

Statisztikai Szemle

A KÖZPONTI STATISZTIKAI HIVATAL
TUDOMÁNYOS FOLYÓIRATA

SZERKESZTŐBIZOTTSÁG:

DR. BOD PÉTER ÁKOS, DR. BOZSONYI KÁROLY, ÉLTETŐ ÖDÖN, DR. HARCSA ISTVÁN,
DR. HUNYADI LÁSZLÓ, DR. HÜTTL ANTÓNIA (főszerkesztő), DR. JÓZAN PÉTER,
DR. LAKATOS MIKLÓS, DR. RAPPAI GÁBOR, SÁNDORNÉ DR. KRISZT ÉVA,
DR. SIPOS BÉLA, DR. SPÉDER ZSOLT, SZABÓ PÉTER, DR. VARGHA ANDRÁS,
DR. VITA LÁSZLÓ, DR. VUKOVICH GABRIELLA (a Szerkesztőbizottság elnöke)

93. ÉVFOLYAM 11–12. SZÁM 2015. NOVEMBER–DECEMBER

*A Statisztikai Szemlében megjelenő tanulmányok
kutatói véleményeket tükröznek, amelyek nem esnek szükségképp egybe
a KSH vagy a szerzők által képviselt intézmények hivatalos álláspontjával.*

Utánnomás csak a forrás megjelölésével!

ISSN 0039 0690

Megjelenik havonta egyszer
Főszerkesztő: dr. Hüttl Antónia
Osztályvezető: Dobokayné Szabó Orsolya
Kiadja: a Központi Statisztikai Hivatal
A kiadásért felel: dr. Vukovich Gabriella
2015.137– Xerox Magyarország Kft.

Szakreferensek: dr. Németh Zsolt, dr. Laczka Éva
Szerkesztők: Bartha Éva, dr. Kondora Cosette, Visi Lakatos Mária
Tördelőszerkesztők: Bartha Éva, Simonné Káli Ágnes
A Folyóiratszemle összeállítója: Lencsés Ákos

Szerkesztőség: Budapest II., Keleti Károly utca 5–7. Postacím: Budapest, 1525. Postafiók 51.

Telefon: 345-6908, 345-6546

Internet: www.ksh.hu/statszemle

E-mail: statszemle@ksh.hu

Kiadó: Központi Statisztikai Hivatal, Budapest II., Keleti Károly utca 5–7.

Postacím: Postafiók 51. Budapest, 1525. Telefon: 345-6000

Előfizetésben terjeszti a Magyar Posta Rt. Hírlap Üzlet (1089 Budapest, Orczy tér 1.).

Előfizethető közvetlen a postai kézbesítőknél, az ország bármely postáján,
valamint e-mailen (hirlapelofizetes@posta.hu) és faxon (303-3440).

További információ: 06-80-444-444

Előfizetési díj: fél évre 6 000 Ft, egy évre 10 800 Ft

Beszerezhető a KSH Információs szolgálatán (Budapest II., Fényes Elek u. 14–18. Telefon: 345-6789)

Tartalom

Tanulmányok

Az adatvédelem és a statisztika kapcsolatának jogi szabályozása – <i>Lakatos Miklós</i>	1017
A statisztikai adatvédelem és -hozzáférés szabályai az Európai Unióban – <i>Nagy Eszter</i>	1051
Az információszabadság szerepe, jelentősége, aktuális kihívásai – <i>Révész Balázs – Buzás Péter</i>	1070
Az adat-hozzáférési igények és az adatvédelmi kötelezettségek egyensúlyáról – <i>Vereczkei Zoltán</i>	1089
Nyilvános mikroadatfájlok összeállításának főbb jellemzői, különös tekintettel az adatvédelmi szempontokra – <i>Kristóf Péter</i>	1112
Az EU-SILC- és az EU-LFS-adatok nyilvános mikroadatfájljai – <i>de Wolf, Peter-Paul</i>	1140
A célzott adatsere módszere a térstatisztikában – <i>Nagy Beáta</i>	1152

Műhely

A Központi Statisztikai Hivatalban kezelt statisztikai adatok gyakorlati hozzáférése – <i>Mátyás-Bodovics Edina</i>	1170
A statisztikai táblázatokban található kis előfordulási számokkal kapcsolatos angol bírósági ítélet – <i>Alexin Zoltán</i>	1177

Fórum

Beszélgetés Péterfalvi Attilával a Nemzeti Adatvédelmi és Információszabadság Hatóságának elnökével – <i>Lakatos Miklós</i>	1183
A Nemzeti Adatvédelmi és Információszabadság Hatóság 2014. évi beszámolója – <i>Lakatos Miklós</i>	1190
Hírek, események	1197

Szakirodalom

Folyóiratszemle

Martin, D. – Cockings, S. – Harfoot, A.: Földrajzi keretrendszer fejlesztése népszámlálási munkahelyadatokhoz – (<i>Kristóf Péter</i>).....	1202
Brandt, M.: Az európai mikroadat-hozzáférés fejlesztése – (<i>Pécs Mária</i>)	1204
Kiadók ajánlata	1206
Társfolyóiratok	1209

Az adatvédelem és a statisztika kapcsolatának jogi szabályozása

Lakatos Miklós,
a KSH szakmai főtanácsadója
Email: Miklos.Lakatos@ksh.hu

A statisztika és az adatvédelem kapcsolata már több mint száz éve tárgya a jogi szabályozásnak. A társadalmi-gazdasági fejlődés szükségszerűen maga után vonja e szabályozás folyamatos frissítését. A szerző cikkében egyrészt áttekinti az említett hosszú időszak változó törvényeit, különös tekintettel a jelenleg hatályos statisztikai törvényre, másrészt felhívja a figyelmet arra, hogy a jelen korszakban és a közeljövőben is szükségessé válik az időközben elavult jogszabályok helyett újak megalkotása a statisztika és az adatvédelem tárgyában.

TÁRGYSZÓ:
Adatvédelem.
A statisztika története.
Statisztikai törvény.

Az emberi civilizáció fejlődése során az államok kialakulása, a gazdasági tevékenység sokrétűvé válása nem történhetett volna meg a szükséges információk nélkül. Az ókori és a középkori államok információéhségét főként adózási és katonai szempontok határozták meg. Azonban a személyiségi jogok védelme ezekben a történelmi korokban még fel sem merült, ez az eszme csak az európai felvilágosodás korától kezd hatni és az állami jogalkotási folyamatokba beépülni.

Az ipari forradalom következtében előállott robbanásszerű gazdasági fejlődés hatására jelentős társadalmi változások zajlottak Európa fejlett régióiban. Nem függetlenül a különböző társadalmi, politikai mozgalmak hatásától, előtérbe került az egészség- és nyugdíjbiztosítás, melynek megszervezése és kiterjesztése a munkavállaló népség egészére naprakész, hiteles információtömeget igényelt. Az ipari termeléssel, a kereskedelmi forgalom növekedésével, a szolgáltatási szektor jelentőségének növekedésével, az országok közötti gazdasági, társadalmi kapcsolatok erősödésével további jelentős mennyiségű információ-halmazra volt szükség.

A begyűjtendő információk többsége egyedi, „nevesített”, személy és/vagy gazdasági szervezet adataként kerül be a rendszerbe, lévén az adatfelhasználás főként redisztribúciós, igazgatási jellegű. Azokat az adatgyűjtéseket, amikor a közfeladatot ellátó, valamint egyéb, jogszabály által meghatározott szervek igazgatási feladatainak ellátására, jogok és kötelezettségek megállapítására, végrehajtására gyűjtik az egyénekre vagy gazdasági szervezetekre vonatkozó azonosítható, egyedi információkat, adminisztratív célúnak nevezzük.

A társadalmi, gazdasági fejlődés során azonban egyre fontosabbá váltak az ún. statisztikai célú adatgyűjtések, melyeknek nem célja, hanem eszköze az egyedi adatok gyűjtése, minek segítségével a bizonyos ismérvek összesítése révén összeállított adatállományok alkalmasak lehetnek a társadalmi, gazdasági és környezeti jelenségek egészének bemutatására, elemzésére. Az összesítések által előállított adatállományok csak úgy jöhetnek létre, ha egyedi (azonosított személyi, illetve vállalati) adatok felvételére és gyűjtésére kerül sor. A statisztikai rendszer adatszolgáltatóira, a megfigyelés egységeire vonatkozó, egyedi azonosítható adatok bizalmas kezelése – a statisztikai titok megtartása –, a statisztikai tevékenység hagyományosan általános, elismert normája.

A statisztikai adatok bizalmas kezelése szoros összefüggésben van azzal a kívánalommal, hogy a statisztikai adatok a megfigyelések tárgyának valós állapotáról adjanak hírt. A statisztikára jellemző, hogy a kiinduló információ nem közvetlenül, hanem valamely adatszolgáltató bevallása útján érhető el. Még ha az információkérésre adandó válasz kötelező jellegét törvény is írja elő, a statisztikai adatfelvétel

akkor is nagymértékben függ a válaszadók együttműködési hajlandóságától. Ha az adatszolgáltató bízhat abban, hogy egyedi adatai a statisztikai rendszeren belül maradnak és azokat kizárólag statisztikai célokra használják fel, az általa adott információ felhasználása nem lesz ártalmas számára, akkor nagyobb valószínűséggel válaszol a valóságnak megfelelően.

A jelen dolgozat terjedelmi korlátai nem engedik meg, hogy e témakör minden részletét bemutassa, célja csupán az, hogy az adatvédelem és a statisztika hazai szabályozásának legfontosabb jellemzőiről adjon összefoglaló képet.

1. Történeti előzmények

A statisztikai tevékenység előképei a népesség-összeírások, a népszámlálások voltak. Népesség-összeírásokat már az ókorban is végeztek. A magyarországi összeírások és népszámlálások főleg adóztatási céllal készültek és így csak a népesség adóköteles részét, gyakran nem is a lélekszámot, hanem az adóalapként funkcionáló telkeket, portákat vették számba. Az 1867. évi kiegyezést megelőzően az első általános, a népesség egészére kiterjedő összeírást Magyarországon II. József rendelte el. Az 1784. július 14-én kiadott népszámlálási rendeletben még nem találunk ún. adatvédelmi szabályokat, de a jogalkotó már megpróbált olyan iránymutatást adni, amely biztosítani kívánta a lakosságot arról, hogy az összeírásnak nem lesznek negatív következményei.

„A tisztviselők pedig, és más az egyes törvényhatóságok által az összeírás végrehajtására kiküldött személyek, mindegyik községben, ahol az összeírást megkezdik, hívassák össze maguk elébe a lakosokat s olvassák azt fel az ott honos nyelven, és nyilvánítsák ki a népnek a Mi nevünkben, hogy ez az összeírás egyáltalában nem újoncok szedése okából, hanem egyedül a közjó érdekében történik, amely feltétlenül szükségessé teszi a népesség számának pontos ismeretét; egyszóval minél keményebben munkálkodjanak azon, hogy miként Mi egyedül a köz szeretetétől vezetve fordulunk a Helytartótanácsához, akképpen az összeírandók lelkéből is teljesen kigyomlátsassék minden kedvezőtlen benyomás az összeírásról.” (*Klinger–Kepecs* [1990] 11. old.)

A magyar statisztikai szolgálat kialakulására, megerősödésére és rendszerszintűvé válására alapvető hatást gyakoroltak az 1870. évtől tízévenként megtartott népszámlálások. A népszámlálások, de különösen az 1870. évi népszámlálás tervezése, vég-

rehajtása, az adatok feldolgozása, a tapasztalatok összegzése jelentősen elősegítette az önálló országos magyar statisztikai hivatal megalakulását, a statisztikai munka szakszerűvé válását. Az 1870. évi népszámlálás a statisztika és az adatvédelem kapcsolatára is hatással volt, hiszen valamilyen formában utalni kellett a népszámlálás során gyűjtött adatok védelmére. Az 1870. évi népszámlálást elrendelő 1869. évi III. törvénycikk a népszámlálásról még nem rendelkezett adatvédelmi kérdésekről, az e témával kapcsolatos normaszöveget *Keleti Károly*, a hivatal létrehozója és első igazgatója sugallatára egy alacsonyabb szintű rendelkezés, a törvényhatóságoknak küldött miniszteri felhívás tartalmazta, felhívta a figyelmet arra, hogy ha olyan „balítéletek” merülnének fel, hogy a népszámlálás „oly célzatú puhatólás akarna lenni, amely új terhek kiszabására, adó felemelésére fogna vezetni, ezeket el kell oszlatni”. (*Thirring* [1983] 8. old.) A magyar statisztika történetében ez az első utalás az adatszolgáltató védelmére egy nagy állapotfelvétel során. Ezt a kitételezt az 1880. és az 1890. évi népszámlálásoknál hasonló módon megfogalmazták.

Az 1870-es évtizedben kialakult és megszilárdult a magyar statisztikai szervezet, 1871-ben állandó szervezettel, pontosan rögzített munkakörrel és azzal a jogosítvánnyal alakult meg az önálló országos magyar királyi statisztikai hivatal, hogy közvetlenül érintkezzék a tárgy szerint érdekelt miniszterek útján megindított statisztikai adatgyűjtésekben az adatszolgáltató hatóságokkal. A statisztikai hivatal tevékenységét tovább erősítette, hogy az első évek tapasztalatai alapján megalkották az országos statisztika ügyének szervezéséről szóló első, 1874. évi XXV. törvényt, amely az intézményeket, gazdasági szervezeteket kötelezte a statisztikai adatok szolgáltatására. (*KSH* [1994c] 5. old.) A törvény a magánszemélyek adatszolgáltatási kötelezettségére nem terjedt ki, adatvédelmi kérdésekről még nem rendelkezett, így népszámlálást elrendelni csak külön törvény megalkotásával lehetett.

Ezt a hiányosságot a Magyar Királyi Központi Statisztikai Hivatalról szóló második, 1897. évi XXXV. törvény pótolta, amely már a régi kereteken túlnőtt statisztikai hivatal munkásságához minden téren szilárdabb alapot adott, számos máig ható, adatvédelemmel kapcsolatos rendelkezést is tartalmazott. Az 1897. évi XXXV. törvény 9. §-a az adatszolgáltatásra korábban is kötelezett közhivatalokon és hatóságokon kívül a magánszemélyekre is kimondja az adatszolgáltatás kötelezettségét, a 3. § rendelkezésének megfelelően az Országgyűlés által elfogadott munkaterv korlátai között. Így ez a rendelkezés már a népszámlálásokra is kiterjedt. Az adatszolgáltató kötelezettségének megszegése esetén a törvény 11. §-a biztosítékokat tartalmaz arra, hogy a be nem szolgáltatott adatot a statisztikai hivatal a mulasztó költségére a helyszínen begyűjtse. A 13. § pótolta a szándékosan elkövetett hamis vagy valótlan adatszolgáltatást, függetlenül attól, hogy magán- vagy nem magánszemély követi azt el. Ehhez kapcsolódik a 14. §-ban leírt szabályozás, amely a hamis hírek terjesztéséből, vagy egyéb okból nem a tervezett időpontban végrehajtott vagy meghíúsult adatfelvétel miatti kártérítési költségekről rendelkezik.

Adatvédelmi szempontból az egyik legfontosabb törvényszakasz a 15. §, ugyanis itt található az a rendelkezés, amely a statisztikai adatszolgáltatás „lelkét” képezi, nevezetesen az, hogy az adatfelvétel során begyűjtött személyes adatok vagy üzleti titkot képező információk nem áramolhatnak ki büntetlenül a statisztikai információs rendszerből.

„15. § A központi statisztikai hivatal alkalmazottja vagy a jelen törvény értelmében szolgáltatandó statisztikai adatok gyűjtésével, vagy feldolgozásával megbízott más egyén, ki ily minőségében tudomására jutott egyéni természetű statisztikai vagy egyéb adatokat, akár alkalmaztatásának vagy megbízatásának idejében, akár annak megszűnte után magánosoknak vagy oly hatóságoknak, melyek azok átvételére jogosítva nincsenek, elbeszél, kiszolgált, felmutat, megtekinteni enged vagy a jelen törvény ellenére köztudomásra juttat – amennyiben cselekedete súlyosabb büntetőjogi beszámítás alá nem esik –, kihágást követ el és 2 hónapig terjedhető elzárással, valamint 600 koronáig terjedhető pénzbüntetéssel büntetendő. Fennmarad ezenkívül a fegyelmi felelősség és az okozott kár megtérítésének kötelezettsége.” (KSH [1994c] 8. old.)

Már ebben a törvényben erős hangsúlyt helyeztek arra, hogy egyéni jellegű statisztikai bevallások az adó kivetésénél alapul nem szolgálhatnak, és a statisztikai hivatal a tudomására jutott adatokat – mai szóval a regisztertípusú adatok kivételével – mindenkor csak területi vagy tárgy szerinti összefoglalásban teheti közzé. Tehát már a XIX. és XX. század fordulóján korszerű módon megfogalmazták az egyedi adatok közzétételének tilalmát.

