos (helytelen felhasználás, visszaélés) felhasználási halmazát.

Felmerülhet a kérdés, hogyan vezet az oktatási adatok használata a tanulási teljesítmények növekedéséhez? Az adatok segítségével az iskolavezetők és a pedagógusok megfelelő tanulási célokat képesek felállítani diákjaik számára; monitorozni és ellenőrizni tudják, hogy a diák elérte-e a kitűzött tanulási célt; továbbá segíti a diákok is önmaga tanulásának ellenőrzésében. Amennyiben a tanuló mégsem érné el a kitűzött tanulási célt, a pedagógus további döntéseket hozhat a szükséges változtatások érdekében. Az oktatási adatokat ezáltal ott vizsgálják, ahol az iskola és a pedagógus folyamatosan monitorozza oktatási módszerének tanulói teljesítményre gyakorolt hatását, és szükség esetén változtat addigi tevékenységén. Ez azonban eltérő módokon jelenik meg a gyakorlatban.

A bemutatott példákban közös vonásként jelenik meg az adatok értelmezésének nehézsége. Minden esetben elmondható, hogy az egyes országok nagyméretű adatbázissal rendelkeznek a tanulók családi hátteréről, iskolai előmeneteléről, kompetenciáiról és lexikális tudásáról, viszont mindenhol gondot jelent az adatok összekapcsolása a pedagógiai gyakorlattal. Kivétel nélkül az összes esetben felmerül a kérdés a pedagógusok részéről, hogy mit kezdjenek az óriási méretű adathalmazokkal, hogyan fordítsák le azokat a helyi kontextusra, és voltaképpen mit tegyenek a tanulói teljesítmények fejlődéséért. Ezeket a kérdéseket más és más módokon közelítik meg az egyes oktatási rendszerek, saját oktatási kultúrájukhoz igazítva a megoldásokat.

A kötet szerzői pontokba szedték az adatok optimális szintű használatához kapcsolódó szükségleteket. Mindenekelőtt a pedagógusoknak képessé kell válniuk az önálló adatgyűjtésre, elemzésre, interpretálásra és felhasználásra; vagyis invesztálni kell a pedagógusok továbbképzésébe. Kanadában és Belgiumban a

pedagógusok adatfeldolgozói kompetenciájának növelése céljából támogatói rendszert alakítottak ki. Bevezettek egy minden iskola számára szabadon elérhető *help desk* szolgáltatást, képzett kutatók és pedagógiai tanácsadók irányításával külső és belső tréningeket szerveztek, és olyan felhasználói felületeket hoztak létre, ahol a pedagógusok egymással is meg tudják osztani kérdéseiket, tapasztalataikat. A különböző támogatások eltérő eredményeket hoztak, így például a kutatók vezette tréningek résztvevőinek nőtték az adathasználattal kapcsolatos ismeretei, míg a pedagógiai tanácsadók vezette továbbképzések résztvevői megtanulták, hogyan lehet a saját iskolai kontextusukban interpretálni az információkat.

Hollandiában és Angliában jelenleg problémaként jelenik meg a pedagógusok adathasználattal szembeni negatív hozzáállása. A pedagógusok szemléletbeli változása még nem történt meg maradéktalanul, többségük továbbra is saját pedagógiai tapasztalataira hagyatkozik. A szemléletváltást nehezíti az elszármaztatottság érzése: a tanárok többsége központi ellenőrzésként éli meg az adatgyűjtés fogalmát; úgy vélik, a kormány nem bíz meg pedagógiai szakértelmükben. Ennek ellensúlyozására az egyes központok igyekeznek az adathasználat támogató, segítő jellegét hangsúlyozni, emellett újra és újra megismétlik, hogy az adatok nem használhatók fel a pedagógusok hibáztatására, megszégyenítésére vagy szankcionálására. Ennek ellenére az adatvezérelt döntéshozatal buktatójának továbbra is a fejlesztés és az elszámoltathatóság kettősségét és kényes egyensúlyát tartják.

A rendelkezésre álló adatok maximális kihasználásához elengedhetetlen az iskolavezetés és a pedagógusok együttműködése, az információ szabad áramoltatása. Hollandiában, Angliában és Dél-Afrikában iskolai és tantárgyi szintű adatfeldolgozó csoportokat alakítottak ki pedagógusokból, amelyek feladata igen sokrétű: adatokra alapozott megbeszéléseket szerveznek; elősegítik a szakmai fejlődést, a tantervi tervezést és fejlesztést; rendszeres monitoro-

zást végeznek, és ezek alapján célokat állítanak fel. Dél-Afrikában az adatelemzés logisztikájára több mintát is alkalmaznak az iskola oktatási kultúrájától függően. Az adatok lépcsőzetes terjesztését és feldolgozását például olyan intézményeknek javasolják, ahol a pedagógusok különböző szerepekkel vesznek részt az adatfeldolgozásban, míg a csapatmunkát inkább olyanoknak, ahol további analíziseket és különböző adatok triangulációját is bevonják az elemzési folyamatba.

Sok esetben feszültséget okoz, ha az iskolán belül nem tisztázott, hogy kinek a feladata az információk összegyűjtése és feldolgozása. Egy angol felmérés szerint az oktatási adatok kezelését az esetek kétharmadában egy idősebb kolléga, míg az esetek fennmaradó részében több idősebb kolléga végzi. A megkérdezett pedagógusok véleménye arról, hogy ki kezelje az adatokat, azonban már más képet mutat. A többség úgy véli, hogy az adatkezelést a szakvezető tanároknak vagy a senior kollégák egy csoportjának kellene végeznie, míg az adatok értelmezéséért jobb, ha az egyes szaktanárok a felelősek. Ugyanakkor az is kiderült a felmérésből, hogy a pedagógusok nem egyformán jutnak hozzá a különböző oktatási adatokhoz, és eltérő mértékben tudják hasznosítani (értelmezni) azokat – az életkor és a pedagógiai tapasztalat egyenes arányosságot mutat az adatfelhasználás fokával.

A bemutatott példák összegzéseként a szerzők megalkották az adathasználat egy elméleti modelljét. A modell három főbb részre tagolódik: 1. adottságok és korlátok (iskolai szerkezet és kontextus; adatok és adatrendszerek; felhasználói jellemzők); 2. adathasználat (nincs adathasználat, kívánt felhasználás, nem kívánt felhasználás); valamint 3. eredmények (az érintettek, pl. iskolavezetők, pedagógusok tanulása; a tanulók tanulása). Mindhárom területre befolyással bír az adott oktatási rendszer politikája. A politikai kontextus befolyásolja az adathasználat adottságait és korlátait, melyek viszont kihatnak az érintett résztvevők tanulási folyamatára, vagyis az eredmények szintjére.

A különféle adottságok és korlátok befolyásolják továbbá az adatok használatát, ami visszafelé is igaz, vagyis kétirányú kapcsolatról van szó (179. o.).

A kötet célja az oktatási adatok felhasználási lehetőségeinek bemutatása döntéshozók, kutatók, iskolavezetők és pedagógusok számára, hét nemzet reprezentatív példáján keresztül. A szerzők kíváncsággal említi az adathasználat politikai ideológiától mentes természetét, illetve az adatvezérelt döntések tisztán fejlesztő jellegű szándékát. Az összegyűjtött minták alapján úgy vélik, hogy az oktatási adatokra alapozott döntéshozatal erőteljes és fenntartható fejlődéshez vezet. Többször is hangsúlyozzák, hogy a bemutatott országoknak eltérő az oktatási rendszere, ezáltal nem is egy-egy rendszer komplex átvételét javasolják, hanem azok átalakítását a saját oktatási környezetnek megfelelően.

Ugyan a kötet különálló tanulmányokból áll, mégis érdemes azokat egymást követően olvasni. Az elméleti fejezetek kiváló alapot nyújtanak a reprezentatív példák megértéséhez, melyeket a fejezetek végén található reflektív kérdések segítségével akár munkacsoportokban is feldolgozhatók. A kötet segítségével pedig akár a mai magyar adathasználati jellemzőket is be tudjuk azonosítani.

(Kim Schildkamp, Mei Kuin Lai & Lorna Earl (eds): *Data-based Decision Making in Education. Challenges and Opportunities*. Springer, 2013, Dordrecht, London. 216 p.)

Bús Enikő

ADATBŐSÉG AZ EURÓPAI FELSŐOKTATÁSI TÉRSÉGBEN

Az 1999-ben a Bolognai Nyilatkozattal útjára indított európai felsőoktatási reformfolyamatot – értelemszerűen – kezdettől fogva nagy apparátussal működtetett rendszeres statisztikai adatgyűjtések kísérik. Az Európai Felső-

oktatási Térség (EFT) létrehozását, a kitűzött célok elérésében történt előrehaladást szisztematikusan követik ezek az adatgyűjtések. Országsoros feldolgozások készülnek, amelyek külön érdekessége, hogy egyre több európai országot érintenek (1999-ben Bolognában 29 ország csatlakozott a kezdeményezéshez, majd folyamatos "létszámemelkedéssel" 2015-ben már 47 ország volt a tagok között, amikor ismét bővült a kör, mégpedig Belorusszia csatlakozásával).

Két kiadványsorozat érdemel kiemelt figyelmet. Az egyik a *Stocktaking Report* sorozat, amely az oktatásért (felsőoktatásért) felelős miniszterek konferenciái számára adott áttekintést az előző tanácskozás óta történetekről. A reform első tíz évében két évente került sor a Miniszteri Konferenciákra, és 2003-2009 között készültek *Stocktaking Report*-ok. A reform második szakaszában áttértek a három évente történő értékelésre, amit az is jelez, hogy 2012-ben és 2015-ben már *Bologna Process Implementation Report* a kiadvány címe.

A másik sorozatot az Európai Egyetemi Szövetség (European University Association – EUA) adja ki *Trends* címmel. Pontosabban 1999-től 2007-ig nevezték így a szériát, és a kiadványok ugyancsak a Miniszteri Konferenciák előtt, azok számára készültek. A 2009-es Konferenciára már nem állítottak össze ilyen értékelést, viszont 2010-ben az egész tízéves folyamat áttekintését végezték el. A következő, egyben legújabb kiadvány 2015-ben született, most már ismét az esedékes Miniszteri Konferencia előtt.

A két sorozat lényegében ugyanarról a jelenségről, kérdéskörrel szól, viszont más-más megközelítésben. A *Stocktaking* (ill. *Implementation*) *Report* a Bologna-folyamat csúcsszerve, a Nemzetközi Bologna Csoport (Bologna Follow-up Group) felkérésére készül a nemzeti Bologna Bizottságok által összeállított országjelentések és más információgyűjtések eredményeinek felhasználásával. Elsősorban kvantitatív módon dolgozza fel az adatokat, amelyeket sokoldalúan és korrekten elemez

is. A megelőző két (most már három) év történéseit, eredményeit tárgyalja, aminek alapján meg lehet fogalmazni a következő két (most már három) év feladatait. A *Trends Report*ok többféle információgyűjtési módszer bevetésével készülnek, elsősorban a tagintézményektől, az egyetemektől begyűjtött adatokra, véleményekre alapoznak, tehát az intézményi szintre fókuszálnak, és a változásokra helyezik a hangsúlyt, jelentős mértékben bevetve kvalitatív eszközöket is. Mindkét sorozat törekszik a stabilizásra a jelentések struktúráját illetően, hogy lehetőség legyen longitudinális elemzésekre, miközben a felhasznált indikátorok körét a szakmailag indokolt mértékben és jelleggel időnként módosítják. A *Stocktaking* (illetve *Implementation*) *Report*oknál erőteljesebb a törekvés a stabilizásra, a *Trends Report*ok viszont egy-egy fő téma köré csoportosítják az elemzésüket, ezért annak logikáját tükrözi a szerkezetük. Az előbbire jellemző valamilyen "kincstári optimizmus" és a nemzeti, illetve az EFT-szintű kormányzati szempontok követése, az utóbbira a kritikusabb megközelítés, az akadémiai és az intézményi szempontok kiemelése.

A 2015-ös év alkalmat ad arra, hogy a két kiadványt egymás mellett, egymáshoz képest is áttekintsük, és az általuk bemutatott hatalmas adathalmazt különböző elemzésekre felhasználjuk.

Az idei *Implementation Report* összeállításába – a sorozat folyamatos professzionálisizálódásának keretében – kiemelt szerepet kaptak az egyre nagyobb teret nyerő, nagy európai statisztikai adatgyűjtési és -elemző intézmények, az Eurostat, az Eurostudent és az Eurydice. Felépítése viszont követi a sorozat hagyományait. Hét fő tématerülettel foglalkozik: a fokozatok, végzettségek rendszerével; a képzési reform fő eszközeivel (a képesítési keretrendszerekkel, az oklevélmelléklettel és ECTS típusú kreditrendszerrel); a minőségbiztosítással; a szociális dimenzióval; az életen át tartó tanulással; a tényleges kimenet kér-

désével és a foglalkoztathatósággal; valamint a nemzetköziesedéssel, a mobilitással.