„17. § Egyéni természetű statisztikai bevallások és az adatgyűjtés vagy ellenőrzés alkalmával a központi statisztikai hivatal tudomására jutott egyéb adatok az adók kivetésénél alapul nem vehetők.

A központi statisztikai hivatal a statisztikai adatokat – a czimtarak körébe tartozó közlemények és azon adatok kivételével, a melyek törvény alapján vagy máskülönben úgyis nyilvánosságra hozatnak – nem egyénenkint, hanem mindenkor csak területi vagy tárgy szerinti összefoglalásban teheti közzé.” (KSH [1994c] 8. old.)

Az 1897. évi XXXV. törvény rendelkezéseit összefoglalva megállapíthatjuk, hogy ez a törvény már tartalmazza azokat az alapvető adatvédelmi szabályokat, amelyek a jelenlegi adatvédelmi szabályozást is jellemzik. Mint említettük, a népszámlálásokat külön törvény rendelte el. Ez időtől fogva az adatvédelmi garanciákat főleg a

statisztikáról szóló éppen hatályos törvények tartalmazták, a népszámlálásokat elrendelő törvény csupán utalt ezekre a biztosítékokra. Ez a jogi megoldás mind a mai napig kisebb-nagyobb megszakításokkal jellemzi a népszámlálással kapcsolatos jogalkotást.

Az 1920-as évtizedben, az első világháborút követő válságok leküzdése után, viszonylag gyors ütemben rendeződtek a gazdasági-társadalmi viszonyok. A megváltozott határok között működőképes közigazgatási rendszer jött létre, amelynek fontos része volt a statisztikai szakszolgálat intézményrendszere. Egyre jelentősebbé vált a központi statisztikai hivatalon kívül folyó statisztikai tevékenység is, rendszeressé váltak más hivatalos szervek adatgyűjtései, amelyek között az együttműködés gyakran nem volt kellőképpen biztosítva. Ezért a kormányzat elhatározta, hogy módosítja (illetve új törvénnyel helyettesíti) az 1897. évi XXXV. törvényt. Az új törvény megalkotásának fő indoka tehát az volt, hogy a statisztikáról szóló törvény ne csak a hivatal tevékenységét szabályozza, hanem terjedjen ki az egész statisztikai szakszolgálatra. A témánk szempontjából fontos adatvédelmi szabályok azonban alapvetően nem változtak, az új törvény más paragrafusszámmal, illetve eltérő formai megoldással, de szinte változtatás nélkül vette át a korábbi a szabályokat.

Az Országgyűlés elfogadta az 1929. évi XIX. törvénycikket a Hivatalos Statisztikai Szolgálatról, amely az Országos Törvénytár 1929. május 22-i 11. számában jelent meg. Mivel az adatvédelemmel kapcsolatos törvényszövegek nagy része egyezett a korábbi törvény szövegével, ezért ezeknek a törvényszakaszoknak az idézésétől eltekintünk.

1948-tól alapvetően megváltozott Magyarország politikai, társadalmi és gazdasági berendezkedése, megtörtént a kommunista hatalomátvétel, kialakult az a szovjet típusú rendszer, amely az ezt követő negyven évet jellemezte. Az egész jogrendszer követte ezt az átalakulást, így a statisztikai adatszolgáltatási rendszert szabályozó jogi környezetet is át kellett formálni. Az 1952. december 25-én kihirdetett 1952. évi VI. törvény az állami statisztikáról, a kor szellemének megfelelően, nem adott garanciákat a személyes adatok védelmére, és hatályon kívül helyezte az e témát részletesen tárgyaló 1929. évi XIX. törvényt. Így például különleges megoldást kellett alkalmazni, az 1960. évi népszámlálás jogi előkészítésénél, mert figyelembe kellett venni azt a tényt, hogy a statisztikai szolgálatról szóló 1952. évi VI. törvény nem tartalmazta azokat az adatvédelemmel kapcsolatos szabályokat, amelyekre a korábbi népszámlálásokat elrendelő törvények hivatkoztak. Így az 1960. évi népszámlálást elrendelő jogszabály (1959. évi 30. sz. törvény) rendelkezett a bizalmas adatok kezeléséről. Ezt a megoldást az 1970., 1980. és az 1990. évi népszámlálást elrendelő jogszabályok is átvették.

Az 1960-as évtizedben, a társadalmi-gazdasági konszolidációt követően, igény jelentkezett arra, hogy oldják a statisztikáról szóló 1952. évi VI. törvény centralizációt szolgáló kötöttségeit, és a kor követelményei szerint szabályozzák a statisztikai adatszolgáltatási rendszert. Ennek megfelelően, az 1897. évi XXV. törvény magánszemé-

lyeket érintő néhány biztosítékát visszaemelték az 1973. évi V. törvénybe. Rendelkeztek arról, hogy milyen szintű jogszabály kötelezhet magánszemélyt statisztikai adatszolgáltatásra. (12. §) (KSH [1994c] 55. old.)

Jogtörténeti szempontból ez az 1952 előtti állapothoz képest visszalépés volt, hiszen magánszemélyre vonatkozó kötelező adatszolgáltatást a törvénynél (törvényerejű rendeletnél) alacsonyabb szintű jogszabály is elrendelhetett. Kedvező viszont, hogy az 1973. évi statisztikai törvényben történő szabályozás legalább tiszta helyzetet teremtett azáltal, hogy a korábbi statisztikai törvényekből visszahozták azokat a rendelkezéseket is, amelyek az adatközlés feltételeiről szóltak. Külön rendelkeztek a nem természetes személyek egyedi adatainak közléséről. (14. §) Ez a rendelkezés is előrelépés az 1952. évi VI. törvényhez képest, mert legalább törvényi szinten szabályozta ezt a problémát. Az más kérdés, hogy az ún. jogi személyeket érintő egyedi statisztikai adatok nyilvánosságát a nyugati demokráciákban, ebben az időben már sokkal pontosabban szabályozták. Magyarországon majd csak az 1993. évi XLVI. törvény fog erről a témáról korszerű módon rendelkezni.

Az 1973-as törvény megalkotásakor szabályozták azt is, hogy magánszemélyek egyedi adatait csak statisztikai célra szabad felhasználni (15. §), ami azt jelentette, hogy a jogszabályba beépítették a „statisztikai cél” fogalmát, amelyet úgy lehet értelmezni, hogy a statisztikai adatfelvételek során a magánszemélyeket azonosító személyes adatokat nem lehet közzélni. (KSH [1994c] 72. old.) Tehát az 1973. évi V. törvény az 1952. évi VI. törvényhez képest korszerűbb módon fogalmazta meg a személyes adatok védelmével kapcsolatos témakört.

2. A statisztikáról szóló 1993. évi XLVI. törvény elfogadása

A rendszerváltozás hatására alapvető változások következtek be a jogrendszerben, az új alkotmányra alapozva sorra születtek olyan törvények, amelyek átalakították az állami és a magánszféra intézményrendszerét, kialakították a piacgazdaságot és a jogállamot szabályozó jogi struktúrát. Ez az átalakulás igen nagy hatással volt a statisztikai szolgálatra, a statisztikai adatgyűjtés teljes folyamatára. Nyilvánvaló, hogy a statisztikai tevékenységnek és az e tevékenységet szabályozó jogi környezetnek összhangban kell lennie a társadalmi, gazdasági, politikai viszonyokkal. Ez az összhang a rendszerváltozás után fokozatosan épült ki.

Még mielőtt ismertetnénk a statisztikai törvény előkészítésének folyamatát, röviden be kell számolni a KSH-nak¹ egy szintén alapvető jelentőségű kezdeményezését.

¹ KSH: Központi Statisztikai Hivatal.

ről. Mint ismeretes a rendszerváltás előtt több évtizeden át a KSH-t jelölte ki az akkori kormányzat a számítástechnika alkalmazásáért, fejlesztéséért felelős intézménynek. Ez lehetett az oka annak, hogy a KSH-ban dolgozó szakemberek (például adatvédelmi szakemberként hírnevet szerzett *Könyves Tóth Pál*) jobban érzékelték azt az egyre nagyobb veszélyt, amely az informatika fejlődése révén a számítástechnika által tárolt adatok alanyaira várt, tekintettel arra, hogy azokban az években nem volt olyan jogszabály, amely az állami szervek által gyűjtött és tárolt személyes adatok kezelését kellő szakszerűséggel szabályozta volna. Jellemző a korabeli helyzetre, hogy kissé rejtve kellett a jogszabály tervezetén dolgozni, információs (informatikai) törvénytervezetnek nevezték az anyagot, mert még nem akarták teljesen felfedni azt a tény, hogy olyan törvényre van szükség, mely alkotmányos alapjogokat szabályoz. Végül a rendszerváltást megelőző évben a Minisztertanács 3022/1989. határozata a személyes adatok kezeléséről és a közérdekű adatok nyilvánosságáról címmel a kormány 1989. január 26-án jóváhagyta a törvényjavaslat elveiről készített előterjesztést és annak az elveket részletező mellékletét, egyúttal elrendelte, hogy a törvénytervezetet a KSH elnöke és az igazságügy-miniszter 1990. december 31-ig terjessze a kormány elé. Időközben, a szabad országgyűlési választásokat követően, megalakult az új kormány és a törvénytervezet elkészítése immáron a KSH nélkül, az Igazságügyi Minisztériumban folyt tovább.

Visszatérve az új statisztikai törvény előkészületeinek munkafázisaira, 1990. év tavaszán kormányrendelet (78/1990. (IV. 25.)) született arról, hogy ki kell dolgozni az új statisztikai törvény főbb elveit, téziseit. A KSH-ban kidolgozott elveket, téziseket széles körű egyeztetéseken (szakértői testületek, tárcaértekezletek) tárgyalták meg, miközben megalakult az új kormány, a személyi és jogszabályi környezet is jelentősen megváltozott. A kormány 1991. február 28.-i ülésén elfogadta a leendő statisztikai törvény főbb elveit, szerkezetét, és ezzel egy hosszú törvényalkotási folyamat vette kezdetét. A KSH-ban elkészítették a törvénytervezet normaszövegét, melyet széles körű egyeztetés után, 1992. január 15-én a kormány benyújtott a Parlamentnek, és csaknem ugyanazon a napon az Országgyűlés elé terjesztette a személyes adatok védelméről és a közérdekű adatok nyilvánosságáról szóló (adatvédelmi) törvénytervezetet. A KSH azt javasolta, hogy a két törvényjavaslatot együtt tárgyalja meg az Országgyűlés, tekintettel arra, hogy a két törvény szorosan összefügg egymással. Sajnos ennek a javaslatnak a végrehajtása nem sikerült, a megtárgyalandó törvényjavaslatokkal nagymértékben megterhelt parlamenti jogalkotási munka miatt az adatvédelmi törvény tervezetének tárgyalását előbbre hozták. Ennek valószínű oka az is lehetett, hogy politikai szempontból az adatvédelmi törvény elfogadását a kormány és az ellenzék fontos kérdésnek tartotta, tekintettel arra, hogy ez a törvény az ún. kétharmados törvények sorába tartozott, a statisztikai törvény pedig ilyen minősítést nem kapott. Ez a tény a későbbiekben több problémát okozott a két törvény utóéletében, ugyanis a kétharmados törvény megváltoztatása mindig nagyobb nehézség-

gel jár, és így jogtechnikai szempontból az esetek többségében, a kívánatos változások során, általában a statisztikai törvénynek kellett igazodnia az adatvédelmi törvényhez. Ez az igazodás már a statisztikai törvény elfogadásakor is megvalósult.

A statisztikai törvénytervezet benyújtása a Parlamentnek (1992. január 15-én) és általános vitára becsatolása (1993. február 17-én) közötti egy év történéseit nem részletezem, csupán megjegyzem, hogy ezen interregnum alatt a KSH vezetésének a statisztikai munka folyamatosságának biztosítása érdekében néhány ideiglenes jogi megoldást is kellett alkalmaznia (például kormányrendelettel kellett életbe léptetni az 1993. évi OSAP²-ot). Nem részletezve a törvénytervezet vitáját, csak jelzem, hogy a képviselők igen aktívak voltak: 126 önálló képviselői indítványt kellett feldolgozni.

A képviselői indítványok a statisztikai adatgyűjtés szabályozására és az adatvédelem rendelkezéseivel való összehangolására irányultak, egymásnak ellentmondó indítványok is érkeztek, és voltak olyan javaslatok is, melyek a normaszöveg jobb megértését szolgálták. Az Országgyűlés főleg azokat a képviselői indítványokat támogatta, melyek az adatvédelmi törvényben foglaltakra tekintettel módosították, illetve kiegészítették a törvényjavaslat szövegét. A statisztikai törvény legitimitását nagyban elősegítette, hogy a Parlament 209 igen, 2 nem szavazattal és 3 tartózkodással 1993. április 27-én elfogadta az 1993. évi XLVI. törvényt (továbbiakban: statisztikai törvény – stt.)³ (*Bánszegi* [1993] 745. old.).

3. Az adatvédelem jelenleg hatályos jogi szabályozása

Természetesen az elfogadás óta többször módosították, de ezek a változtatások a jogszabály lényeges elemeit nem érintették. A jelenlegi változat összhangban áll a 223/2009/EK európai parlamenti és tanácsi rendelettel az európai statisztikáról, valamint az 557/2013. EU-bizottsági rendelettel a bizalmas adatokhoz való tudományos (kutatási) célú hozzáférés szabályairól. Meg kell még említeni az ESR⁴ működésére vonatkozóan az Európai Statisztika Gyakorlati Kódexét, melyet az Európai Unió Statisztikai Programbizottsága 2005. február 24-én fogadott el, és amely az óta (2011-ben) egy alapos revízió is átesett. A Gyakorlati Kódex 5. elve szól a statisztikai adatok bizalmas kezeléséről.

A statisztikai törvény szerkezete adatvédelmi szempontból logikus és áttekinthető. Egyes részei a statisztikai adatok nyilvánosságáról, mások az egyedi adatok védelméről

² OSAP: Országos Statisztikai Adatgyűjtési Program.

³ http://net.jogtar.hu/jr/gen/hjegy_doc.cgi?docid=99300046.TV

⁴ ESR: Európai Statisztikai Rendszer.

és annak biztosítékairól szólnak. A jelen dolgozatban főleg ezekkel a részekkel foglalkozunk oly módon, hogy figyelembe vesszük a törvény végrehajtási rendeletének vonatkozó normaszövegét, sőt – a gyakorlati problémák bemutatása céljából – a KSH 2011-ben alakult Adatvédelmi Bizottságának állásfoglalásaira is kitékintünk.

3.1. A statisztika feladata és célja

Egy statisztikai törvény alapvető feladata, hogy meghatározza a statisztikai cél fogalmát. Erre azért van szükség, mert számos esetben más a jogi helyzete egy statisztikai, mint egy nem statisztikai, például igazgatási célú adatgyűjtésnek, és mert a személyes adatok kezelésének vonatkozásában az adatvédelmi törvény egyik alaptétele a célhoz kötöttség, ami megköveteli a statisztikai cél megfogalmazását.

„Stt. 1. § 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. A cél elérése érdekében e törvény az adatok statisztikai módszerekkel történő felvételét, feldolgozását, tárolását, átadását, átvételét, elemzését, szolgáltatását, közlését, valamint közzétételét (a továbbiakban: statisztikai tevékenység) szabályozza.”

A statisztikai cél részletes meghatározása a szakirodalomban foglaltak legteljesebb leszűkítésével is oldalakat venne igénybe. A lényegét már *Fekete Gyula* akkori MDF⁵-képviselő, sok éven át KSH munkatárs is, az országgyűlési vitában megfogalmazta, miszerint „a statisztika a tömegesen előforduló jelenségek, számadatok révén történő leírása és elemzése.” (*KSH* [1994c] 123. old.) Fontos a normaszöveg azon része, mely a statisztikai adatelőállítási folyamat összes fázisát felsorolja, melyet adatvédelmi szempontból is figyelembe kell venni, ugyanis a statisztikai adatok védelmét a statisztikai tevékenység összes felsorolt fázisában is biztosítani kell.

A statisztikai törvény előkészítése kapcsán már említettük, hogy milyen szoros összefüggés volt az adatvédelmi törvénnyel, mely abban is kifejezésre jutott, hogy a statisztikai törvény 1. §-a az adatvédelmi törvényre való utalással megadja az adatvédelemmel kapcsolatos alaphangot, azaz nyomatékosan jelzi, hogy a törvény rendelkezéseinek meg kell felelniük az adatvédelmi törvény szabályainak.