A fokozatok, végzettségek rendszere minden bizonnyal a legnagyobb jelentőségű téma, amely megalapozza a többi megértését is. Kardinalis kérdés ugyanis, mennyiben és hogyan érvényesítették a részt vevő országok a lineáris (többfokozatú, többciklusú) felsőoktatási rendszerre való áttérést. A tiszteletre méltóan sokoldalú elemzés fő tanúsága, hogy az egyes országok közötti, kezdettől fogva tapasztalt különbségek minden jel szerint stabilan fennmaradtak. Eltérő az egyes fokozatok megszerzésének időbeli hossza (és ECTS-tartalma), az akadémiai és a szakmai irányultságú képzési programok kezelése (megkülönböztetése vagy ennek hiánya), az osztatlan képzések súlya (ezek az ún. professzionális vagy szabályozott szakterületeken, mint az orvosképzés és a jogászképzés, általánosan elterjedtek, a tanárképzési programoknál is több helyen jelen vannak). A BA/BSc programok esetében jellemző a törekvés a képzés hosszának növelésére – a tapasztalatokra alapozott igényeket követve. Mozgásban van, és különböző mértékben jelentkezik a rövid (általában szakmai) programok befogadása a felsőoktatási fokozatok rendszerébe. Nagyon a különbségek abban is, hogy a hallgatók mekkora része folytatja tanulmányait az első fokozat megszerzése után mesterfokon: az egynegyed alatti aránytól a háromnegyed felettiig alakul a mutató. A kötet összeállítói nem minősítik a helyzetet, de ez nem is az ő feladatuk. Igen összetett okokra vezethető vissza a vázolt kép, aminek minősítése igen messzire, az EFT alapfilozófiájának értelmezéséig vezetne.

Nagy figyelmet fordít a kötet a tanulmányok sikeres elvégzésére, más megfogalmazásban a hallgatói lemorzsolódás témájának adatyszerű megközelítésére. Valójában csak survey jellegű rendszeres kohorsz vizsgálatokkal lehetne követni a hallgatók tanulmányi útját. Mivel ilyen munkákat rendszeresen úgyszólván sehol nem végeznek, más megoldásokkal kell próbálkoznunk. A kiadvány összeállítói 17 országtól kaptak használható adatokat egy egyszerű

és csak megközelítő képet ígérő indikátor megalkotásához. Igen nagy különbségekkel (48 és 88% között) alakul azon hallgatók aránya, akik „időben” elvégzik tanulmányaikat. Ez rossz hír, hiszen a tanulmányi rendszeren történő gyors áthaladás minden felsőoktatási szereplő érdeke lenne. Természetesen az igen összetett jelenség megértése, a konkrét adatok értelmezése még várat magára, de legalább komolyan felvetődött a kérdés.

Az utóbbi évek fejleményeit jelzi, hogy kiemelt figyelmet kap a kötetben az ún. szociális dimenzió, amely ezúttal a hallgatókat érintő kérdésekre irányul. Egyre gazdagabbá válik az az indikátorrendszer, amely alapvető szociális összetételüket, előtanulmányaikat, anyagi helyzetüket írja le. Tényszerűen kibontakozik a hallgatói populáció – eddig is „sejtett” – sokfélesége. Olyan gyakorlatias ügyekre is kitér, mint az előzetes tanulmányok beszámítása, a hallgatói szolgáltatások jellemzői és az igen bonyolult és változó tandíjrendszer.

A hallgatói és tanári (alkalmazotti) mobilitás kezdettől fogva kiemelt témája az EFT-nek, hiszen a Bologna-reform egyik elsődleges célja az volt, hogy a képzési rendszerek összehangolásával könnyebbé váljon a nemzetközi mobilitás, főleg az EFT-n belül. A Miniszteri Konferenciák és előkészítő anyagaik mindig elégedetlenséget fejeztek ki az eredményekkel kapcsolatban ezen a területen, nem tér el ettől a 2015-ös kötet sem. Bár kissé emelkedett a nemzetközi mobilitásban részt vevő hallgatók aránya 2012 óta, de ez az arány továbbra is igen alacsony (szinte mindenütt 10% alatt van, EFT-szinten pedig 4-5% körüli). Ma már egész indikátorrendszerrel méri a jelenséget, de ez sem tudja a maga összetettségében leírni azt, ráadásul nem minden országban történik meg a szisztematikus adatgyűjtés. Itt egy különös helyzettel találkozunk. Mondhatni a nemzetközi mobilitás mértéke az egyetlen indikátor, amellyel egyszerűnek és egyértelműnek látszó módon értékelni lehet a reform sikerességét. Az EFT második évtizedének kezdetekor a miniszterek abban állapodtak meg, hogy a cik-

lus végére, 2020-ra, a nemzetközi mobilitásban részt vevő hallgatók arányának el kell érnie a 20%-ot. Az elemzés aggodalmát fejezi ki, hogy a jelek szerint erre alig van esély.

Megállapítható, hogy miközben a kötet a korábbinál finomabb, bővített eszközrendszerrel dolgozik, továbbra is beleütközik az adathiányból adódó problémákba, részben egyes országok kimaradása következtében, részben pedig a meglévő indikátorok terén továbbra is fennálló szűkössége miatt. Új elem viszont, hogy mintha a megszokottnál erősebb kritikai attitűdöt képviselne.

A Trends Reportok mindig kapnak címet is, amely arra a kérdéskörre utal, amelyet kiemelten vizsgál a jelentés. A 2015-ös jelentés címe „Tanulás és tanítás az európai egyetemeken”, és alapvetően arra a kérdésre keresi a választ, mennyiben valósultak meg a hallgatóközpontú tanulás/tanítás erősítésével kapcsolatos, korábban meghirdetett tervek.

A fő információforrást a tagegyetemeknek és rektori konferenciáknak, illetve további érintett, érdekelt szervezeteknek, stakeholdereknek kiküldött kérdőívek adták. (451 használhatóan kitöltött kérdőív érkezett vissza. A visszaküldési arány 49% volt az EUA tagintézmények körében, amit nehéz minősíteni, mivel a korábbi hasonló adatgyűjtéseknél mindig valamelyest változó volt a megkérdezettek köre és a tematika.)

A válaszoló intézmények összetételét és ezzel a reprezentativitás értelmezhetőségét bemutatva a szerző az európai egyetemeket a következő – nem hivatalos, de tanulságos – típusokba sorolja: többkarú (multidiszciplináris) egyetemek, szakegyetemek, műszaki egyetemek, alkalmazott tudományok egyetemei, speciális főiskolák, nyitott egyetemek.

A Report öt nagy fejezetből áll. Az I. fejezet a változó környezetet mutatja be, amelyben a felsőoktatásnak küzdenie kell, hogy megtartsa, erősítse pozícióit. Talán ez a legértékesebb fejezet, kiválóan foglalja össze az utóbbi évek súlyos történéseit: ami a pénzügyi válság, a kedvezőtlen demográfiai folyamatok, az illegális mig-

ráció Európába, az egyre erősebb globális verseny, és az informatikai forradalom kiteljesedése. Mindezek negatívan hatottak a felsőoktatásra, a diplomások helyzetére (így az ifúságra, ezen belül a diplomás munkanélküliség növekedése, vagy a diplomások áramlása Európán belül és más földrészekről: Délről Északra és Keletről Nyugatra). A kooperációs kényszer és a verseny intenzitása egyszerre erősödik (előbbi az oktatás, utóbbi a kutatás területén jellemzőbb), ami új stratégiák kidolgozását kényszeríti ki.

A II. fejezet az európai és a nemzeti szakpolitikákkal foglalkozik. Megállapítja, hogy az EU és az EFT „kifelé” próbálja demonstrálni, hogy van saját politikája, miközben sokféle, fragmentált és dinamikus reform zajlik nemzeti keretekben. A legjobb példa erre az igen eltérő finanszírozási reformok sora, ami oda hatott, hogy erősödött az aszimmetrikus akadémiai és hallgatói mobilitás. A pénzügyi források apadása az intézmények működésének bürokratikus jellegét erősíti fel, csökkenti a kollegiális döntések terét.

A III. fejezet az intézményi stratégiákról és a változó összetételű hallgatósról szól. Az intézmények mozgásterét nagymértékben befolyásolja a hallgatói létszám. Ebben a tekintetben általában további növekedés tapasztalható, de Közép- és Kelet-Európa több országában stagnálás és csökkenés váltotta fel a korábbi expanziót. A hallgatók összetétele egyre változatosabbá válik (főleg életkor és előzetes tanulmányok szempontjából), de közben a társadalmi esélykülönbségek fennmaradnak (a szerző egy OECD adatra hivatkozik, miszerint a felsőfokú végzettségű szülők gyermekeinek esélye az érettségi utáni továbbtanulásra 4,5-szer nagyobb az alacsonyabb végzettségű szülők gyermekeinek esélyeihez képest). Mindez tulajdonképpen bevezetés a IV. fejezet témájához, a tanulás és tanítás kérdéséhez. A fejezet a Bologna-reform hatásának rövid bemutatásával kezdődik, sommásan megállapítva azt, hogy a három fokozat tantervei túl gyorsan, nem eléggé átgondoltan készültek, és főleg a BA/BSc programokra jellemző, hogy az eredeti elképze-

lésekhez képest igen rugalmatlanok. A reform több eleme mára elavult (ilyen pl. az Oklevél-melléklet rendszere). A képesítési keretrendszerek kidolgozása és bevezetése a vártnál lassabban történik, és ugyancsak rugalmatlanságot jelentő elemeket hordoz.

Az V. fejezetben a szerző megfogalmazza a következő évtized egyetemeire vonatkozó vízióját. Eszerint a hallgatói populációban továbbra is növekedni fog a nem tipikus életkorúak, a munka mellett tanulók, a kifejezetten munkaerő-piaci ismereteket igénylők aránya, amire a diverzifikált, hallgatóközpontú tanulási és tanítási módszerek elterjesztésének gyorsításával kell válaszolni. A felsőoktatásban további teret nyer az üzleti típusú gondolkodás, elmosódnak a határok a magán és a közösségi (állami) szektor között. Az egyre több adminisztratív elem-

mel terhelt oktatási és a kutatási tevékenységet korszerű, ún. para-akadémiai stáb kiépítésével fogják megtámogatni, amely az akadémiai és az adminisztratív stáb között foglal helyet. Mind ehhez közös európai tervre van szükség. A tárgyalt öt évben ugyanis kevésbé volt kezdeményező az európai szint, a reformokat elsősorban nemzeti keretekben indították. Most ismét aktívnak kell lennie az EFT-nek.

(*European Commission/EACEA/Eurydice: The European Higher Education Area in 2015: Bologna Process Implementation Report. Publications Office of the European Union, 2015, Luxembourg. 300 p. – André Sursock: Trends 2015: Learning and Teaching in European Universities. EUA, 2015, Brussels. 128 p.*)

Hrubos Ildikó

ÖSSZEFOGLALÓK

Adatbőség

VEROSZTA ZSUZSANNA

ADMINISZTRATÍV ADATOK TÁRSADALOMKUTATÁSI KEZELÉSE

A tanulmány célja számba venni az adminisztratív adatok társadalomkutatási felhasználásának elméleti és gyakorlati feltételeit és következményeit. Ennek során összefoglaljuk azokat a tényezőket, amelyek az adminisztratív adatok iránti megnövekedett figyelmet táplálják, majd megvizsgáljuk ezen adattípus természetét és viszonyát a kutatási céllal gyűjtött adatokhoz. Ezt követően az adminisztratív adatokon nyugvó társadalomtudományi kutatás szükséges validálási kritériumait és eljárásait tekintjük át, illusztratív példákat is beemelve. Az elemzés során emellett az adatbázisok összekapcsolásában rejlő lehetőségek tipizáló áttekintését végezzük el.

KULCSSZAVAK: ADMINISZTRATÍV ADAT, ADATÖSSZEKAPCSOLÁS, ÉRVÉNYESSÉG, ADATBÁZIS, ELEMI ADAT

NAGY PÉTER TIBOR

A RÉGI ÉS ÚJ TÖRTÉNETI ADATOK TÁRSADALMI TERMÉSZETE

A tanulmány három - eltérő méretű - altémát fejt ki: a régi típusú, de új adatok keletkezését; az új típusú, új adatok keletkezését és a régi adatok újrahasznosulását. Minden egyes altémán belül először azzal foglalkozik, hogy az „adatok új világa” (mely valószínűleg a számítástechnikával, s különösen az internettel vette kezdetét) hogyan alakítja át - funkcionalistán szólva - a tudományos témákat és módszereket, majd arra, hogy - konfliktusos paradigmában fogalmazva - a tudományos közösségek milyen érdekviszonyok alapján kezelik, illetve ismerik el egyáltalán releváns adatnak az új információs helyzet „termékeit”. A vizsgálat során tehát az adatfeltárás, adatrendezés és adatelemzés jelenségvilágát egy komplex tudományszociológiai összefüggésrendszer részeként törekszünk értelmezni.