A 2011. év tavaszán elfogadott, a korábbi alkotmány helyébe lépő Alaptörvény részben megváltoztatta a később meghozandó új adatvédelmi törvény (2011. évi CXII. törvény – az információs önrendelkezési jogról és információszabadságról,

⁵ MDF: Magyar Demokrata Fórum.

továbbiakban: Infotv.⁶) státusát, már nem tekintette a megszülető új adatvédelmi törvény egészét ún. „sarkalatos” (kétharmados) törvénynek, hanem csak egy részét, mely az új független hatóságra (NAIH⁷) vonatkozott. Ezzel a megoldással az Infotv. lényegében ún. „feles törvénnyé” vált, elvesztette kódexjellegét, mely szerint az Infotv. meghatározza a témával kapcsolatos általános szabályokat, az ún. szektorális törvények pedig az adott területre vonatkozó konkrét adatvédelmi rendelkezéseket foglalják magukban, utalva az adatvédelmi törvényre. Ez az oka annak, hogy a statisztikai törvény 1. §-a úgy módosult, hogy kivették az utalást az Infotv.-re. (Az Infotv. a jelentős változások ellenére a korábbi adatvédelmi törvénnyel megtartotta a folytonosságot, amit az is bizonyít, hogy a személyes adatok védelmével és a közérdekű adatok nyilvánosságával kapcsolatos rendelkezések jelentős részét szinte szó szerint átvette, tehát a korábbi törvény leglényegesebb elemei az Infotv.-ben is jelenkeznek.) (Lakatos–Nagy [2012] 39. old.)

– A statisztikai cél kizárólagossága szempontjából az Infotv. 13. §-ának (2) bekezdése egy fontos rendelkezést tartalmaz:

„Infotv. 12. Személyes adatok felhasználása statisztikai célra

13. § (2) A statisztikai célra felvett, átvett vagy feldolgozott személyes adatok – ha törvény eltérően nem rendelkezik – csak statisztikai célra kezelhetők. A személyes adatok statisztikai célra történő kezelésének részletes szabályait külön törvény határozza meg.”

A 13. § (2) bekezdése szerint törvény rendelkezhet arról, hogy statisztikai adatokat más célra is fel lehessen használni. Ez a rendelkezés lehetővé teszi például a statisztikai adatok kutatási célú felhasználását (az európai uniós szabályokkal és a tudományos közösség igényeivel összhangban), ugyanakkor fokozott óvatosságot követel a jogalkotótól, hogy ne kerülhessen sor olyan törvények elfogadására, amelyek statisztikai adatok hatósági, ellenőrzési célú felhasználásáról rendelkeznek.

A statisztika és az adatvédelem kapcsolatát erősíti az Infotv. közérdekű adatokkal kapcsolatos fogalma, mely egyértelművé teszi, hogy valamennyi, a hivatalos statisztikai adatgyűjtésből származó adat főszabályként közérdekűnek minősül.

3.2. Fogalmak az Infotv. alapján

Mint említettük, az 1993. évben alkotott statisztikai törvény számos ponton kapcsolódott az 1992-ben elfogadott adatvédelmi törvényhez, különösen a fogalomrendszerre igaz ez. Az 1993. évi statisztikai törvény nem tartalmaz külön értelmezési

⁶ http://net.jogtar.hu/jr/gen/hjegy_doc.cgi?docid=A1100112.TV

⁷ NAIH: Nemzeti Adatvédelmi és Információszabadság Hatóság.

rendelkezéseket, azonban magában a jogszabályban számos ponton van ilyen jellegű meghatározás (például 8. §-ának (4), 17. §-ának (2) bekezdése). A statisztikai törvény szempontjából azért nagyon fontos az Infotv. megújított 3. §-a, mert számos értelmező rendelkezés a statisztikai törvény adatvédelmet érintő alkalmazása szempontjából is jól hasznosítható. Terjedelmi korlátok miatt most csak néhány fontosabb fogalomra utalunk.

Az Infotv. követi az információs jogok magyarországi szabályozásában megszilárdult fogalomrendszert. A meghatározások többségét változtatás nélkül átveszi a korábbi adatvédelmi törvényből, de néhány ponton eltér a korábbiakban adott egyes fogalom-meghatározásoktól, sőt néhány új fogalmat is bevezet, ezek azonban a lényeges pontokban nem változtatják meg a fogalomrendszert.

„Infotv. 3. § E törvény alkalmazása során:

1. érintett: bármely meghatározott, személyes adat alapján azonosított vagy – közvetlenül vagy közvetve – azonosítható természetes személy;

2. személyes adat: az érintettel kapcsolatba hozható adat – különösen az érintett neve, azonosító jele, valamint egy vagy több fizikai, fiziológiai, mentális, gazdasági, kulturális vagy szociális azonosságára jellemző ismeret –, valamint az adatból levonható, az érintettre vonatkozó következtetés.”

A „személyes adat” fogalma lényegében megegyezik a korábbi törvényben megfogalmazott tartalommal. Az érintettre vonatkozó fogalom kifejtésének szinte minden szava lényegi elemet hordoz, közülük ezúttal csak az „azonosított” és az „azonosíthatóság” meghatározására utalunk. A kialakult vélemény szerint valaki akkor „azonosítható”, ha elég információ áll rendelkezésre elkülönült létezésének, egyénként való léte tényének tükrözéséhez, és akkor válik „azonosítottá”, ha ugyancsak elég információ áll rendelkezésre másoktól valamilyen módon történő megkülönböztetéséhez, felismeréséhez. Lényeges fogalom ezen belül a „közvetlen”, illetve „közvetett” azonosíthatóság. A „közvetlen” azonosíthatóságon az érintett neve, címe és a hozzákapcsolt azonosítói jel értendő, ezen felül minden más olyan ismeret, mely az érintettet azonosíthatja, az „közvetett azonosítást” jelent. A statisztika szempontjából is alapvető fogalomról van szó, amely az azonosíthatóságának ezt a megközelítését nemcsak a természetes személyekre, hanem a jogi személyekre is – és általánosságban fogalmazva, minden statisztikai egységre – alkalmazza.

A statisztikai információrendszer működtetése során, az elmúlt közel húsz évben, sokszor felmerült az a kérdés, hogy az érintett közvetett azonosíthatóságának hol van a határa, mikor lehet azt mondani, hogy bizonyos adatokból felismerhetők az érintettek. Ennek eldöntéséhez eddig támpontot adott az Európa Tanács Adatvédelmi Egyezménye.

nyének indoklása, nevezetesen az, hogy „az azonosítható személy olyan személyt jelent, aki könnyen azonosítható, nem foglalja magába azokat, akiket csak nagyon bonyolult módszerrel lehet azonosítani”. (*InfoFilia* [1991–1993] I. kötet 3.40. old.) Az Infotv. ezt az értelmezési kérdést próbálja szabályozni, amikor a 4. § (3) bekezdésében kimondja, hogy az érintettel akkor helyreállítható a kapcsolat, ha az adatkezelő rendelkezik azokkal a technikai feltételekkel, amelyek a helyreállításhoz szükségesek. Ez az új jogszabályhely megkönnyíti a jogalkalmazók dolgát, hiszen alátámasztja azt a gyakorlatot, amely szerint minden egyes esetben, minden adatkérő, -kezelő esetében külön-külön kell vizsgálni az azonosíthatóság kérdését, és így a jogszabály alkalmazása rugalmasabbá válik. A statisztika területén már jelenleg is ezt a megközelítést alkalmazzuk, így igen szerencsés, hogy azt a jelen jogszabály is rögzíti.

3.3. A statisztikai törvény hatálya

Az Infotv.-vel való kapcsolódás miatt van jelentősége az Stt. hatályának

„Stt. 2. § A törvény hatálya kiterjed

- a) az 1. §-ban meghatározott statisztikai tevékenységet folytató természetes és jogi személyekre, valamint a jogi személyiséggel nem rendelkező szervezetekre;
- b) a nem statisztikai célú adatgyűjtéseket végző és adatokat kezelő szervekre és szervezetekre, a statisztikai célra is felhasználásra kerülő adataik vonatkozásában.”

A statisztikai törvény hatályossági köre tehát kiterjed többek között a Hivatalos Statisztikai Szolgálathoz tartozó szervek statisztikai tevékenységére, a szolgálathoz nem tartozó szervek adatgyűjtésére, adatszolgáltatására (társadalombiztosítás, cégnyilvántartás, ingatlan-nyilvántartás, adó stb.), ha azok adataikat statisztikai célra használják. Már e szakasznál (2. § b) pont) fontos lehatárolási ismérv a statisztikai cél fogalma. A statisztikai törvény 2. §-val kapcsolatban ismét megjegyezzük, hogy míg az adatvédelmi törvény hatálya csak a természetes személyekre terjed ki, addig a statisztikai törvény hatálya szélesebb körre, a nem természetes személyekre is kiterjed.

3.4. A statisztikai célú adatkezelés elkülönülése a Hivatalos Statisztikai Szolgálat intézményeinél, további adatvédelmi garanciák

Az 1993. évi statisztikai törvény eredetileg csak a Központi Statisztikai Hivatalról mondta ki a szakmai önállóságot, ezért a nemzetközi elvárásoknak megfelelően, ki kellett terjeszteni ezt minden Hivatalos Statisztikai Szolgálathoz tartozó szervekre,

kimondva e szervek kötelezettségét, hogy az igazgatási célú információs rendszereiktől függetlenül működtessék statisztikai rendszerüket:

„Stt. 3/A. § A Hivatalos Statisztikai Szolgálathoz tartozó szerv vezetőjének biztosítani kell a statisztikai tevékenység működésének önállóságát és a statisztikai tájékoztatás függetlenségét.”

Ez a rendelkezés biztosítja azt, hogy a Hivatalos Statisztikai Szolgálat tagjai elkülönítetten kezeljék a statisztikai céllal felvett adatokat. Adatvédelmi szempontból is nagyon fontos rendelkezésről van szó, mivel az adatszolgáltatónak tudnia kell, hogy ha például az adóhatóság felügyeletét ellátó minisztérium végez statisztikai célú adatfelvételeket, akkor a felvett adatokat nem használják fel adminisztratív, netán adózási célokra.

3.5. Statisztikai adatgyűjtések adatvédelmi vonatkozásai

A statisztikai törvény 8. § (1) bekezdése az adatszolgáltatás szempontjából az adatgyűjtéseket kötelező vagy önkéntes voltak alapján határolja el egymástól:

„Stt. 8. § (1) Statisztikai adatgyűjtéseket *a*) adatszolgáltatási kötelezettség előírásával, vagy *b*) önkéntes adatszolgáltatás alapján lehet végrehajtani.”

Önkéntes adatszolgáltatáson alapuló adatgyűjtéseket bárki, bármikor szervezhet. Ezzel kapcsolatban csak egyetlen előírás: az adatszolgáltató figyelmét az adatgyűjtő köteles felhívni arra, hogy az adatszolgáltatás önkéntes és annak nem teljesítése semmiféle hátrányos következménnyel nem jár. (16. §)

Az Infotv. fogalomrendszerével összhangban a statisztikai törvény is rendelkezik a különleges adatokról (8. § (4)) mégpedig oly módon, hogy ezek az adatok csak személyazonosításra alkalmatlan módon gyűjthetők a 8. § (4) *a*) pontban felsorolt adatfajták (például az érintett faji eredetére) csak önkéntesen, a *b*) pontban felsoroltaké (például egészségi állapotra) ezen kívül törvény által elrendeltek alapján is.

Ezzel is kapcsolatban a statisztikai törvénynek adatvédelmi garancia szempontjából is fontos szakaszának 8. § (3) bekezdése:

„Stt. 8. § (3) Természetes személytől személyes adatára vonatkozó kötelező adatszolgáltatást csak törvény rendelhet el.”

Erről a kérdésről az 1929. évi XIX. tv. 14. §-a is rendelkezik, csak nem ilyen egyértelmű módon. (Az éves munkaterv vonatkozásában határozták meg azt, hogy a magánszemély adatszolgáltatási kötelezettségét hogyan lehet elrendelni.)

A statisztikai törvénynek ez a rendelkezése is indokolta, hogy a KSH hagyományos adatgyűjtései közé tartozó népmozgalmi adatok gyűjtését a 10. §-ban részletesen szabályozza oly módon, hogy definiálja a népmozgalmi eseményt (születés, házasságkötés, válás, halálozás; (10. § (1)), majd felsorolja azokat az ismérveket, amelyek a statisztikai felmérés céljából szükségesek (10. § (2)). (Például: név, lakcím, nem, foglalkozás, munkahely, népmozgalmi eseménnyel összefüggő egészségi állapot stb.). A 10. § (3) szakasza tartalmazza azt a garanciális szabályt, hogy:

„Stt. 10. § (3) A népmozgalmi eseménnyel kapcsolatban gyűjtött adatok közül a nevet, a lakcímet és a tartózkodási helyet az adatok teljessége és összefüggése ellenőrzésének befejezését követő nyolc napon belül törölni kell.”

A 10. § (4) pedig rendelkezik arról, hogy a népmozgalmi eseményekkel kapcsolatos adatokat mely szervek szolgáltatassák (például anyakönyvezetők, bíróságok).

Míg a természetes személyekkel kapcsolatos adatgyűjtést csak törvény rendelhet el, addig a nem természetes személyekkel kapcsolatos adatgyűjtést kormányrendelet kötelező módon előírhat. A statisztikai törvény 8. § (6.) bekezdése erről szól, azzal, hogy a végrehajtási rendelet 6. §-a további részletszabályokat is felsorol. (Például a statisztikai adatgyűjtés kérdőívei milyen garanciális tájékoztatási információkat tartalmazzanak.)

Az 1929. évi XIX. tv. 15. §-a rendelkezik arról, hogy az adatszolgáltatási kötelezettség nem akármilyen adatok szolgáltatására, hanem a valósághoz hű információk adására vonatkozik. (Amint jeleztük, ez az egyik ok, mely oly fontossá teszi az adatszolgáltatók bizalmának elnyerését.) A jelenleg hatályos statisztikai törvény 9. § (1) bekezdése a következőképpen szabályozza ezt a témát:

„Stt. 9. § (1) Az adatszolgáltatásra kötelezett az előírt adatokat a valóságnak megfelelő tartalommal, megadott határidőben és meghatározott módon, térítésmentesen köteles szolgáltatni.”

3.6. A statisztikai adatok nyilvánossága, egyedi adatok védelme.

A statisztikai törvény 17–21. §-ai szabályozzák a statisztikai adatok nyilvánosságát, védelmét és továbbítását. A statisztikai törvénynek ez a része kapcsolódik legérősebben az Infotv.-hez, ezért e témakörrel részletesebben kell foglalkoznunk.

Az Infotv. megadta a főbb irányokat a személyes adatok védelme és a közérdekű adatok nyilvánossága tárgyában. A statisztikai törvénynek, már csak időbeni eltolódása miatt is, a korábban megalkotott adatvédelmi törvényhez és az egyedi adatok hagyo-

mányos védelméhez igazodva kellett szabályozni a tájékoztatási rendszert. Ennek megfelelően a statisztikai törvény 17. §-a a statisztikai információrendszer szempontjából alapvető kérdést szabályoz, melynek első szakasza főszabályként leszögezi:

„Stt. 17. § (1) A Hivatalos Statisztikai Szolgálathoz tartozó szervek által végrehajtott adatgyűjtések eredményei — a (2) bekezdésben foglalt kivétellel — nyilvánosak. A nyilvánosságra hozásról e szervek saját hatáskörükben gondoskodnak.

(2) Nem lehet nyilvánosságra hozni a minősített adatokat, valamint — a 18. §-ban foglaltak kivételével — a statisztikai célt szolgáló, a természetes és a jogi személy, valamint a jogi személyiséggel nem rendelkező adatszolgáltatóval kapcsolatba hozható adatot (a továbbiakban: egyedi adat).”

Erre ráerősít a statisztikai törvény 20. §-a:

„Stt. 20. § Az egyedi adatok más jogszabályok alkalmazása szempontjából magántitoknak minősülnek. Az ezek védelmére vonatkozó szabályok betartásáért a statisztikai tevékenységet végző, valamint az abban közreműködő személyek felelősek.”

A magántitkokra vonatkozó szabályok megszegéséért az elkövető fegyelmi, polgári és bizonyos esetekben büntetőjogi felelősséggel tartozik. (E rendelkezés is bizonyítja, hogy milyen fontos a statisztikai tevékenységet végzők alapvető jogi felkészítése, nevezetesen, hogy tisztában legyenek jogaikkal és kötelezettségeikkel.)

Mint jeleztük, a statisztikai törvény nem sorol fel értelmező rendelkezéseket, azonban néhány helyen tartalmaz ilyen jellegű szöveget. A 17. § (2) szakasza lényegében megadja az egyedi adat fogalmát.

Lényeges a fogalom pontos értelmezése, mert számos jogkövetkezménye van annak, hogy a statisztikai céllal kezelt adat egyedinek minősül-e vagy sem. E fogalom értelmezése szorosan kapcsolódik az Infotv.-ben használt személyes adat fogalmához. Míg a személyes adat csak a természetes személyekre vonatkozik, addig az egyedi adat ezen túlmenően a nem természetes személyek adatait is magába foglalja. (E megkülönböztetésre azért is szükség van, mert az adatvédelmi törvény hatálya csak a természetes személyekre terjed ki.) E fontos fogalom gyakorlati alkalmazását segítik a KSH adatvédelmi szabályzatának 3 §-ban közölt értelmező rendelkezései, például a felfedés, a felfedési kockázat, a közvetlen azonosítás, az azonosító vagy a közvetett azonosítás fogalmai.

A KSH adatvédelmi szabályzata (KSH [2013]) igyekszik segítséget adni a statisztikai törvény 17. § (2) szakaszának értelmezéséhez. Az Infotv. fogalomrendszerére épít-

ve a kiterjesztett tulajdonságú egyedi adat meghatározását több fogalom meghatározásával járja körül. Különösen fontosak a közvetlen, közvetett azonosítással kapcsolatos fogalmak. Az Infotv.-nél alkalmazott „kapcsolatba hozható adat” „az adatokból levonható, az érintettre vonatkozó következtetés” szöveg értelmezése jelent a gyakorlati alkalmazás során kihívást. Különösen azért, mert nemcsak a természetes személyek, hanem a nem természetes személyek azonosíthatóságáról is szó van.

Az egyedi adatok védelmének egy garanciális szabályát tartalmazza a statisztikai törvény végrehajtási rendeletének 19. §-a:

„Stt. Vhr. 19. § Összesítve sem lehet nyilvánosságra hozni — az stt. 18. §-ának (2) bekezdésében és a 16. §-ban foglaltak kivételével — olyan adatot, amelynél az adatszolgáltatók száma háromnál kevesebb.”

A Vhr. 19. §-a az eddig ismertetett összes garanciákon túl kimondja, hogy összesítésben sem lehet nyilvánosságra hozni – a közérdekű adat és az adatszolgáltató hozzájárulása kivételével – olyan adatot, amelynek megszerzésénél az adatszolgáltatók száma háromnál kevesebb. Ez az előírás azt a célt szolgálja, hogy ne lehessen következtetni az adatszolgáltatóra és így kikövetkeztetni a személyére vonatkozó adatokat, mert ez, különösen a gazdasági élet területén, az üzleti titok megsértéséhez vezethetne. (Amely adott esetben kártérítési kötelezettséget vonhat maga után.).

A Vhr. e rendelkezése a gyakorlati alkalmazás során szintén kihívást jelent. Többek között kiemelendő a tudományos szféra utóbbi években egyre intenzívebben jelentkező igénye a statisztikai mikroadathoz való tudományos célú hozzáférésre. Amennyiben a statisztikai törvény e rendelkezését szó szerint értelmeznénk, úgy jelenleg nem nyílna lehetőség a statisztikai mikroadatokhoz való tudományos célú hozzáférésre. Így a jelenlegi statisztikai törvény a mikroadat szintű tudományos célú hozzáférés tekintetében is jelentős átalakításra, modernizálásra szorul.

3.7. Egyedi adatok nyilvánosságra hozatalának esetei

Még a statisztikai célú adatfelvételeknél is lehetnek olyan helyzetek, mikor kivételt kell tenni az egyedi adatok nyilvánosságra hozásával kapcsolatban. Ezt szabályozza a statisztikai törvény 18. §-a.

3.7.1. Írásbeli hozzájárulás

„Stt. 18. § (1) Egyedi adat csak statisztikai célra használható, mással csak akkor közölhető, és abban az esetben adható át, valamint hozható

nyilvánosságra, ha ehhez az adatszolgáltató előzetesen írásban hozzájárult. Ez a korlátozás nem vonatkozik az azonos szervek belül statisztikai tevékenységet végző személyek egymás közötti adatközlésére.”

A 18. § (1) szakasza – az Infotv.-hez hasonlóan – írásbeli hozzájárulás esetén teszi megengedhetővé az egyedi adatok közlését. Garanciális elem, hogy a statisztikai törvény végrehajtási rendeletének 16. §-a részletesen szabályozza, hogy az adatszolgáltató írásbeli hozzájárulásának milyen fontos kritériumai vannak, például a hozzájárulás csak meghatározott célra és időtartamra adható. A hozzájárulás tehát akkor tekinthető szabályosnak, ha az adatszolgáltató önkéntesen, megfelelő tájékoztatását követően, határozottan és félreérthetetlenül adja beleegyezését egyedi adatainak felhasználásához.

A statisztikai célú felvételek esetében viszonylag gyakran előfordul, hogy az adatszolgáltató valamely, saját szempontjából fontos okból, például adategyeztetés miatt, esetleg évek múltán kikéri azokat az információkat (kérdőíveket), melyeket statisztikai adatszolgáltatás során a KSH-nak eljuttatott. A KSH az adatszolgáltatás során ezeket az információkat ellenőrzi, feldolgozza, minőségileg javítja; így előfordulhat, hogy az eredetileg beküldött információ módosult. Ezért a KSH adatvédelmi szabályzatának 7. § (3) b) pontja úgy rendelkezik, hogy csak az eredeti, az adatszolgáltató által beküldött (kérdőíven rögzített) információ szolgáltatható vissza az adatkérő számára. Amennyiben a kért információ a beküldéskori állapotában már nem áll rendelkezésre, úgy a kért információt a KSH nem tudja visszajuttatni.

3.7.2. Hivatalos Statisztikai Szolgálat szervein belüli adatáramlás

A statisztikai törvény 18. § (1) szakasza szól még arról is, hogy az azonos statisztikai szervek, például a KSH-n belül, az egyedi adatokhoz a statisztikai tevékenységet végző személyek hozzáférhetnek. Ez logikus szabály, hiszen a statisztikai adatok felvétele, feldolgozása során, több munkafolyamat végzése feltételezi az egyedi adatok ismeretét. Azonban olyan nagy létszámú szervezet esetében, mint a KSH ezt a lehetőséget is érdemes szabályozni, kontroll alatt tartani. A KSH Adatvédelmi Bizottsága több állásfoglalásában is kifejtette, hogy ebben az esetben is érvényes a célhoz kötöttség elve, tehát csak határozott feladat elvégzéséhez lehet a munkatársak részére hozzáférést biztosítani. Ezen felül ezt a hozzáférést naplózni, illetve dokumentálni kell. Valamennyire e témához kapcsolódik, hogy a jogalkotó – szintén a statisztikai munka hatékony végzése érdekében – engedélyezi azt, hogy – ezúttal csak a Hivatalos Statisztikai Szolgálat tagjaira vonatkozóan –, a szervek, így a KSH is, a saját statisztikai adatállományait statisztikai célra összekapcsolhatják (21. § (5)).

3.7.3. Tudományos (kutatási) célú hozzáférés

Csak részben tartozik az egyedi adatok megismerhetősége körébe az az eset, amikor egy kutató ellenőrzött körülmények között, tudományos célból hozzáférhet a közvetlen azonosítóktól megfosztott, közvetetten azonosítható statisztikai célból felvett adatokhoz (különösképpen mikroadat-állományokhoz). A statisztikai mikroadatokhoz való tudományos (kutatási) célú hozzáférés keretében az egész uniós szabályozásban kiemelten kezelt kérdéséről van szó. Itt is tetten érhető az az egyensúlykeresés, mely a statisztikai adatok megismerhetőségéről, illetve védelméről szól. Személyes adatok vonatkozásában itt két alkotmányos alapjog ütközéséről van szó, a kutatás szabadságának és a személyes adatok védelmének kívánalma feszül egymásnak. A statisztikai mikroadat-állományok igen sokrétű információtartalommal bírnak, ezeknek csak egy részük kerül nyilvánosságra – jellemzően aggregált módon –, melyek a részletes, jó minőségű kutatásra vonatkozó igényeknek nem felelnek meg.

A statisztikai törvény nagy hiányossága, hogy közvetlenül nem szól a statisztikai adatokhoz való tudományos célú hozzáférés kérdéseiről, csak közvetetten az 557/2013/EK⁸ bizottsági rendelettel összeolvasva lehet azt a kérdést megoldani. A lényeg az, hogy a KSH, mint a statisztikai adatok meghatározó kezelője ún. biztonságos környezetben (kutatószobában, távoli hozzáférés, illetve távoli végrehajtás formájában), akkreditált kutatók részére biztosítja tudományos célból a statisztikai adatok kutatásának lehetőségét. Terjedelmi okokból nem ismertetjük a teljes folyamatot, ezt egyébként a KSH belső adatvédelmi, adathozzáférési szabályzatai megteszik. Tudományos célú hozzáférés esetén a kutatás eredményét a KSH szakemberei annak kiadása előtt felfedés elleni védelmi szempontból ellenőrzik (ezt a folyamatot hívja a nemzetközi szakirodalom ún. output checking-nek), hogy a kutatási eredmény egyedi adatot ne tartalmazzon, illetve annak alapján ne nyíljon lehetőség egyedekre vonatkozó új, eddig nem ismert információ megismerésére.

3.7.4. Közérdekű feladatokat ellátó szerv adatainak nyilvánossága

„Stt. 18. § (2) Közérdekű feladatot ellátó szerv, illetőleg egyesület, valamint a költségvetési szerv ezen tevékenységére vonatkozó egyedi adat az adatszolgáltató írásbeli hozzájárulása nélkül is nyilvánosságra hozható.”

A rendszerváltást követő években jelentős igény mutatkozott arra, hogy az állam működése legyen átlátható, ott ahol közpénzt költenek, legyen nyilvános az adott

⁸ EK: Európai Közösség.

szervezet működése. A statisztikai törvény kodifikátorai is abból indultak ki, hogy a statisztikai célból gyűjtött adatok esetében, ha az adat közérdekű feladatot ellátó szervtől származik, akkor a tevékenységével kapcsolatban lehessen egyedi információhoz jutni.

A KSH Adatvédelmi Bizottságához is több olyan állásfoglalás-kérés érkezett, melyben az adatfelhasználók e törvényi szakaszra hivatkozva igényeltek egyedi információkat. Az adatvédelmi bizottság véleménye szerint e jogszabályhely alapján kizárólag az állami forrásból finanszírozott tevékenységre vonatkozó adatok adhatók ki. Amennyiben ezek nem különíthetők el a magánforrásból támogatott tevékenység adataitól, úgy az egyedi adatokra vonatkozó adatkérés nem teljesíthető. Az elmúlt hosszú időszak tapasztalatai jelzik, hogy e törvényi szakasz normaszövege felülvizsgálatra szorul, tekintettel arra, hogy még a „közérdekű feladatot ellátó szerv” fogalmának jogszabályi definícióját sem sikerült megadni.

3.7.5. Nemzetközi szervezet részére történő adatszolgáltatás

Az uniós csatlakozással kapcsolatban a statisztikai törvényt úgy kellett módosítani, hogy egyedi adat statisztikai célból jogszabályban meghatározott nemzetközi kötelezettség teljesítése érdekében átadhatóvá váljon a statisztikával foglalkozó európai uniós szervezetek (kiemelten az Európai Bizottság (Eurostat)) részére.

„Stt. 18. § (3) Egyedi adat – az (1) bekezdésben foglaltakon túlmenően – csak statisztikai célból, statisztikai tevékenységgel foglalkozó nemzetközi szervezetnek, jogszabályban meghatározott nemzetközi kötelezettség teljesítése érdekében adható át az adatszolgáltató írásbeli hozzájárulása nélkül.”

Fontos megjegyezni, hogy csak abban a körben kell az egyedi adatok átadását megtenni, mely meghatározott nemzetközi statisztikai kötelezettségre vonatkozik. Ennek megfelelően az Európai Uniót kívül egyéb nemzetközi szervezet (például ENSZ⁹, OECD¹⁰ stb.) részére nincs egyedi statisztikai adatokra vonatkozó átadási kötelezettség.

⁹ ENSZ: Egyesült Nemzetek Szervezete.

¹⁰ OECD (Organisation for Economic Co-operation and Development): Gazdasági Együttműködési és Fejlesztési Szervezet.

3.7.6. Statisztikai regiszterekből származó adatok nyilvánossága

Az egyedi adatokkal nemcsak a statisztikai törvény 17. és 18. §-ai foglalkoznak, hanem a törvény más jogszabályhelyei között is találunk ilyen rendelkezéseket. A statisztikai törvény 1993. évi elfogadása után egyre nyilvánvalóbbá vált, hogy nagy igény van az adatszolgáltatók nyilvántartását tartalmazó, ún. statisztikai regisztertípusú adatok iránt, például a statisztikai adatgyűjtések esetén, ahol a mintavétel alapjai is lehetnek a statisztikai regiszteradatokból előállított ún. felvételi keretből származó információk.

„Stt. 6. § (2) A statisztikai regiszter az adatszolgáltatók nyilvántartása érdekében a jogi személyiségű gazdasági szervezet, a gazdasági tevékenységet (vállalkozást) folytató természetes személy és jogi személyiséggel nem rendelkező szervezet nevét, statisztikai azonosítóját, statisztikai csoportképző ismérveinek kódját, valamint az azok meghatározásához szükséges alapadatokat, székhelyét és levelezési címét, telefon- és telefax-számát, valamint telepének címét, tevékenységi körét, nyitvatartási jellegét (állandó jelleggel, idejéveljelleggel működő telep) tartalmazza.

(3) A statisztikai regiszter (2) bekezdés szerinti tartalma – a csoportképzés alapjául szolgáló alapadatok kivételével – nyilvános. A Hivatalos Statisztikai Szolgálathoz tartozó szervek részére – az Országos Statisztikai Adatgyűjtési Programban szereplő adatgyűjtéseik megszervezése érdekében – a regiszter tartalma az ehhez szükséges mértékben kiegészíthető.”

A KSH jelenleg több statisztikai regisztert is működtet, ezek közül az egyik legismertebb a GSZR¹¹, mely a statisztikai törvény 6. § (2) által meghatározott körben egyedi információkat szolgáltat. Az ún. statisztikai regisztertípusú adatok iránti igény kiegészül azzal, hogy például a gazdasági szervezetek adatai bizonyos – egyedileg nem azonosítható –, csoportképző ismérvek szerint is nyilvánosságra kerüljenek. Ilyen csoportképző ismérv például a létszám- és az árbevételi kategória. Az elmúlt években bizonytalanság volt abban, hogy ezek a kategóriák milyen részletezettségűek legyenek. Az adatkérők más és más kategóriák szerint kérték a statisztikai regiszterekből származó adatokat. A KSH Adatvédelmi Bizottsága végül úgy egységesítette az adatkiadási gyakorlatot, hogy állásfoglalásában kimondta, csak olyan csoportosításban lehet a statisztikai regiszterekből adatokat kiadni, melyeket a KSH a honlapján nyilvánosságra hoz. Így minden adatkérő csak ebben a csoportosí-

¹¹ GSZR: Gazdasági Szervezetek Regisztere.

tásban kaphatja meg a statisztikai regiszteradatokat, miáltal kivédhető, illetve csökkenthető a felfedés kockázata.

3.7.7. Passzív adatvédelem

Az Európai Statisztikai Rendszer tagjaként végzett statisztikai munka egyik fontos területe a külkereskedelmi statisztika. Az egyes termékek külkereskedelmi célú forgalmazásával kapcsolatos statisztikáknál gyakran előfordulhat az adatszolgáltató azonosíthatósága. Ugyanis egyes bontásokban a statisztikai egységek száma olyan kevés, hogy a Vhr.-ben lefektetett közlés szabályokba ütközne. Ez a probléma nem magyar sajátosság; több EU-tagállam is hasonló nehézségekkel nézett szembe. Ezért a csak erre a statisztikai területre a vonatkozó uniós jogszabályok (Az Európai Parlament és a Tanács 638/2004/EK és 471/2009/EK rendelet) megfordították azt az egyedi adatokkal kapcsolatos főszabályt, hogy az egyedi adatokat nem lehet nyilvánosságra hozni. Ez a gyakorlatban azt jelenti, hogy ha az adatszolgáltató úgy ítéli meg, hogy valamely termék exportjáról vagy importjáról közölt információ lehetővé teszi külkereskedelmi forgalmának azonosítását, kérheti adatának olyan védelmét, hogy az adatszolgáltató azonosítása ne legyen lehetséges. Ezt nevezi a szakirodalom ún. passzív adatvédelemnek. Itt tehát arról van szó, hogy alapszabályként a KSH az érintett adatszolgáltató közvetett beazonosíthatósága ellen az érintett adatállományokat nem védi (közvetlen azonosítás passzív adatvédelem esetén sem megengedett), csak akkor, amennyiben az adatszolgáltató írásban tiltakozott saját adatainak közvetett beazonosíthatóságú közzététele ellen.

3.7.8. A Magyar Nemzeti Bank részére történő adatátadás

Az Európai Unió működésének jellegzetessége, hogy az Európai Statisztikai Rendszeren kívül egy másik prominens statisztikai hálózata is van, mégpedig a KBER¹², melynek tagjai – jellemzően a pénzügyi statisztikák tekintetében – statisztikai jellegű munkákat is végeznek. Érthető módon a nemzeti statisztikai hivataloknak és a nemzeti bankoknak – statisztikai feladataik minél hatékonyabb ellátása érdekében – kapcsolatban kell lenniük egymással. Ezt szabályozza az MNB¹³-ről szóló 2013. évi LXXXIX. törvény is, melynek 30. §-a kimondja, hogy statisztikai célból a jegybanki információs rendszer jogosult a KSH-tól személyes adatnak nem minősülő egyedi adat átvételére. Ez lényegében azt jelenti, hogy itt jogi személyek – főleg a pénzügyekkel kapcsolatos – adatainak átadásáról van szó. Maga a jegybanktörvény, az MNB belső szabályzatai és a KSH-val kötött együttműködési megállapodások is

¹² KBER: Központi Bankok Európai Rendszere.