KULCSSZAVAK: ADAT, ADATBŐSÉG, ADATROBBANÁS, TUDOMÁNYSZOCIOLÓGIA, KONFLIKTUSOS PARADIGMA, FUNKCIONALISTA PARADIGMA, KUTATÁSMÓDSZERTAN, DIGITALIZÁLÁS

GÁRDOS ÉVA

ADATOK ÉS KEZELÉSÜK A HIVATALOS STATISZTIKÁBAN

Jó minőségű statisztikai mutatók előállításához nagyon fontos a megfelelő adatforrás kiválasztása. A tanulmány bemutatja a hivatalos statisztikai munkában használható adatforrás-típusokat és alkalmazási területeiket, kitérve a másodlagos adatokra, különös tekintettel az adminisztratív adatok egyre növekvő jelentőségére. Részletes képet kaphat az olvasó a különböző fajta adatok felhasználási feltételeiről, ezek között is nagy nyomatókat kap a hazai és az EU-s jogszabályi háttér. Emellett röviden szó esik a statisztikai termékek és adat-előállítási folyamat minőségi kritériumairól, valamint az adatvédelem szabályairól, és a felfedés elleni védelem során figyelembe veendő szempontokról. A tanulmány átfogó képet ad a hivatalos statisztikai rendszer működéséről és bepillantást enged a KSH tevékenységét meghatározó legfontosabb irányelvekbe is.

KULCSSZAVAK: STATISZTIKAI ADAT, MÁSODLAGOS ADAT, ADMINISZTRATÍV ADAT, STATISZTIKAI MINŐSÉG, STATISZTIKAI FELFEDÉS ELLENI VÉDELEM

SZÉKELY IVÁN

KÖZADATOK ÉS NYILVÁNOS ADATBÁZISOK: A HOZZÁFÉRÉS KÉRDÉSEI

A közadatokhoz való szabad hozzáférés a demokratikus jogállamok alapértéke és alapvető joga, mégis állandó érdekütközések színtere és gyakorlati problémák felvetője. A tanulmány bemutatja a nemzetközi és magyar jog fogalomrendszerét és megközelítésének korlátait a közadatok hozzáférhetősége terén, ismerteti legfontosabb jogintézményeit, majd külön fejezetet szentel az ingyenesség kérdésének, esettanulmányok elemzésével. A szerző bemutat néhány releváns hazai és nemzetközi közadat-hozzáférést segítő technológiát és alkalmazást, majd felvet néhány ritkábban elemzett szempontot a közadatok, nyilvános adatbázisok felhasználási lehetőségei terén. A tanulmány zárásképpen röviden kitér az adatbőség oktatási vonatkozásaira, mind a diákok, mind a pedagógusok szempontjából.

KULCSSZAVAK: INFORMÁCIÓSZABADSÁG, KÖZÉRDEKŰ ADAT, ADATBÁZIS, INGYENESSÉG, KÖZADAT-HOZZÁFÉRÉST SEGÍTŐ TECHNOLÓGIÁK

HALÁSZ GÁBOR

ADATOKKAL TÖRTÉNŐ KORMÁNYZÁS: DILEMMÁK ÉS PERSPEKTÍVÁK

E tanulmány az adatokkal vagy indikátorokkal történő kormányzás dilemmáit és perspektíváit elemzi, különös tekintettel a szakpolitikai koordináció Európai Uniót jellemző gyakorlatára, valamint a tényeken alapuló szakpolitikai döntéshozatal európai normáira. Bemutatja az adatok és indikátorok kormányzási eszközként történő használatával járó előnyöket és az ennek nyomán jelentkező kihívásokat. A tények és adatok közötti különbséget hangsúlyozva bemutatja a tényekre épülő politikai nyomásgyakorlás és meggyőzés gyakorlatának előnyeit, hangsúlyozva ennek korlátait is. A tanulmány utal azokra az új lehetőségekre, amelyek a nagy mennyiségű mikro-szintű, longitudinális adatot gyűjtő adatrendszerek létrejötte nyomán keletkezhetnek.

KULCSSZAVAK: OKTATÁSPOLITIKA, KORMÁNYZÁS, ADATOKKAL TÖRTÉNŐ KORMÁNYZÁS, TÉNYEKEN ALAPULÓ SZAKPOLITIKA, EURÓPAI UNIÓ, SZAKPOLITIKAI KOORDINÁCIÓ, NAGY ADATRENDSZEREK

SZÉLL KRISZTIÁN**AZ ADATHOZZÁFÉRÉS ÉS FELHASZNÁLÁS NEMZETKÖZI GYAKORLATAI**

A tanulmány egy angolszász (Anglia) és egy skandináv (Svédország) kultúrkörből származó országban tárja fel a közoktatási célú adatgyűjtések alapján előálló adatbázisokhoz való hozzáférés és az adatfelhasználás gyakorlatait. A tanulmány arra keresi a választ, hogy a vizsgált nemzeti adatrendszerek (1) milyen célból gyűjtenek adatokat és fejlesztenek indikátorokat, (2) a célok elérése érdekében milyen szintű és tartalmú adatgyűjtések valósulnak meg, valamint (3) milyen hozzáférést biztosítanak az egyes szereplők számára a háttéradatbázisokhoz. A tapasztalatok alapján a fejlett adatrendszerekben az adatgyűjtések és indikátorfejlesztések céljai között a költségvetés-tervezési és tanügyigazgatási szempontok mellett hangsúlyosan megjelennek a fejlesztési, kutatási és értékelési szempontok is. Az egységes alapelvekre épülő és több évtizedre visszatekintő adatkultúrával rendelkező nemzeti adatrendszerek elsősorban egyéni szintű adatokat gyűjtenek. Adott tehát a lehetőség – és a szándék – az anonimizált formában megvalósuló adatintegritációra, hozzátéve, hogy a fejlett adatrendszerek különös figyelmet fordítanak az adatvédelem és az adathozzáférés közötti egyensúly megteremtésére. Összességében elmondható, hogy az adatgyűjtési és hozzáférési alapelvek lefektetése és társadalmi elfogadottsága erősíti az adatok célhoz kötött és széles körű felhasználását.

KULCSSZAVAK: ADATHOZZÁFÉRÉS, ADATGYŰJTÉS, INDIKÁTOROK, FEJLESZTÉS, ÉRTÉKE-LÉS

SALOMVÁRI GYÖRGY**KÖZNEVELÉS AZ ADATOK BŰVÖLETÉBEN – AZ ADATTERMELÉS FORRADALMA**

Az adattermelés forradalma a pedagógia területét sem hagyja érintetlenül. A hatására újonnan megjelenő komplex adatvagyon-kezelő rendszerek a felszín karcoló statisztikai megközelítés helyett alapvetően egyéni szintű adatokra épülnek, ezáltal kifejezetten alkalmasak közvetlen pedagógiai célú alkalmazásra: olyan részletezettséggel ismerhetjük meg a tanítás – tanulás folyamatát, eredményességét és hatékonyságát jellemző adatok időbeli változásait, melyre korábban soha nem volt lehetőség.