¹³ MNB: Magyar Nemzeti Bank.

számos garanciális szabályt tartalmaznak az egyedi adatok kezelésével kapcsolatban. A jegybanktörvény azt is kimondja, hogy az egyedi adatokat az MNB sem hozhatja nyilvánosságra. A két statisztikai rendszer kapcsolatából következik, hogy a jegybanktörvény azt is lehetővé teszi, hogy statisztikai célból az MNB adjon át egyedi adatokat a KSH-nak.

3.8. A természetes személyek adatkezelésére vonatkozó garanciák

Az országgyűlési vitákon egyes képviselők különösen fontosnak tartották, hogy a statisztikai célú adatgyűjtések esetében is maximálisan védve legyenek a természetes személyek adatai. Ezt szabályozza – a népmozgalmi adatgyűjtésekhez hasonlóan – a statisztikai törvény 19. §-a:

„Stt. 19. § (1) A természetes személy személyére vonatkozó adatgyűjtésnél az érintett nevét és a lakcímét (a továbbiakban: személyazonosító adat) – kivéve azt, amelynek adathordozóját a levéltári anyag védelmére vonatkozó jogszabály értelmében levéltári őrizetbe kell adni – a statisztikai feldolgozás befejezésekor, az adatok teljességének és összefüggésének ellenőrzését követően, de legkésőbb a tárgyidőszakot követő egy éven belül kell törölni, adatátadás esetén – a 21. § (4) bekezdésében foglalt kivétellel – ezt megelőzően is.”

Ez a megoldás a parlamenti bizottsági viták során merült fel, és átvéve egyes országok gyakorlatát a céllal került a törvénybe, hogy a személyes adatokat megvédjék a statisztikát felhasználók esetleges indiszkréciójától (gondolva az 1941. évi népszámlálás anyagának a kitelepítések végrehajtása céljából történő felhasználására).

A statisztikai információs rendszer tartópillérei közé tartoznak a visszatekintő adatok, az ún. idősorok. A 19. § (1) szakaszának önálló alkalmazása esetén lehetlenné válnának az ún. idősoros vizsgálatok. Ezért gondoskodni kellett olyan szabályozásról, amely ezeket a vizsgálatokat lehetővé teszi, nevezetesen, hogy az egy évnél hosszabb időszakra vonatkozó idősoros vizsgálatok esetében az adatállományt belső azonosítóval kell ellátni, amelyből az érintett személyazonossága nem állapítható meg. Az érintett személyazonosító adatait az adatállománytól elkülönítetten kell kezelni. A 19. § (2) szakaszát természetesen csakis olyan vizsgálatoknál lehet alkalmazni, amelyeknél csakugyan felállíthatók az ún. idősorok.

A 19. § (3) szakasza lehetővé teszi, hogy az új adatoknak az adatállományhoz csatolása idejére a személyazonosító adatok az adatállománnyal ideiglenesen összekapcsolhatók legyenek.

A 19. § (4) bekezdése arra a fontos kérdésre utal, hogy a lakosság teljes (például népszámlálás) vagy részleges körben (például általános mezőgazdasági összeírás)

összeírt adatait milyen jogszabályi rendelkezések alapján lehessen összekapcsolni. Itt statisztikai célú összekapcsolásról van szó, melynek eredményeként a külön-külön felvett adatok információtartalmát meg lehet sokszorozni.

3.9. Statisztikai adatok továbbítása

Főszabályként a statisztikai törvény 21. § (1) bekezdése leszögezi, hogy a Hivatalos Statisztikai Szolgálat tagjai egymásnak átadhatnak adatot, de ezt csak egyedi azonosításra alkalmatlan módon, célhoz kötötten tehetik meg, kivéve, ha a törvény másképpen nem rendelkezik. A KSH tekintetében a statisztikai törvény kivételt tesz, de csak egy irányban. A KSH átvehet egyedi adatot, de nem adhat át ilyen jellegű információt.

Ez a lehetőség a statisztikai információs rendszer centralizált jellegét erősíti és a KSH szerepét, tekintélyét rendkívüli módon növeli, erről szól a 21. § (4) szakasza. Fontossága miatt teljes egészében idézzük:

„Stt. 24. § (4) A KSH jogosult

a) az (1) és (3) bekezdésben meghatározott adatgyűjtést végző szervtől az országos statisztikai adatgyűjtési programból rendelkezésre álló adatokat egyedi azonosításra alkalmas módon,

b) az országos statisztikai adatgyűjtési programban, valamint nemzetközi kötelezettségvállalásban szereplő adatgyűjtéseiben meghatározott adatkörben, az adatgyűjtés céljának és időtartamának megjelölésével, a hatósági, ellenőrzési és gazdasági tevékenységgel járó, továbbá a nyilvános és közhitelű nyilvántartások vezetésével összefüggő adatgyűjtést végző szervtől a rendelkezésre álló személyes és nem személyes adatokat – ha törvény eltérően nem rendelkezik – egyedi azonosításra alkalmas módon, statisztikai felhasználás céljából, az adatvédelemre vonatkozó szabályok megtartása mellett átvenni.

A KSH az adatszolgáltatókat a személyes adatok statisztikai célra történő átvételéről és felhasználásáról hivatalos lapjában előzetesen tájékoztatja. Az adatátvétellel összefüggő térítésre a (2) bekezdés irányadó.”

A 21. § (4) szakaszában foglalt rendelkezések lényege, hogy a KSH nemcsak az Országos Statisztikai Adatgyűjtési Program keretében felvett adatokat veheti át, hanem más, ún. másodlagos adatforrásból származó, nem statisztikai célból gyűjtött adatokat is (kiemelten: adminisztratív adatok átvétele, például cégnyilvántartásból, adónyilvántartásból stb.). Az ilyen adatok átadási kötelezettsége az átadónak nem

okoz felesleges megterhelést, mert már rendelkezésére álló adatállományra vonatkozóan jelent kötelezettséget, ugyanakkor jelentősen csökkenti a statisztikai adatgyűjtésekből származó adatszolgáltatói terheket, mely az Európai Statisztikai Rendszer szintén kiemelt célja.

Az Országgyűlésben vita volt arról, hogy a KSH milyen mértékben vehet át egyedi adatokat. (KSH [1994c] 121. old.) A vita eredményeképpen a statisztikai törvényben a KSH elég tág hatáskörű jogosítványokat kapott, amely a KSH iránti bizalmat mutatta. Az egyedi adatok átadási lehetőségénél a képviselők figyelembe vették, hogy a KSH-ban nem folyik olyan hatósági jellegű tevékenység, amelyből a személyi adatok és az egyedi adatok átvétele során az adatszolgáltató személyére hátrányos következmények adódnának. A (4) szakasz *b*) pontja esetében az adatszolgáltató szerv adatátadására vonatkozó szabályok határozzák meg azokat a személyes adatokat, amelyeket a KSH jogosult statisztikai célra átvenni. Az adatátvételtől előzetesen a KSH-nak hivatalos lapjában (*Statisztikai Közlöny*) tájékoztatni kell az adatszolgáltatókat. Mivel a KSH a természetes személyek személyes adatát is átveheti, és személyes adatot csak törvényi felhatalmazás alapján lehet gyűjteni, a törvényalkotás folyamatában érvényesíteni kell a KSH-nak azt az igényét, hogy ezekhez az adatokhoz hozzájuthasson.

Az elmúlt 22 év további törvénymódosításai közé tartozik, hogy az egyik legfontosabb adminisztratív adatforrásról, az adózással összefüggő információk átvételének körülményeiről az 1993. évi statisztikai törvény külön is rendelkezik, az utóbb beiktatott 25/B. § segítségével:

„Stt. 25/B. § A KSH jogosult az adózással összefüggő tény, adatot, iratot – az adatvédelemre vonatkozó szabályok betartásával, jogszabályban meghatározott nemzetközi adatszolgáltatási kötelezettségének teljesítése érdekében – egyedi azonosításra alkalmas formában átvenni és megőrizni. A természetes személyre és az egyéni vállalkozónak a nem a gazdasági tevékenységéhez kapcsolódó adatára vonatkozó adatok kezelésénél a 19. §-ban foglaltak szerint kell eljárni.”

Az Infotv.-nek és a statisztikai törvénynek is vannak olyan rendelkezései, melyek kapcsolódnak egymáshoz, ezért az Infotv.-nek néhány statisztikai célú adatkezelést érintő rendelkezése a korábbi adatvédelmi törvénybe foglaltakhoz képest kissé változott formában, de fennmaradt.

Az Infotv. önálló cím és paragrafus alatt szabályozza a személyes adatok statisztikai célú adatkezelését.

„Infotv. 12. Személyes adatok felhasználása statisztikai célra
13. § (1) A kötelező adatkezelés keretében kezelt személyes adatokat – ha törvény eltérően nem rendelkezik – a Központi Statisztikai

Hivatal statisztikai célból egyedi azonosításra alkalmas módon átveheti és törvényben meghatározottak szerint kezelheti.”

Az Infotv.-nek ez a megfogalmazása megteremti a kapcsolatot az Infotv. és a statisztikai törvény között, ugyanakkor lehetőséget is ad az eltérésre. Emellett jelentős változás a korábbi adatvédelmi törvényhez képest, hogy az Infotv. általános felhatalmazást ad a KSH részére a személyes adatokat tartalmazó adminisztratív adatállományok statisztikai célú átvételére, ezzel megnyitva az utat például egy regiszteren alapuló népszámlálás megtartása felé. Az eddigi szabályozás főleg a személyazonosításra alkalmatlan módon történő adatátvételt preferálta, illetve az ágazati törvényekben külön szabályozta azokat az eseteket, amikor a KSH személyazonosításra alkalmas módon juthatott az adatokhoz. A jelen szabályozás lényegében általános felhatalmazást ad a KSH-nak a személyes adatok átvételére. Ehhez azonban egy sor törvénytől hatályon kívül kell helyezni a „személyazonosításra alkalmatlan módon” kitétel, illetve az új statisztikai törvényben kell szabályozni a részleteket és megadni a biztosítékokat a személyes adatok átvétele, kezelése tekintetében. A KSH ezzel a bizalmi felhatalmazással akkor tud élni, ha a közigazgatás szervei együttműködnek, már az adminisztratív adatállományok felállításánál, kezelésénél figyelembe veszik a statisztikai célú adatkezelés szempontjait, módszertani ajánlásait.

4. A KSH belső adatvédelmi szabályozása

A statisztikai célból felvett adatok védelmére – különböző szintű jogszabályokból álló –, valóságos „védőrendszer” alakult ki. Ezek közé tartozik a már bemutatott statisztikai törvény és annak végrehajtási rendelete. Nagyon lényeges, hogy olyan jelentős statisztikai célú adattömeget kezelő szervezetnek, mint a KSH, jól működő, átlátható belső adatvédelmi szabályozása legyen. A KSH-ban az adatvédelem kérdéskörét be kellett illeszteni a KSH stratégiai tervezésének folyamatába. Fontos szempont, hogy az adatvédelem és -hozzáférés a statisztikai adatok előállítási folyamatának szoros részét képezi. Olyan jelentős létszámú és feladatkörű intézménynek, mint a KSH, gondosan felépített, koncepcionálisan is kidolgozott adatvédelmi szabályozási rendszerének kell lennie. Különösen így van ez, ha olyan jelentős, nemzetközileg is betagozódott intézményről van szó, mint a KSH. A Hivatal stratégiai dokumentumai között meghatározó szerepük van egy-egy szakterület feladatait meghatározó politikáknak. Az adatvédelem szabályozásának alaphangját a KSH 2020-ig érvényes stratégiai dokumentuma (Stratégia 2020), (KSH [2014a]), illetve a KSH vezetése által 2014-ben elfogadott „A KSH adatvédelmi politikája, 2014” (KSH [2014b]) című dokumentum adja meg.

Az adatvédelmi politika utal a KSH közel másfél százados múltjára, arra, hogy, amit a jelen tanulmány is igyekszik bemutatni, a statisztikai célú adatkezelése során e hosszú időszak alatt mindig tekintettel volt az általa begyűjtött egyedi adatok védelmére, ezáltal lefektetve az adatvédelem 9 alapelvét. Leszögezi, hogy a KSH törekszik az egyedi adatok védelme és az adatokhoz való hozzáféréshez fűződő igények közötti legjobb egyensúly kialakítására és fenntartására. Szól a KSH munkatársainak az adatvédelem iránti elkötelezettségéről, e feladathoz kapcsolódó munkák magas színvonalon történő ellátásáról. Felsorolja az adatvédelem olyan elveit, melynek összefoglalásához segítséget nyújtottak az Európai Statisztika Gyakorlati Kódex elvei, ismérvei.

A KSH belső adatvédelmi szabályozásának legfőbb dokumentuma a jelen tanulmányban már többször hivatkozott adatvédelmi szabályzat. A szabályzat – kódexjelleggel – összefoglalja, egységesíti és átláthatóvá teszi a KSH adatvédelmi gyakorlatát. (A felhasználók e tárgybeli tájékozódását az is elősegíti, hogy ez a fontos szabályzat – az adatvédelmi politika mellett –, teljes tartalmában megtekinthető a KSH honlapján.) A jelentős létszámú és hierarchikus felépítésű KSH-ban valamiképpen rendezni kellett az adatvédelemmel kapcsolatos szervezeti kérdéseket. Ezt tette meg a KSH Szervezeti és Működési Szabályzatában, mely állandóan működő bizottságként felállította a KSH Adatvédelmi Bizottságát. Ehhez kapcsolódva az adatvédelmi szabályzat összefoglalja az Adatvédelmi Bizottság legfontosabb feladatait és rendelkezik az adatvédelem KSH-beli többi szereplőjéről. A KSH-ban folyó adatvédelemmel kapcsolatos munka fontos dokumentumai a KSH Adatvédelmi Bizottságának állásfoglalásai. Ezek az állásfoglalások választ adnak a KSH munkatársainak adatvédelemmel kapcsolatos problémáira, egységesítik a statisztikai adatok kezelésével kapcsolatos adatvédelmi feladatokat és gyakorlatot. Bár ezekben az állásfoglalásokban megfogalmazott ajánlások nem kötelezők, de a tapasztalatok azt mutatják, hogy a KSH munkatársai elfogadják az adatvédelmi bizottság adatvédelemmel kapcsolatos útmutatását, így azok érdemben megjelennek a KSH mindennapi gyakorlatában.

Az adathozzáférési szabályzat részletezi az adathozzáférési csatornák jellemzőit, közli a statisztikai adatok megismerésének szabályait, az ezzel kapcsolatos belső nyilvántartás rendelkezéseit. Ezt követően sorra veszi az adathozzáférési csatornák adatvédelemmel kapcsolatos részletes szabályait. Fontos része a szabályzatnak a tudományos célú hozzáférés és a kutatói akkreditáció rendelkezéseinek összefoglalása, és egyáltalán, a kutatási célú adathozzáférés részletes szabályozása. Tekintettel arra, hogy e szabályzat főleg a KSH-n belüli munkafolyamatok kérdését rendezi, csak a felhasználók számára fontos szabályok összefoglalása jelenik meg a KSH honlapján.

Az adatvédelem, mely tárgyánál fogva az adatainak védelmét jelenti, számos összefüggésben van magának az adatnak a védelmével, mely az adatbiztonság témaköréhez tartozik. Az adatbiztonság kérdése különösen fontossá válik azáltal, hogy az informatika fejlődése révén egyre több adat feldolgozása válik egyre könnyebbé és

így egyre elérhetőbbé, elterjedtebbé. Így ezeknek az adatoknak a biztonságos védelme jelentős kihívást jelent az olyan nagy tömegű adathalmazt kezelő intézmény számára, mint a KSH. Mindazonáltal az adatvédelmet és az adatbiztonságot nem lehet mesterségesen szétválasztani, mert szorosan összefüggenek egymással. Ez a KSH belső szabályozási rendszerében is jelentkezik, mely az adatvédelem kérdéséről a vonatkozó jogi, módszertani, tájékoztatás- és informatikai biztonsági területek együtteseként szól.

A KSH informatikai biztonsági szabályozásának alapját a „KSH informatikai biztonsági politika” című dokumentuma adja meg. Ezen alapvetésben a legfontosabb biztonságpolitikai alapelvek kerülnek felsorolásra (biztonság, sértetlenség, rendelkezésre állás, tudatos kockázatviselés stb.), továbbá összefoglalásra kerülnek a KSH informatikai biztonságának főbb területei. Ezek közé tartozik például az információhoz való hozzáférés kérdése, a szoftverkezelés biztonsága, az elektronikus bűnözés elleni küzdelem stb..

A KSH informatikai biztonsági politikájával összefüggésben készült a KSH informatikai biztonsági szabályzata. E részletes szabályzat kiterjed az összes szóba jöhető informatikai biztonsági kérdésre, az adminisztratív, logikai és fizikai védelmi intézkedésekre.