A lehetőség egyben felelősség is, az adatvagyon létrehozásán túlmenően a felhasználás körülményeinek és korlátainak meghatározása a legfontosabb feladatunk. Sikert akkor érünk el, ha a pedagógia gyermek és tanulóközpontú komplex megközelítést alkalmaz, melynek integráns eleme a pedagógusok adatvagyon felhasználásra épülő tevékenysége - szakértők és oktatáskutatók által „kívülről” támogatva.

KULCSSZAVAK: ADATTERMELÉS, ADATVAGYON, ADATKAPCSOLATI HÁLÓ, ELEMÍ ADATOK, EGYÉNI SZINTŰ ADATOK, ADATVAGYON HASZNOSÍTÁSRA ÉPÍTŐ PEDAGÓGIAI RENDSZEREK

BOGÁROMI ESZTER - MÁTH ANDRÁS**SOKFORRÁSÚ ADATBÁZIS ÉPÍTÉS - BUKTATÓK, NEHÉZSÉGEK, MEGOLDÁSI KÍSÉRLETEK**

A tanulmány a külső forrásból származó adatbázisok felhasználásának folyamatát mutatja be egy példán keresztül.

Az első szakaszban megismerhetjük az ingyenesen és fizetősen hozzáférhető nyilvános, adminisztratív és egyénsoros adatbázisok körét, azok adatfelvételének hullámain, alap-

egységeit, változóit és további tulajdonságait. Továbbá ebben a szakaszban mutatjuk be azt is, milyen megoldások szolgálhatnak arra, ha a számunkra szükséges adatbázisok nem nincsenek előállítva a piacon. A második szakaszban a különböző adatbázisok összefűzésének buktatóit és nehézségeit mutatjuk be, megoldási javaslatokkal együtt. A példában település szinten történt az adatbázisok összekapcsolása. A harmadik szakaszban a kutatás alkalmával tervezett elemzésekben használt módszertanok tükrében vizsgáljuk a rendelkezésre álló változókat és eseteket. A negyedik fejezetben összefoglaljuk a fent bemutatott adatsorok összekapcsolásának kutatási felhasználásának lehetőségeit, eredményeit.

KULCSSZAVAK: TELEPÜLÉS, ADATHASZNOSÍTÁS, GYAKORLATI PÉLDA, KSH, TGI, ADATBÁZISOK ÖSSZEKAPCSOLÁSA

Data abundance

ZSUZSANNA VEROSZTA

THE GROWING USE OF ADMINISTRATIVE DATA IN THE SOCIAL SCIENCES

The study aims to list and evaluate the theoretical and practical conditions and also the consequences of using administrative data in the social sciences. In this context, the first step is to summarize the factors that have fuelled the increased attention being given to administrative data gathering; secondly, we examine the nature of this type of data and compare it with data collected directly for research purposes. Next, we review the validity criteria pertaining to administrative data-based research, using illustrative examples. In addition, during the analysis we provide an overview of and also categorize the possibilities inherent in database linkage.

KEYWORDS: ADMINISTRATIVE DATA, RECORD LINKAGE, VALIDITY, DATABASE, MICRO DATA

PÉTER TIBOR NAGY

THE SOCIAL NATURE OF OLD AND NEW DATA

This study seeks to provide an explanation for three sub-problems. 1. The social determination of creation/the arising of “old-type” new historical data (for example determinative factors for excavations, etc.); 2. the social circumstances of creation/the arising of “new-type” new historical data (for example, historical medicine) 3. re-interpretation/re-structuring of old data (for example, the digitalisation of sources). There are two topics within this area of sub-problems: 1. how the new “data universe” (the internet) has changed the topics and methods utilized in research; 2. how academic communities have perceived and integrated the new data types and how the face of a community has become altered within this process. The first sub-problem is interpreted from a functionalist perspective, the second from a conflict-theoretical one.

KEYWORDS: DATA, ABUNDANCE OF DATA, SCIENTIFIC CLASSIFICATION, SOCIOLOGY OF SCIENCE, SOCIAL CONFLICT THEORY, FUNCTIONALIST THEORY, RESEARCH METHODOLOGY, DIGITALIZATION

ÉVA GÁRDOS

DATA AND DATA MANAGEMENT IN OFFICIAL STATISTICS

From the perspective of producing statistical indicators of good quality, it is important to select the most appropriate data sources. This study presents data sources of different types which can be used in official statistical work as well as the areas of their application, going into detail regarding secondary data, with special attention being given to administrative data and its growing significance. The reader may get a detailed picture of conditions needing to be taken into account when using data of different types; while major emphasis is given to both Hungarian legislation and that of the European Union. Moreover, the qualitative criteria of the statistical product along with the data production

process are also briefly mentioned, as well as data protection rules and also viewpoints that one should consider in the course of disclosure control. The study provides a broad overview of the functioning of the official statistical system and gives insight into the most important policies and guidelines defining the activities of the Hungarian Central Statistical Office.

KEYWORDS: STATISTICAL DATA, SECONDARY DATA, ADMINISTRATIVE DATA, STATISTICAL QUALITY, STATISTICAL DISCLOSURE CONTROL

IVÁN SZÉKELY

PUBLIC DATA AND PUBLIC DATABASES: QUESTIONS OF ACCESS

Free access to public data is a basic value and fundamental right of a constitutional democracy, though it invariably generates both conflicts of interest and practical problems. This study presents the notions used in international and Hungarian law together with the limitations in their approach and, by analysing case studies, it devotes a separate section to discussing questions of 'access free of charge'. The author presents several relevant technologies and applications supporting access to public data and databases, and he raises some less analysed issues relating to public data and database usage. Finally, the study briefly touches upon the educational potential of data abundance both from the students' and the educators' point of view.

KEYWORDS: FREEDOM OF INFORMATION, DATA OF PUBLIC INTEREST, DATABASES, ACCESS FREE OF CHARGE, TECHNOLOGIES SUPPORTING ACCESS TO PUBLIC INFORMATION

GÁBOR HALÁSZ

GOVERNING BY DATA – DILEMMAS AND PERSPECTIVES

This paper analyses the dilemmas and perspectives of governing by data and indicators in the education sector with a special focus being put on the policy coordination practices of the European Union and on the emerging European norm of evidence-based policy-making. It presents some of the most relevant issues related to the use of data and indicators as instruments of governance, exhibiting the advantages and the challenges underpinning this relatively new practice. Stressing the difference between facts and data, it presents the benefits of using facts as tools of political pressure and persuasion, revealing also the limitations of this practice. The paper makes reference to the new opportunities created by the emergence of big databases containing large quantities of micro level, longitudinal data.

KEYWORDS: EDUCATION POLICY, GOVERNANCE, GOVERNING BY DATA, EVIDENCE-BASED POLICY, EUROPEAN UNION, POLICY COORDINATION, BIG DATA

KRISZTIÁN SZÉLL

DATA ACCESSIBILITY IN AN INTERNATIONAL CONTEXT

The study explores the possibilities of data accessibility and the practices of using data in an Anglo-Saxon (England) and in a Scandinavian country (Sweden). The paper tries to answer the following questions: (1) What are the purposes of data collection and indicator development? (2) What level and content of data are collected in order to attain goals? (3) What accessibility is provided regarding databases for stakeholders and publicity? Experience shows that besides budget planning and education administration aspects, research, development and assessment factors are also emphasized among data

collection and indicator development aims. National data systems that have a unified set of principles and a history of data collection looking back several decades mainly gather individual-level data. So the opportunity – and the intent – to integrate anonymized data is given, though developed data systems do put great emphasis on the balance between data protection and data accessibility. All in all, creation and social acceptance of the principles of data collection and accessibility strengthen the aim-assigned and extensive use of data.

KEYWORDS: DATA ACCESS, DATA COLLECTION, INDICATORS, DEVELOPMENT, EVALUATION

GYÖRGY SALOMVÁRI

THE DATA REVOLUTION IN PUBLIC EDUCATION

The data production revolution has not left untouched the area of education. This has caused newly-emerging complex data asset management systems that are fundamentally based on individual-level data instead of the surface abrasive statistical approach, and this is especially suitable when it comes to usage with direct pedagogical purposes: we can now obtain a level of detail regarding teaching-learning process effectiveness and also see the efficacy of typical data changes occurring over time, something which has never previously been possible.

The chance to do this additionally points to a responsibility along with the creation of data asset conditions and limitations for which the ability to determine outcomes is our most important task. Success is achieved when education applies a complex child- and learner-centred approach, which is an integral part of activity-based educators' use of data assets – and this is supported by education experts and researchers from "outside".

KEYWORDS: DATA PRODUCTION, DATA ASSETS, DATA NETWORK OF RELATIONSHIPS, PRIMARY DATA, INDIVIDUAL-LEVEL DATA, UTILIZATION OF DATA ASSETS BUILT ON EDUCATIONAL SYSTEMS

ESZTER BOGÁROMI – ANDRÁS MÁTH

THE POSSIBILITIES OF DATA RECOVERY, LIMITATIONS, PROFITS – PRESENTING AN EXAMPLE

The study demonstrates the process of using databases in a specific example, and our focus is on databases that are available from external sources. In its first part, we reveal the list of free and paid-for databases (used in research) as well as the range of public, administrative and one-line databases; and the section presents different periods of data collection, the basic units, variables and other characteristics. Also shown here is how one may proceed if a data set is not at one's disposal. In our example, we make use of a self-created database that includes all settlements in Hungary – which initially required us to merge several databases, which was not an easy task as many pitfalls and difficulties were encountered. The study also presents some of the possible 'solutions' regarding the way of proceeding and also the most useful results.