A KSH „védőrendszerének” további dokumentumai közé tartoznak az adatkezelés külső körülményeinek biztosításával kapcsolatos belső rendelkezések (például a KSH Tűzvédelmi Szabályzata), az iratkezelés szabályozása (például a KSH iratkezelési szabályzata), a minősített adatok védelme (például a minősített adatok védelmének biztonsági szabályzata).

5. Az adatvédelemmel kapcsolatos országgyűlési beszámolók statisztikai vonatkozásai

Az Infotv. 38. § (4) bekezdés b) pontja alapján a NAIH-nak évenként beszámolót kell készítenie az előző évi tevékenységéről. Ezt a beszámolót az Országgyűlés illetékes bizottsága megtárgyalja és jóváhagyja. Ugyanez volt a korábbi adatvédelmi biztosoknak is a kötelezettsége. Akkor azonban azokat a beszámolókat az Országgyűlés plenáris ülése tárgyalta meg és fogadta el.

Tekintettel arra, hogy az Infotv. hatálya csak a természetes személyekre terjed ki, ezért – érthető módon – a NAIH (és az adatvédelmi biztos) figyelme főleg az olyan adatgyűjtésekre, adatfeldolgozásokra irányul, ahol közvetlen vagy közvetett azonosítható módon kezelik a természetes személyekre vonatkozó adatokat. Ezzel kapcsolatban meg kell jegyezni, hogy a NAIH által vizsgált információk többségét admi-

nisztratív, illetve üzleti célból kezelik, a statisztikai célból történő adatkezelés ehhez képest kisebb arányt képvisel. Ez lehet az egyik oka annak, hogy 2004 év beszámolóját átvizsgálva megállapítható, hogy a statisztikákkal kapcsolatos megállapítások csak néhány év beszámolójában fedezhetők fel. A másik ok pedig az – különösen a KSH vonatkozásában –, hogy a statisztikai célú adatok kezelését olyan apparátus végzi, amelyik közel 150 éve alkalmazza a személyes adatok védelmének modern elveit és gyakorlatát, miáltal kevés olyan beadvány érkezett az évek során, mely kifogásolta volna ezt a tevékenységet.

Az 1990-es évek végén, a közigazgatás korszerűsítése kapcsán, volt egy olyan elképzelés, hogy a nyugdíjbiztosító (ONYP¹⁴), az egészségbiztosító (OEP¹⁵) az adóhivatal (APEH¹⁶) és a KSH bizonyos elemi szintű adatait, – igazgatási célból – kapcsolják össze, és a KSH legyen az adatállomány kezelője. Ezt a tervet az akkori adatvédelmi biztos (*Majtényi László*) határozottan ellenezte, tekintettel arra, hogy a KSH feladatainál fogva nem alkalmas ilyen jellegű tevékenységre, másfelől statisztikai célból kezelt adatot nem szabad igazgatási célra igénybe venni. Ez a vélemény is hozzájárulhatott ahhoz, hogy a tervezett intézkedés elmaradt. Ez az állásfoglalás ezt követően hivatkozási alap volt mindig, amikor statisztikai célú információ igazgatási célra való felhasználásának kérdése felmerült (*Lakatos* [2003] 21. old.).

Szintén precedens értéke volt az adatvédelmi biztosnak az az állásfoglalása, mely jelezte, hogy az 1996. évi mikrocenzus adatállományainak kutatási célú felhasználása esetén is mindent meg kell tenni a felfedés kockázatának minimalizálása érdekében, az adatállományt átvevő kutatóintézetnek különösen védenie kell adatkezelése során ezt az állományt. Az adatvédelmi biztos részletesen le is írja, hogy mit ért ezen a különleges védelmen. 2003-ban az adatvédelmi biztos (*Péterfalvi Attila*) még ennél is tovább ment, amikor az akkori Oktatási Minisztérium megbízott egy kutatócsoportot a roma népesség kutatásával. Akkor az adatvédelmi biztos egyetértett azzal, hogy ezeknek az igen érzékeny adatoknak a feldolgozása, a kutatócsoporttal együttműködve, a KSH-ban történjék meg. Azóta a helyzet sokkal szabályozottabb lett, hiszen a KSH ilyen jellegű statisztikai mikroadatokra vonatkozó tudományos célú hozzáférési igényeket jellemzően anonimizált mikroadat-kiadás, illetve biztonságos környezetben történő hozzáférés (kutatószoba) formájában oldja meg, ahol ellenőrzött körülmények között dolgozhatnak a kutatók.

Alapvető kérdésre adott választ az adatvédelmi biztos, amikor arról volt szó, hogy a KSH által kezelt, az 1945 és 1948 közötti, a német nemzetiségű lakosságot érintő egyedi adatokból a kérelmező saját adatait kikérheti-e a KSH-tól (*Lakatos-Heinz* [2004]). Az adatvédelmi biztos úgy foglalt állást, hogy a kérelmező az eredeti adat-

¹⁴ ONYP: Országos Nyugdíjpénztár.

¹⁵ OEP: Országos Egészségbiztosítási Pénztár.

¹⁶ APEH: Adó- és Pénzügyi Ellenőrzési Hivatal.

gazdához, a BM Bevándorlási és Állampolgársági Kirendeltségéhez fordulva kérje ki az igényelt információt. Ez a probléma többször felmerült a KSH gyakorlatában, jellemzően, amikor a KSH statisztikai célból adminisztratív adatokat vett (vagy vesz) át, és egy adatkérő ebből az adatállományból saját magára vonatkozó adatot kért (kér). A KSH Adatvédelmi Bizottsága is több állásfoglalásában megerősítette, hogy ilyen esetben az adatkérő adathozzáférési igényével az eredeti adatokat kezelő intézményhez forduljon.

Szintén alapvető dilemmára adott választ az adatvédelmi biztos, amikor az egyéni vállalkozók személyes adatainak természetéről szólt. Véleménye szerint üzleti kapcsolataiban az egyéni vállalkozó nem tekinthető természetes személynek, hiszen a gazdasági életben való részvétel is azáltal válik lehetővé számára, hogy a jog bizonyos szempontból cégnek tekintsi. Így ebben a vonatkozásban az egyéni vállalkozással kapcsolatos személyi adatok – például amikor a lakcím egyben a vállalkozó székhelye is – nem tekinthetők személyes adatnak, így kezelésükre, illetve továbbításukra sem vonatkoznak az adatvédelmi törvény rendelkezései.

Az éves beszámolóknak több olyan állásfoglalás, vélemény is található, melyet közvetve a statisztikai adatkezeléseknél is figyelembe lehet venni. Ilyen például az, hogy az internetes adatszolgáltatási módoknál hogyan kell értelmezni az érintett kifejezett hozzájárulását. Nem részletezve a körülményeket, a lényeg az, hogy az interneten akkor tekinthető megadottnak a hozzájárulás, ha egyértelműen bizonyítható, hogy az érintettől származik, engedélyező tartalmú, illetve, hogy az adatszolgáltató érintett adatszolgáltatására vonatkozik. A hozzájárulás érvényességét, illetve az adatkezelés jogszerűségét az adatkezelőnek kell biztosítani, illetve bizonyítani. Amennyiben az adatkezelő igazolni tudja, hogy az adatkezeléshez való hozzájárulást az érintett megadta, az adatkezelés jogszerű. Miután a jövőben egyre több statisztikai adatfelvétel az internetes adatfelvételi mód alkalmazásával fog történni, a NAIH ezen állásfoglalása alapvető jelentőségű.

Az adatvédelmi biztosnak (2001-ben) és a NAIH elnökének (2011-ben) meghatározó szerepe volt ezekben az években végrehajtott két népszámlálás szakszerű, adatvédelmi szempontból példaértékű végrehajtásában. A mindenkor KSH legnagyobb vállalkozásai közé tartozik az aktuális, a teljes magyarországi népesség és lakásállomány összeírásával – közkeletű megnevezéssel a népszámlálással – kapcsolatos munkák tervezése, előkészítése és végrehajtása, majd az adatok feldolgozása és nyilvánosságra hozása. Az 1995-ben megválasztott adatvédelmi biztos (Majtényi) már a kezdetektől fogva kapcsolatban állt a KSH népszámlálási munkatársaival, érdeklődéssel figyelte a népszámlálási tevékenységet. Tette ezt többek között azért, mert az adatvédelem európai szintű szabályozására nagy hatással volt a német Alkotmánybíróság 1983. évi ún. népszámlálási döntése, mely a népszámlálás okán számos adatvédelemmel kapcsolatos kérdést is felvetett. Az adatvédelmi biztos már az 1996. évi mikrocenzus kérdőívét is jóváhagyólag véleményezte és a 2001. évi népszámlálás egész munkafolyamatát figye-

lemmel kísérte (*Lakatos* [2013] 22. old.). Nem részletezve a 2001. évi népszámlálás során végzett adatvédelmi biztosi tevékenységet, jelezzük, hogy interjúiban és egyéb közéleti megnyilvánulásaiban támogatóan tevékenykedett a népszámlálás érdekében. A lakosság részéről érkezett beadványokra adott válaszaiban ugyanezt tette. 2002. február 5-én – már az újonnan megválasztott adatvédelmi biztos, Péterfalvi – összegző, értékelő levelet küldött a KSH elnökének, melynek hangneme jól kifejezi azt az együttműködési készséget, mely a népszámlálás szervezői és az adatvédelmi biztos között kialakult. Ez a pozitív kapcsolat tovább élt a 2005. évi mikrocenzus és a 2011. évi népszámlálás végrehajtásakor. Segítette az együttműködést az is, hogy az utóbbi két jelentős népszámláláskor is részben ugyanaz a személy – Péterfalvi – volt hivatalban, 2011-ben, mint a NAIH elnöke, a népszámlálás összes munkafolyamatát felügyelve támogatta a KSH munkáját.

6. Jövőkép

Végigkövetve az elmúlt közel 150 év statisztikai célú adatkezelések adatvédelemmel kapcsolatos hazai jogi szabályozásának történetét, megállapítható, hogy már elődeink is kiemelt fontosságot tulajdonítottak ennek a témakörnek. Az is nyilvánvalóvá vált ugyanakkor, hogy e hosszú időszak alatt, körülbelül 20-25 évenként, a statisztikai törvényeket meg kellett újítani. Nincs ez másképpen napjainkban sem; látható, hogy a 22 éve hatályban levő 1993. évi XLVI. statisztikáról szóló törvény helyett új jogszabályt kell alkotni, tekintettel arra, hogy a jelenlegi törvény számos ponton elavult és a folyamatosan megújuló európai uniós szabályozással sincs teljesen összhangban. Jelen tanulmánynak nem tárgya a jelen cikk megírásának idején is folyó és egyeztetés stádiumában levő kodifikáció egyes lépéseinek bemutatása, itt csupán csak az adatvédelem témakörének néhány fontosabb leágazására utalunk.

Mindenekelőtt szükség van a statisztikai törvény fogalomrendszerének, különösen az adatvédelemmel kapcsolatba hozható definícióinak a megújítására, egységbe foglalására, a magyar és az európai uniós joganyagokban levő fogalmakkal történő harmonizációjára. Így összehangoltan kell értelmezni például az egyedi adat (bizalmas adat), a közérdekű adat, a mikroadat fogalmait, az ezen adatok kezelésére, védelmére, illetve nyilvánosságra hozatalára vonatkozó definíciókat, az aktív és a passzív adatvédelem fogalmait.

A statisztikai célú adatkezelés teljes garanciarendszerét a statisztikai törvénynek kell tartalmaznia. Mivel az Infotv. megengedi a törvényi eltérést a statisztikai célú felhasználástól, a statisztikai törvényben kell majd pontosan leírni azokat a területe-

ket, amelyek esetében megengedett ez az eltérés, illetve azokat, amelyek esetében az kategorikusan kizárt (hatósági, ellenőrzési cél stb.).

A másik nagy terület, amelynek az Infotv. teremti meg a jogi alapjait, az adminisztratív adatok statisztikai célú felhasználása. Ennek keretében szabályozni kell majd az adminisztratív nyilvántartások kezelőinek kötelezettségét a hivatalos statisztika igényeinek figyelembe vételére a nyilvántartások kialakításakor, módosításakor; az adminisztratív adatok minőségének folyamatos mérését, az adatátvitel ingyenességét, az adminisztratív adatokhoz való tudományos célú hozzáférés lehetőségeit, stb.

A KSH-n kívül a Hivatalos Statisztikai Szolgálathoz tartozó szervezetek is kezelnek statisztikai adatokat. Adatvédelmi szempontból is megnyugtató lenne, ha e szervezeteken belül a statisztikai célú adatkezelések hangsúlyosan elkülönülnének mind szervezeti, mind személyi szinten a többi például adminisztratív célú adatkezelésektől és ezt a tényt a KSH vizsgálná, és e szervezeteket adatvédelmi szempontból is akkreditálná, auditálná.

Rendezni kell a – kiemelten statisztikai mikroadatokhoz való – tudományos célú hozzáférés szabályait, bár erről, mint említettük, külön európai uniós jogszabály rendelkezik, mégis célszerű lenne e kérdést egy jogszabályban a statisztikai törvényben szabályozni.

Az Európai Statisztika Gyakorlati Kódexéhez hasonlóan el kell készíteni a Nemzeti Statisztika Gyakorlati Kódexét és e kódexen belül az adatvédelem kérdéseit is rendezni szükséges. A Nemzeti Statisztika Gyakorlati Kódexe képviseli azt a statisztikai adatot kezelő intézmények működésére, termékeire, folyamataira vonatkozó keretrendszert, melynek keretében az adatok védelmének minden aspektusa (jogi, módszertani, fizikai, stb.) biztosítható. Ebben a tekintetben a Nemzeti Statisztika Gyakorlati Kódexe egyben az adatvédelmi szabályok és egységes iránymutatások újabb garanciája is, mely végső soron a statisztikai adatokba vetett bizalmat erősíti, biztosítva egyúttal a tudományos célú hozzáférés lehetőségét.

A Nemzeti Statisztika Gyakorlati Kódexén felül a statisztikai célból kezelt adatokhoz való tudományos célú hozzáférés módszereinek, csatornáinak és erre épülő szolgáltatások minőségének folyamatos javítása is cél kell, hogy legyen; szoros együttműködésben az adatarhívumokkal és a kutatói szférával.

Az ún. statisztikai regiszter alapú adatok nyilvánosságra hozásának feltételeit pontosabban kellene megfogalmazni, hogy a közhiteles és mindenki számára elérhető nyilvántartásokban levő információ a statisztikai regiszterekből is nyilvánosságra hozható legyen. Ebben az esetben határozottabban kellene szabályozni, hogy az adatigénylő mit és milyen szabályok szerint kérhet a KSH-tól és mit a közhiteles nyilvántartást kezelő intézménytől.

Irodalom

- BÁNSZEGI K. [1993]: Beszámoló a hatodik statisztikai törvény megalkotásáról. *Statisztikai Szemle*. 71. évf. 741–745. old.
- BÁNSZEGI K. – LAKATOS M. [1994]: Információszabadság – Adatvédelem – Statisztika. (III.) *Statisztikai Szemle*. 72. évf. 10. sz. 761–777. old.
- BÁNSZEGI K. – LAKATOS M. [2000]: Információszabadság – Adatvédelem – Statisztika (VII.) *Statisztikai Szemle*. 79. évf. 12. sz. 444–455. old.
- BUDAY L. (szerk.) [1911]: *A Magyar Királyi Központi Statisztikai Hivatal munkássága (1871–1911)*. Magyar Statisztikai Közlemények. Új sorozat 36. köt. Pesti Könyvny. VI. Budapest.
- InfoFilia* [1991–1993]: *Informatika – Jog – Közigazgatás. Nemzetközi dokumentumok I.–V.*, Magyar Adatvédelem és Információszabadság Alapítvány. Budapest.
- KLINGER A. – KEPECS J. (szerk.) [1990]: *A magyar népszámlálások előkészítése és publikációi*. I. köt. Központi Statisztikai Hivatal. Budapest.
- KLINGER A. – KEPECS J. (szerk.) [1991]: *A magyar népszámlálások előkészítése és publikációi*. II. köt. Központi Statisztikai Hivatal. Budapest.
- KLINGER A. – KEPECS J. (szerk.) [1995]: *A magyar népszámlálások előkészítése és publikációi*. III. köt. Központi Statisztikai Hivatal. Budapest.
- KÖNYVES TÓTH P. [2010]: Az adatvédelmi törvény metamorfózisai. *Fundamentum*. XIV. évf. 2. sz. 53–61. old.
- KSH (KÖZPONTI STATISZTIKAI HIVATAL) [2013]: 27/2013. számú KSH utasítás a Központi Statisztikai Hivatali adatvédelmi szabályzatáról. Budapest. http://www.ksh.hu/docs/szolgáltatások/adatigenyles/ksh_adatvedelmi_szabalyzat.pdf
- KSH [2014a]: *Központi Statisztikai Hivatal Stratégia 2020*. Budapest, http://www.ksh.hu/docs/bemutakozas/hun/strategia_2020.pdf
- KSH [2014b]: *A KSH adatvédelmi politikája, 2014*. Budapest. http://www.ksh.hu/docs/bemutakozas/hun/adatvedelmi_politika_2014.pdf
- KSH [1994c]: *A magyar statisztikai törvények*. Budapest.
- LAKATOS M. [1993]: A személyi szám és a statisztikai információ-rendszer. *Statisztikai Szemle*. 71. évf. 10. sz. 815–826. old.
- LAKATOS M. [1994]: Információszabadság – Adatvédelem – Statisztika. (I.) *Statisztikai Szemle*. 72. évf. 7. sz. 547–559. old.
- LAKATOS M. [2000]: Az adatvédelem jogi szabályozása a magyar népszámlálások történetében. *Statisztikai Szemle*. 79. évf. 10–11. sz. 794–812. old.
- LAKATOS M. [2003]: Információszabadság – Adatvédelem – Statisztika (VIII.) *Statisztikai Szemle*. 83. évf. 1. sz. 5–27. old.
- LAKATOS M. – HEINZ E. [2004]: *A Központi Statisztikai Hivatal szerepe a német lakosság kitelepítésében. A Magyarországi németek kitelepítése és az 1941. évi népszámlálás*. MST Statisztika-történeti Szakosztálya, KSH Levéltár. Budapest.
- LAKATOS M. – NAGY E. [2012]: Információszabadság – Adatvédelem – Statisztika (X.) *Statisztikai Szemle*. 90. évf. 1. sz. 18–41. old.
- THIRRING L. [1983]: *Az 1869–1980. évi népszámlálások története és jellemzői*. I. rész. Központi Statisztikai Hivatal. Budapest.

- THIRING L. [1988]: *Az 1869–1980. évi népszámlálások története és jellemzői*. II. rész. Központi Statisztikai Hivatal. Budapest.
- THIRING L. [1989]: *Az 1869–1980. évi népszámlálások története és jellemzői*. III. rész. Központi Statisztikai Hivatal. Budapest.
- VÉGVÁRI J. [1991]: Az új magyar statisztikai törvény koncepciója. *Statisztikai Szemle*. 69. évf. 6. sz. 413–424. old.

Summary

In Hungary, the relationship between statistics and data protection has been the subject of legislation for more than a century. The socio-economic development necessarily involves the constant revision of legal rules. This study, on one hand, reviews the changing legislation of this long period, in particular the statistical law in force and, on the other hand, points out that instead of the outdated regulations, the adoption of new ones is required in statistics and data protection.

A statisztikai adatvédelem és -hozzáférés szabályai az Európai Unióban

Nagy Eszter,
a KSH szakmai tanácsadója
E-mail: Eszter.Nagy@ksh.hu

A cikkben bemutatott európai uniós szabályozások, eljárások és gyakorlat alapján jól érzékelhető, hogy az adatvédelem kiemelt jelentőséggel bír az európai statisztikákban. Mindezekre tekintettel az elmúlt időszakban elsődleges cél volt az ezzel kapcsolatos uniós és nemzeti szabályozás összehangolása, egységes elvek alapján történő kialakítása, amelyet az adatkezelési gyakorlatok minél nagyobb mértékű közelítése is kísért. Az európai uniós és a nemzeti szabályok egymásra épülése, és azok azonos elvek és szempontok szerint történő kialakítása egyszerre szolgálja az adatszolgáltatók érdekeinek a védelmét és a felhasználóknak egyenlő hozzáférési lehetőségét az adatokhoz az Európai Statisztikai Rendszeren belül.

TÁRGYSZÓ:
Európai Statisztikai Rendszer (ESR).
Bizalmas adat.
Tudományos célú hozzáférés.

Az európai statisztikák fejlesztését, előállítását és nyilvánosságra hozatalát az Európai Unióban az ESR¹ végzi. Az ESR tagjai a nemzeti statisztikai hivatalok, a tagállami szinten európai statisztikák előállítását végző vagy abban közreműködő egyéb nemzeti hatóságok, valamint az Európai Bizottság statisztikáért felelős hatósága, az Eurostat. Az európai statisztikák előállítása során az ESR tagjainak egységes elveknek kell megfelelniük annak érdekében, hogy azok működése és teljes adat-előállítási folyamatai megfeleljenek a felhasználók igényeinek, az adatszolgáltatók érdekeinek és az ESR minőségi elvárásainak. Ezeket az egységes elveket mind tagállami, mint európai uniós szinten érvényesíteni kell.

Az adatszolgáltatók érdekeinek védelme szempontjából ezen alapelvek között kiemelt jelentőséggel bír az adatok bizalmas kezelésének elve. A statisztikai adatkezelés során mindenkor biztosítani kell az adatszolgáltató által statisztikai célra szánt adatok védelmét, ugyanakkor, ezzel egyidejűleg, a felhasználók számára a lehető legszélesebb körű hozzáférést kell nyújtani az előállított statisztikai adatokhoz.

A jelen cikk célja, hogy bemutassa a statisztikai adatkezelésre és -hozzáférésre vonatkozó európai uniós szabályozás és gyakorlat legfontosabb elemeit.

1. Általános adatvédelmi szabályozás

A statisztikai adatkezelés keretszabályai az Európai Unióban is alapvető szinten megjelennek összefüggésben a személyes adatok védelméhez való joggal. Tekintettel arra, hogy a statisztika is kezel személyes adatokat, annak szabályozását nem lehet elszakítani az adatvédelem általános garanciáitól, ugyanakkor figyelemmel kell lenni a statisztikai cél sajátos jellegére és az azzal összefüggő tájékoztatási tevékenységhez fűződő közérdekre.

Mindezekre tekintettel a személyes adatok feldolgozásában az európai statisztikai adatvédelem alapvető szabályait az Európai Parlamentnek és a Tanácsnak az egyének védelméről és az ilyen adatok szabad áramlásáról szóló 95/46/EK-irányelvvel határozza meg.

Az irányelv a tagállamokkal való összhang megteremtését tűzi ki célul, annak érdekében, hogy a személyes adatok feldolgozásakor a tagállamok gondoskodjanak a természetes személyek alapvető jogainak és szabadságjogainak védelméről, különös

¹ ESR: Európai Statisztikai Rendszer.

tekintettel a magánélethez fűződő jogokra. A jogszabályi formából következően fogalmakat és keretszabályokat határoz meg, a részletes rendelkezések kidolgozását részben a tagállamok nemzeti szabályozására, részben pedig egyéb ágazati uniós szabályozásra bízta.

Ilyen uniós jogszabály az Európai Parlament és a Tanács 45/2001/EK-rendelete az egyének védelméről a személyes adatok közösségi intézmények és szervek által történő feldolgozásakor, valamint az ilyen adatok szabad áramlásáról, amely nemcsak a statisztikára vonatkozik, hanem általános jellegű keretszabályként határoz meg az irányelvvel összhangban alapvető garanciákat, az európai uniós szintű adatkezelések, tehát az ESR-en belüli statisztikai adatkezelésben. A statisztika szempontjából fontos, hogy a személyes adat fogalmán túl, egymással teljesen összhangban, mindkét jogszabály meghatározza az adatkezelő, valamint az adatfeldolgozó, és a harmadik személy fogalmát.

A 95/46/EK-irányelv a fogalmak tisztázásán túl lefekteti az adatvédelem legfontosabb elveit, követelményeit is. Ennek megfelelően mindenekelőtt rögzíti a célhoz kötöttség elvét, amikor kimondja, hogy a személyes adatok feldolgozását tisztességesen és törvényesen kell végezni, azok gyűjtése csak meghatározott, egyértelmű és törvényes célból történhet, és további feldolgozása nem végezhető e célokkal összeférhetetlen módon. A 95/46/EK-irányelv mindezekhez hozzáteszi, hogy a statisztikai feldolgozást ilyen törvényes célnak tekinti, ezáltal megteremtve a jogalapot a statisztikai adatfelvételekhez.

Ez a felhatalmazás ugyanakkor nem korlátlan, és az irányelv szabályai is előírják, hogy az adatkezelőnek statisztikai célú adatkezelés során is meg kell teremtenie a megfelelő garanciákat, különösen annak biztosítása érdekében, hogy az adatokat semmilyen más célból ne dolgozzák, és ne használják fel egy konkrét személlyel kapcsolatos intézkedések vagy döntések alátámasztásához. Ezzel kimondja a statisztikai cél kizárólagosságának a követelményét, ami védi az adatszolgáltatót attól, hogy az általa statisztikai célra szolgáltatott adatát bármely egyéb, számára akár pozitív, akár hátrányos jogkövetkezménnyel járó más eljárásban felhasználják. A 95/46/EK-irányelv által kimondott másik fontos elv, hogy az adatkezelés nemcsak célját tekintve, hanem időben is korlátozott, azaz az adatkezelés céljának megfelelő ideig lehet a személyes adatokat ilyen jellegüket megőrizve kezelni, azok tárolásának olyan formában kell történnie, amely az érintettek azonosítását csak az adatok gyűjtése vagy további feldolgozása céljainak eléréséhez szükséges ideig teszi lehetővé.

A 95/46/EK-irányelv ugyanakkor elismeri, hogy a statisztikának – tekintettel a hosszabb időszakok vizsgálatára vonatkozó igényre – ezen a téren speciális szükségletei vannak, az ezekre vonatkozó eltérő szabályozás megállapítását a tagállamokra bízta.

A 45/2001/EK-rendelet a közösségi intézmények vonatkozásában már részletesebben rendelkezik a statisztikai célú adatkezelés sajátosságairól, amikor kimondja,

hogyan az említett garanciák biztosítása érdekében az adatot kezelő közösségi intézmény vagy szerv rendelkezik arról, hogy a statisztikai célból hosszabb ideig tárolandó személyes adatokat anonim adatként kell tárolni, vagy, ha ez nem lehetséges, az érintettek személyazonosságát titkosítani kell. Ez a rendelkezés, amelyet a magyar statisztikai törvény is alkalmaz, megteremti annak a lehetőségét, hogy a személyes adatokat egy belső azonosító képzésével anonimizálja az adatkezelő, majd a meghatározott statisztikai cél érdekében – amely lehet mintaválasztás, új adatoknak az állományhoz való csatolása stb. – azokat az azonosítókkal ismét összerendelve, a célhoz szükséges mértékben és ideig, ismét azonosíthatóvá tegye az állományt. Mindez ugyanakkor csak szigorúan igazolt statisztikai célból történhet. A célhoz kötöttség és a megfelelő tárolás követelményeinek teljesítése mellett a 95/46/EK-irányelv, valamint a 45/2001/EK-rendelet egyaránt megköveteli, hogy az adatkezelés az adatszolgáltató számára átlátható módon történjen. Az adatszolgáltató az adataival való információs önrendelkezési jogot a megfelelő tájékoztatás birtokában gyakorolhassa, és előírja, melyek azok az alapvető információk, amelyeket az adatszolgáltatást megelőzően az adatszolgáltató tudomására kell hozni (pl. az adatkezelés célja, az adatszolgáltatás kötelező vagy önkéntes jellege). A tájékoztatáskérés joga bármikor megilleti az érintettet, nemcsak az adat felvételekor, hanem az adatfeldolgozással összefüggésben is, ahogy azt mind a 95/46/EK-irányelv, mind a 45/2001/EK-rendelet kimondja. A megfelelő tájékoztatás mellett az érintett – szintén az információs önrendelkezési jogával összefüggésben – bármikor kérheti adatai helyesbítését vagy zárolását, törlését. Ennek a jognak a gyakorlása – különös tekintettel a törlés, zárolás kérésének jogára – a gyakorlatban jelentősen megnehezítené a statisztikai tevékenységet. Ennek felismerése alapján a közösségi intézmények adatkezelésével összefüggésben a 45/2001/EK-rendelet a statisztika területére kivételszabályt fogalmaz meg, és nem teszi kötelezővé ezeknek a szabályoknak az alkalmazását, amennyiben a statisztikai adatkezelésre vonatkozó garanciák megfelelően biztosítottak.

2. Az európai statisztikák uniós szabályozása

A korábban ismertetett két keretszabály jellegű jogszabályon kívül kifejezetten csak az európai statisztikák fejlesztésére, előállítására, nyilvánosságra hozatalára vonatkozik az Európai Statisztika Gyakorlati Kódexe (a továbbiakban Gyakorlati Kódex), valamint az európai statisztikákról szóló 223/2009/EK európai parlamenti és tanácsi rendelet.

A Gyakorlati Kódex 2005. május 25-én a Bizottság ajánlásaként született meg, és 15 alapelvben, valamint azok megvalósításához szükséges ismérvekben határoz-

za meg az európai statisztikákkal szemben támasztott alap- követelményeket, melyeknek az ESR valamennyi tagjának meg kell felelnie.

A Gyakorlati Kódex nem jogszabály, ezért közvetlen jogi kötőerővel nem rendelkezik sem az Eurostat, sem a tagállami statisztikai hivatalok, sem más hatóságok vonatkozásában.

Hangsúlyozni kell ugyanakkor, hogy a 223/2009/EK-rendelet a statisztika alapelvei között felsorolja a Gyakorlati Kódexben foglalt elvek közül a legfontosabb elveket, emellett utal arra, hogy az európai statisztikákat a Gyakorlati Kódexben foglaltakkal összhangban kell fejleszteni, előállítani és nyilvánosságra hozni, így áttételesen ugyan, de annak alkalmazása kötelező az ESR-en belül. A Gyakorlati Kódexben foglalt követelmények megvalósulását rendszeresen felülvizsgálják az ESR tagjainál és az auditok eredményeit nyilvánosságra hozzák.

A 223/2009/EK-rendelet mint jogszabály közvetlenül alkalmazandó a tagállamokban, az abban foglaltakat minden, az európai statisztikák előállításával, nyilvánosságra hozatalával kapcsolatos tevékenység során alkalmazni kell. A rendeletben foglaltak így a nemzeti statisztikák előállítására is hatással vannak, hiszen a tagállami jogi szabályozás nem lehet ellentétes az európai uniós rendeletben foglalt szabályokkal, ezért az a magyar statisztikai törvény háttérjogszabályának is tekinthető.

A Gyakorlati Kódexnek és a 223/2009/EK-rendeletnek egyaránt az a fő célja, hogy egységes szabályrendszert állítson fel a statisztika előállítói számára, legyen az Eurostat vagy a tagállami intézmények. Az egységes elvek érvényesítése fontos a megfelelő minőség biztosítása érdekében, az adatvédelem területén pedig a védelem azonos szintjét garantálja az adatszolgáltatók számára, és így biztonságossá teszi a statisztikai adatok átadását az ESR tagjai között, de legfőképpen az Eurostat felé. Az uniós jogi szabályozásban foglaltak biztosítják, hogy az adatszolgáltató adata az Eurostat kezelésében a védelem ugyanolyan szintjét élvezze, mint a KSH² kezelésében.

Figyelembe véve az adatvédelem kiemelt jelentőségét, külön elv vonatkozik rá a Gyakorlati Kódexben. A Gyakorlati Kódex 5. elve mondja ki a bizalmas adatkezelés követelményét mint elvet. Ennek alapján az ESR tagjai a legmesszebbmenőkig tiszteletben tartják az adatszolgáltatók (háztartások, vállalatok, közigazgatási intézmények és más válaszadók) jogait, biztosítják az általuk szolgáltatott információk bizalmas kezelését és kizárólag statisztikai célokra történő felhasználását.

Az elvhez öt ismérv tartozik, melyek felölelik az adatvédelem összes fontos területét az intézményi, jogi környezetet, beleértve a belső szabályozó eszközöket, illetve a fizikai, logikai, informatikai biztonságot. Az ismérvek között kiemelten megjelenik a tudományos célú adat-hozzáféréshez való igény, melyet az adatvédelemhez fűződő érdek sérelme nélkül kell biztosítani.

² KSH: Központi Statisztikai Hivatal.

A Gyakorlati Kódex mellett az európai statisztikákról szóló 223/2009/EK-rendelet rögzíti jogszabályban a statisztikai alapelvek között a bizalmas adatkezelés elvét, amelyen az egyedi statisztikai egységekre, közvetlenül statisztikai célokra vonatkozó, vagy közvetve, adminisztratív vagy egyéb forrásokból beszerzett bizalmas adatok védelmét kell érteni, amely magában foglalja a megszerzett adatok nem statisztikai célú felhasználásának és jogellenes nyilvánosságra hozatalának tilalmát is.

3. A bizalmas adat fogalma az európai statisztikában

A bizalmas adatnak mint védendő tárgynak fogalmát 223/2009/EK-rendelet határozza meg. Ennek alapján bizalmas adat minden olyan adat, amely lehetővé teszi a statisztikai egységek közvetlen vagy közvetett azonosítását, és ez által egyedi információt közöl. A statisztikai egység azonosíthatóságának megállapításakor figyelembe kell venni mindazokat az eszközöket, amelyeket egy harmadik személy ésszerűen igénybe vehet az említett statisztikai egység azonosításához. A jogszabály tehát a személyes adatok (egyedi információk) védelménél használt érintett helyett az adat alanyaként a statisztikai egységet nevezi meg, az ő érdekeinek védelme indokolja a bizalmas adatkezelést. A statisztikai egység a 223/2009/EK-rendelet szerint a megfigyelés alapegysége, azaz egy természetes személy, egy háztartás, egy gazdasági szereplő vagy egyéb vállalkozás, amelyre az adatok vonatkoznak. A statisztikai egység tehát nem minden esetben egyezik meg közvetlenül az adatszolgáltatóval, hiszen például adminisztratív adatátvétel esetén jellemző gyakorlat, hogy nem a megfigyelési egység szolgáltatja az adatokat a statisztikai hivatal felé, hanem az adminisztratív nyilvántartást vezető szerv.