KEYWORDS: SETTLEMENT, DATA RECOVERY, PRACTICAL EXAMPLES, HUNGARIAN CENTRAL STATISTICAL OFFICE, TGI, DATABASE MERGING

EDDIGI SZÁMAINK

1992/1 | Iskola és egyház
SZERKESZTŐ: KOZMA TAMÁS

1992/2 | Pénz – piac – iskola
SZERKESZTŐ: LUKÁCS PÉTER

1993/1 | Munkanélküliség és oktatás
SZERKESZTŐ: LISKÓ ILONA

1993/2 | Kisebbségek
SZERKESZTŐ: FORRAY R. KATALIN

1993/3 | Felsőoktatás
SZERKESZTŐ: SETÉNYI JÁNOS

1993/4 | Pedagógusok
SZERKESZTŐ: SZABÓ LÁSZLÓ TAMÁS

1994/1 | Mérleg, 1990–1994
SZERKESZTŐ: LUKÁCS PÉTER

1994/2 | Vezetők
SZERKESZTŐ: DRAHOS PÉTER & GÁL FERENC

1994/3 | Tanterv
SZERKESZTŐ: SZEKENYI PÉTER

1994/4 | Tankönyv
SZERKESZTŐ: NAGY PÉTER TIBOR

1995/1 | Önkormányzatok
SZERKESZTŐ: HALÁSZ GÁBOR & NAGY MÁRIA

1995/2 | Ifjúság
SZERKESZTŐ: GÁBOR KÁLMÁN

1995/3 | Vizsgák
SZERKESZTŐ: NAGY PÉTER TIBOR

1995/4 | Elit
SZERKESZTŐ: CSÁKÓ MIHÁLY

1996/1 | Szakképzés
SZERKESZTŐ: LISKÓ ILONA

1996/2 | Iskolaszervezet
SZERKESZTŐ: LUKÁCS PÉTER

1996/3 | Nők
SZERKESZTŐ: FORRAY R. KATALIN

1996/4 | Európa
SZERKESZTŐ: KOZMA TAMÁS & SZIGETI MIKLÓS GÁBOR

1997/1 | Hátrányos helyzet
SZERKESZTŐ: LISKÓ ILONA

1997/2 | Iskolán kívüli képzés
SZERKESZTŐ: TÓT ÉVA

1997/3 | Régiónk
SZERKESZTŐ: IMRE ANNA

1997/4 | Internet
SZERKESZTŐ: CZEIZER ZOLTÁN

1998/1 | Mérlegen
SZERKESZTŐ: KOZMA TAMÁS

1998/2 | Mentálhigiéné
SZERKESZTŐ: PAKSI BORBÁLA

1998/3 | Pályaválasztás
SZERKESZTŐ: LISKÓ ILONA

1998/4 | Nat
SZERKESZTŐ: SETÉNYI JÁNOS

1999/1 | Felnőttoktatás
SZERKESZTŐ: HINZEN, HERIBERT

1999/2 | Cigányok
SZERKESZTŐ: FORRAY R. KATALIN

1999/3 | Minőség
SZERKESZTŐ: HORVÁTH ZSUZSANNA

1999/4 | Agresszió
SZERKESZTŐ: GÁBOR KÁLMÁN & LISKÓ ILONA

2000/1 | Felsőoktatás, tömegoktatás
SZERKESZTŐ: HRUBOS ILDIKÓ & POLÓNYI ISTVÁN

2000/2 | Kisebbségek Közép-Európában
SZERKESZTŐ: KOZMA TAMÁS & RADÁCSI IMRE

2000/3 | Tankönyv
SZERKESZTŐ: GÁL FERENC

2000/4 | Nyelvtudás, nyelvvoktatás
SZERKESZTŐ: IMRE ANNA

2001/1 | Oktatás – politika – kutatás
SZERKESZTŐ: KOZMA TAMÁS

2001/2 | Fogyatékos fiatalok
SZERKESZTŐ: ILLYÉS SÁNDOR

2001/3 | Értékek
SZERKESZTŐ: LISKÓ ILONA

2001/4 | Ezredforduló
SZERKESZTŐ: KOZMA TAMÁS

2002/1 | Mérlegen, 1990–2002
SZERKESZTŐ: LUKÁCS PÉTER

2002/2 | Diplomások
SZERKESZTŐ: NAGY PÉTER TIBOR

2002/3 | Család
SZERKESZTŐ: SOMLAI PÉTER

2002/4 | Taneszközpolitika
SZERKESZTŐ: GÁL FERENC

2003/1 | Felsőoktatási reformok
SZERKESZTŐ: HRUBOS ILDIKÓ & POLÓNYI ISTVÁN

2003/2 | Felvételi
SZERKESZTŐ: NAGY PÉTER TIBOR

2003/3 | e-Learning
SZERKESZTŐ: TÖRÖK BALÁZS

2003/4 | Európai Unió
SZERKESZTŐ: BAJOMI IVÁN

2004/1 | Alternatív oktatás
SZERKESZTŐ: LISKÓ ILONA & TOMASZ GÁBOR

2004/2 | Műveltség
SZERKESZTŐ: SÁSKA GÉZA

2004/3 Pedagógusképzés SZERKESZTŐ: NAGY MÁRIA	2010/1 Mérleg, 2006–2010 SZERKESZTŐ: KOZMA TAMÁS
2004/4 Politikai szocializáció SZERKESZTŐ: CSÁKÓ MIHÁLY	2010/2 Fiatalok SZERKESZTŐ: FEHÉRVÁRI ANIKÓ & SZEMERSZKI MARIANNA
2005/1 Budapest SZERKESZTŐ: NAGY PÉTER TIBOR	2010/3 Felsőoktatás és foglalkoztathatóság SZERKESZTŐ: HRUBOS ILDIKÓ
2005/2 Hallgatói mobilitás SZERKESZTŐ: HRUBOS ILDIKÓ	2010/4 Oktatás és politika SZERKESZTŐ: BAJOMI IVÁN
2005/3 Egyházak és oktatás SZERKESZTŐ: PUSZTAI GABRIELLA & RÉBAY MAGDOLNA	2011/1 Ideológiák SZERKESZTŐ: SÁSKA GÉZA
2005/4 Óvodák SZERKESZTŐ: BAJOMI IVÁN & TÖRÖK BALÁZS	2011/2 Külföldiek SZERKESZTŐ: POLÓNYI ISTVÁN
2006/1 Mérleg, 2002–2006 SZERKESZTŐ: KOZMA TAMÁS & LISKÓ ILONA	2011/3 Átalakuló szakmák SZERKESZTŐ: FEHÉRVÁRI ANIKÓ
2006/2 Képzés és munkaerőpiac SZERKESZTŐ: POLÓNYI ISTVÁN & GYÖRGYI ZOLTÁN	2011/4 Menedzserizmus SZERKESZTŐ: KOZMA TAMÁS & VEROSZTA ZSUZSANNA
2006/3 Ötvenhar SZERKESZTŐ: NAGY PÉTER TIBOR & SÁSKA GÉZA	2012/1 Magyar kisebbségek az oktatásban SZERKESZTŐ: PAPP Z. ATTILA
2006/4 Változó egyetem SZERKESZTŐ: HRUBOS ILDIKÓ	2012/2 Látszat és való SZERKESZTŐ: LUKÁCS PÉTER
2007/1 Előítéletek SZERKESZTŐ: ERŐS FERENC	2012/3 Értékelés és politika SZERKESZTŐ: HORVÁTH ZSUZSA & FEHÉRVÁRI ANIKÓ
2007/2 Ekvivalenciától a kompetenciáig SZERKESZTŐ: POLÓNYI ISTVÁN	2012/4 Tantárgyak és társadalom SZERKESZTŐ: NAGY PÉTER TIBOR
2007/3 Felsőoktatók SZERKESZTŐ: NAGY PÉTER TIBOR	2013/1 Centralizáció, decentralizáció, demokrácia SZERKESZTŐ: SÁSKA GÉZA
2007/4 Társadalmi nemek SZERKESZTŐ: FORRAY R. KATALIN & KÉRI KATALIN	2013/2 Egészség és oktatás SZERKESZTŐ: POLÓNYI ISTVÁN
2008/1 Minőségügy a felsőoktatásban SZERKESZTŐ: POLÓNYI ISTVÁN	2013/3 Tudáselosztás, tudásmonopóliumok SZERKESZTŐ: HRUBOS ILDIKÓ
2008/2 Informális tanulás SZERKESZTŐ: TÓT ÉVA	2013/4 Iskolázottság SZERKESZTŐ: NAGY PÉTER TIBOR
2008/3 Veszélyes iskola SZERKESZTŐ: SÁSKA GÉZA	2014/1 Mérleg, 2010–2014 SZERKESZTŐ: KOZMA TAMÁS
2008/4 Támogató programok SZERKESZTŐ: BAJOMI IVÁN	2014/2 Felsőoktatási expanzió SZERKESZTŐ: POLÓNYI ISTVÁN
2009/1 Felsőoktatás és tudománypolitika SZERKESZTŐ: NAGY PÉTER TIBOR & POLÓNYI ISTVÁN	2014/3 Vidékfejlesztés és oktatás SZERKESZTŐ: KOZMA TAMÁS
2009/2 Tehetség SZERKESZTŐ: BALOGH LÁSZLÓ	2014/4 Versenyképesség és felsőoktatás SZERKESZTŐ: HRUBOS ILDIKÓ & TELCS ANDRÁS
2009/3 A „bolognai tanárképzés” SZERKESZTŐ: PUKÁNSZKY BÉLA	2015/1 Pedagógus Életutak - Karrierek SZERKESZTŐ: BIRO ZSUZSANNA HANNA
2009/4 Rendszerváltás és oktatáspolitikai, 1989–2009 SZERKESZTŐ: SÁSKA GÉZA	2015/2 PISA – Kritika és védelem SZERKESZTŐ: KOZMA TAMÁS