Az adat alanyának meghatározása mellett a legfontosabb szempont, amelyet egy adott statisztikai adat bizalmas jellegének meghatározásánál vizsgálni szükséges, az azonosíthatóság kérdése. Az előbbi fogalom alapján, amennyiben egy adat lehetővé teszi a statisztikai egység azonosítását, azt bizalmas adatnak kell tekinteni és ennek megfelelően kezelni. Az azonosítás azt jelenti, hogy egy adott statisztikai egység a hozzá rendelt adatok alapján egyértelműen felismerhető, vagy hozzá lehet jutni a statisztikai egység egy vagy több közvetlen azonosítójához.

Az 223/2009/EK-rendelet 3. cikkének 9–10. pontja az azonosítás két különböző formáját különbözteti meg, melyeket a következőképpen definiál:

„A közvetlen azonosítás a statisztikai egység azonosítása annak nevével vagy címevel, illetve egy nyilvánosan hozzáférhető azonosító számáról.

Közvetett azonosítást jelent a statisztikai egység azonosítása bármilyen, a közvetlen azonosítástól eltérő módon.”

A gyakorlatban ez a vizsgált adatállomány vagy az adatkérő birtokában levő egyéb adatállomány két vagy több változójának együttes információja mint kulcs alapján levont következtetés segítségével történő azonosítást jelent. Mindezen túl, a bizalmas adatok esetében az azonosítás mellett adatvédelmi szempontból vizsgálni kell azt az esetet is, amikor nem csupán a statisztikai egység felismerése történik meg, hanem a statisztikai egységről olyan információ kerül nyilvánosságra, amely addig nem volt nyilvánosan elérhető, ismert. Ezt az információszerzést nevezi a statisztikai adatvédelmi módszertan felfedésnek, illetve az adatvédelem ilyen jellegű kockázatának kezelési módszertanát pedig felfedés elleni védelemnek.

A bizalmas adatok védelme két fontos célt szolgál. Egyrészt védi a statisztikai egységeket a jogosulatlan felfedés, jogellenes nyilvánosságra hozatal ellen, másrészt biztosítja, hogy a statisztikai célra felvett bizalmas adatokat kizárólag statisztikai célra használják fel.

4. A statisztikai cél kizárólagosságának elve

A bizalmas adatok védelmének érdekében a 223/2009/EK-rendelet a statisztikai cél kizárólagosságát már a preambulum bekezdései között kimondja, majd megismétli azt a bizalmas adatkezelés elvének tartalmában, de az adatvédelmi rendelkezéseket tartalmazó VII. fejezet első cikkében is, ezzel kezdve az adatvédelmi szabályok részletes meghatározását (20. cikk (1) bekezdés). A statisztikai cél kizárólagossága – ahogy azt a leírtak is rögzítették – a személyes adatok védelmére vonatkozó általános szabályokkal összhangban biztosítja, hogy a bizalmas adatok felhasználásával ne kerüljön sor a statisztikai egység vonatkozásában jogkövetkezmény megállapítására, azaz a bizalmas adatok ne legyenek felhasználhatók nem statisztikai célra, kiemelten hatósági, bírósági, egyéb hivatalos eljárásokban.

E főszabály alól egy kivételt enged meg a 223/2009/EK-rendelet, amely összefügg az információs önrendelkezési joggal. Ennek alapján lehetőség van a bizalmas adatok nem statisztikai célú felhasználására, amennyiben ahhoz a statisztikai egység egyértelmű beleegyezését adta. Ennek a beleegyezésnek a gyakorlat szerint előzetesnek és az adott célhoz kötöttnek kell lennie és csak az adott cél megvalósításához szükséges időre szólhat. Ez azt jelenti, hogy a hozzájárulás például csak meghatározott bírósági eljárásra vonatkozhat, nem lehet azt általánosan, bármilyen jövőbeli eljárásra vonatkozóan és időkorlát nélkül megadni.

Ugyanakkor a bizalmas adatokat még statisztikai célból sem lehet közzétenni, ezért a Gyakorlati Kódex és a 223/2009/EK-rendelet is kimondja a bizalmas adatok jogellenes közzétételének, nyilvánosságra hozatalának tilalmát.

5. Kivételek szabályozása a bizalmas adatkezelésben

A 223/2009/EK-rendelet bizalmas adatok nyilvánosságra hozatalára vonatkozó szigorú tilalom alól bizonyos esetekben mégis kivételt enged.

5.1. A statisztikai egység beleegyezése

Az egyik ilyen kivétel megegyezik a statisztikai cél kizárólagossága alól tett kivétellel, és lehetővé teszi a bizalmas adatok nyilvánosságra hozatalát akkor, ha ahhoz a statisztikai egység egyértelmű beleegyezését adta. Ennek a beleegyezésnek a formájára, tartalmára is az előbb leírtak vonatkoznak, azaz a beleegyezésnek esetinek, előzetesnek és célhoz kötöttnek kell lennie. Ezt a beleegyezést abban az esetben is az adatfelvételt végző tagállami statisztikai hivatalnak vagy más statisztikai hatóságnak kell megkérnie a statisztikai egységtől, ha az adat közlése az Eurostatnál történik. Ennek megfelelően a tagállamoknak a bizalmasságra utaló jelzéssel kell továbbítaniuk az adatokat az Eurostathoz.

A gyakorlatban számos esetben felmerül ugyanakkor, hogy bizonyos adatkörök esetében az beleegyezések nem nyújtanak megfelelő megoldást az adatközlés megvalósítására. Ezekben az esetekben olyan adatállományok érintettek, amelyeknél a statisztikai egységek száma, vagy az adat jellege miatt, a felfedési kockázatra tekintettel, akár országos területi bontás esetén sem lenne lehetőség adatközlésre. Annak érdekében, hogy a felhasználók érdekei se sérüljenek túlzottan, a gyakorlat jelenleg megoldást próbál kialakítani arra, hogy miként lehetséges már a statisztikai adatfelvétel során megadni az adatközléshez szükséges beleegyezést.

5.2. Passzív adatvédelem

A beleegyezés mellett lehetőség van a bizalmas adatok nyilvánosságra hozatalára az ún. passzív adatvédelem eseteiben is. A statisztikai adatvédelem alapvető megközelítése – utalva a leírtakra – az, hogy a bizalmas adatok nem adhatók át, nem hozhatók nyilvánosságra, kivéve, ha a statisztikai egység, az adat alanya ahhoz kifejezett hozzájárulását adta. A hozzájárulás, beleegyezés a statisztikai egység részéről aktív tevékenységet, ún. jognyilatkozat tételét kívánja meg.

Ahogy arra már utaltam, vannak ugyanakkor olyan adatkörök, melyek esetében az adatszolgáltatók összetétele, vagy az adatok jellege miatt még magas aggregáltsági szint mellett is jelentős a felfedési kockázat, ezért azok bizalmasnak tekintendők és ennél fogva felfedés elleni védelemmel ellátandók. Ezek az adatok

nem lehetnének közölhetők, de így az általuk közvetített információ nem jutna el a felhasználókhhoz, az ő érdekeik túlzottan sérülnének az adatvédelemhez fűződő érdek miatt.

Az európai jogalkotó ennek az érdekkonfliktusnak a feloldására vezette be egyes statisztikai területeken a passzív adatvédelmet, melynek esetén az előbbi mechanizmus megfordul, és a közvetett azonosítást lehetővé tevő bizalmas adatok közölhetők, amennyiben az adatközlés ellen a statisztikai egység nem tiltakozott.

A passzív adatvédelem alkalmazását európai uniós jogszabálynak kell elrendelnie. Jelenleg ezt a fajta szabályt a külkereskedelem-statisztika területén alkalmazzák csak, elsősorban a termékek külkereskedelmi forgalma esetében. Azt, hogy mely esetekben áll fenn a közvetett azonosítás révén történő felfedés kockázata, az adatot gyűjtő tagállami statisztikai hivatalok, vagy más statisztikai hatóságok határozzák meg. Ugyanakkor azt, hogy az egyes adatszolgáltatók vonatkozásában értelmezhető-e a kockázat, nem a statisztikai hivatal, hanem maga az adatszolgáltató mérlegeli, és dönti el, hogy ennek alapján kívánja-e, hogy a statisztikai hivatal az adatát felfedés elleni védelemmel lássa el. Erre való tekintettel a jogszabályi rendelkezés önmagában nem nyújt elég garanciát az adatszolgáltatóknak a passzív adatvédelem gyakorlati megvalósulása szempontjából.

Tekintettel arra, hogy a közvetett felfedési kockázat fennállását és az erre tekintettel történő, közzétételre vonatkozó tiltakozást magának az adatszolgáltatónak kell mérlegelnie, így elengedhetetlen, hogy az adatot szolgáltató statisztikai egység tisztában legyen ennek a speciális adatközlési szabálynak a meglétével és alkalmazásával, és a megfelelő tájékoztatás birtokában, ésszerű időn belül, írásban tiltakozhasson adatai nyilvánosságra hozatala ellen. Ennek alapján a KSH honlapján tájékoztatja az adatszolgáltatókat a szükséges lépésekről. Ezzel kapcsolatban hangsúlyozni kell, hogy az uniós adatközlések esetén sem az Eurostatnál kell tiltakozni, hanem a tiltakozási jog érvényesítésének lehetőségét még tagállami szinten kell biztosítaniuk a statisztikai hivataloknak és ennek megfelelően továbbítani az adatokat az Eurostat-hoz, megjelölve a bizalmasan kezelendő cellákat az adatállományon belül. A tiltakozás, csakúgy, mint a korábban említett beleegyezés, szintén nem lehet utólagos, általános és határozatlan időre szóló. A KSH általános szabályként az adatvédelmet meghatározott termékre, a kérelem benyújtásának ideje szerinti tárgyév adataira biztosítja. Az adott év február 15-ig beérkezett és elfogadott kérelmek alapján a védelem a januári adatoktól folyamatos. Az ezt követően, de még a naptári éven belül benyújtott, és az adatvédelmi követelményeknek megfelelő kérelmek esetén a KSH a tájékoztatási naptár szerint következő havi adatközléstől alkalmazza azt, a tárgyév visszamenőleges hónapjaira is. A naptári év végét követően, utólagos adatvédelmi kérelmek nem fogadhatók el.³

³ http://www.ksh.hu/docs/adatgyujtesek/intrastat/adatvedelem_2015.pdf

5.3. A bizalmas adatok cseréje az ESR és a KBER⁴ tagjai között

A bizalmas adatok továbbításának tilalma alóli másfajta kivételt jelent a 223/2009/EK-rendelet 21. cikkében foglalt rendelkezés, amely lehetővé teszi a bizalmas adatok cseréjét, átadását az ESR, illetve a KBER tagjai között.

Ezek az említett rendelkezések nem önmagukban állók, alkalmazásuk során figyelemmel kell lenni a jelen írás elején ismertetett 45/2011/EK-rendeletre, amely a személyes adatok közösségi intézmények közötti továbbításának szabályait rögzíti. Ennek megfelelően a bizalmas statisztikai adatok cseréje, továbbítása is csak célhoz kötötten történhet meg, azaz kizárólag akkor, ha az az európai statisztikák fejlesztése, előállítása, nyilvánosságra hozatala érdekében szükséges és indokolt. Az ilyen célból történő adatcserét több szempont is indokolja. Egyrészt az Eurostathoz történő adat-továbbítás elengedhetetlen az uniós szintű adatközlések megvalósításához, de egyes esetekben költséghatékonysági, statisztikai minőségi szempontok indokolhatják azt is, hogy tagállami statisztikai hivatalok között kerüljön sor ilyen adatátadásra. Egyes statisztikai területek, mint például a külkereskedelem-statisztika, ugyanazon adatszolgáltató tevékenységét több tagállam területén is megfigyelik, ami lehetővé teszi tükörstatisztikai vizsgálatok elvégzését, ezáltal javítva a statisztika minőségét. Az ilyen adatcsere nem jelent kockázattövekedést az adatszolgáltatók szempontjából, hiszen az adatok védelme – tekintettel a Gyakorlati Kódexben és az uniós jogszabályokban előírt egységes elvárásokra – minden ESR-tag esetében megfelelően biztosított, és kizárható a jogellenes nyilvánosságra hozatal vagy a statisztikai céltól eltérő célra való felhasználás veszélye.

Ennek ellenére, az ESR-en belüli adatcsere egyik legfontosabb kérdése éppen a célhoz kötöttség meglétének fennállása. A 223/2009/EK-rendelet kimondja, hogy amennyiben európai jogszabály úgy rendelkezik, az adatátadás kötelező és nem tagadható meg nemzeti jogszabályra való hivatkozással. Ilyen kötelező adatátadást eddig jellemzően csak az Európai Bizottság (Eurostat) felé írtak elő az egyes statisztikai területeket szabályozó jogszabályok.

Mérlegelési jogkörbe esnek ugyanakkor a már említett tagállami statisztikai hivatalok vagy egyéb statisztikai hatóságok közötti adatcserek. Ezek általában nem rendszerezettek, egy-egy eseti vizsgálatra vonatkoznak és így indokoltságukat is esetről esetre kell megvizsgálnia az adatátadó szervnek. Az adatátadás során biztosítani kell az átadott adatok megfelelő kezelését, ezért az adatátadás részletes szabályait az átadó és az átvevő intézmény közötti megállapodások szokták rögzíteni.

A másik kör, amely számára a jogalkotó biztosította a bizalmas adatcsere lehetőségét a KBER, melynek tagjai az Európai Központi Bank és a tagállami jegybankok. A KBER elkülönül az ESR-től, annak ellenére, hogy maga is jelentős statisztikai tevé-

⁴ KBER: Központi Bankok Európai Rendszere.

kenységet végez, és az általa előállított statisztikák is európai statisztikának minősülnek. A KBER statisztikai tevékenységére nem terjed ki a 223/2009/EK-rendelet, sem a Gyakorlati Kódex, hanem azt külön jogszabály szabályozza.⁵ Ugyanakkor a KBER statisztikai tevékenységének szabályozása az adatvédelmi garanciákat tekintve megegyezik a Gyakorlati Kódexben és a 223/2009/EK-rendeletben foglaltakkal, maga a jogszabály mondja ki, hogy a KBER által alkalmazott statisztikai elveknek meg kell felelniük a 223/2009/EK rendelet által meghatározott elveknek. Az adatsere vonatkozásában a KBER-re vonatkozó külön jogszabályi rendelkezések tükörszabályozásai az ismertetett, európai statisztikákra vonatkozó rendelkezéseknek, ezáltal biztosítva a bizalmas adatok megfelelő szintű védelmét mindkét rendszeren belül. A bizalmas adatok megismerésével kapcsolatos következő speciális esetkör a tudományos célú hozzáférés lehetősége, melyről a jelen cikkben külön fejezet szól.

6. Szervezeti, szabályozási elvárások az adatvédelem területén

Az adatok védelmére vonatkozó, ismertetett szabályok érvényesítése az egyes adatkezelők feladata, és ezzel kapcsolatban mind a Gyakorlati Kódex, mind a 223/2009/EK-rendelet elvárásokat fogalmaz meg az Eurostat, illetve az ESR többi tagja iránt. Mindezek mellett a személyes adatok kezelésére vonatkozóan továbbra is alkalmazni kell a 95/46/EK- irányelvben és a 45/2001/EK-rendeletben foglaltakat is.

Ennek megfelelően a 223/2009/EK-rendelet a 20. cikk (4) bekezdésében rögzíti, hogy saját megfelelő hatáskörükön belül az nemzeti statisztikai hivatalok és más nemzeti hatóságok és az Eurostat megteszik az összes szükséges szabályozási, adminisztratív, technikai és szervezeti intézkedést a bizalmas adatok fizikai és logikai védelmének biztosítása érdekében (felfedhetőség elleni védelem).

Ez a rendelkezés azért fontos, mert meghatározza az adatvédelem legfőbb területeit. Az adatvédelem ennek alapján jogi (szabályozási), módszertani, valamint informatikai és egyéb fizikai biztonsági intézkedések alkalmazásának összessége, melyek közül önmagában egyik sem elégséges, hanem kizárólag megfelelő szintű, együttes alkalmazása nyújt garanciát a statisztikai egységek jogainak érvényesítéséhez.

Ezt a rendelkezést fejti ki részletesebben a Gyakorlati Kódex az öt ismértve keretében.

I. A statisztikai adatok bizalmas kezelését jogszabályban kell biztosítani.

⁵ A Tanács 2533/986/EK-rendelete az Európai Központi Bank által végzett statisztikai adatgyűjtésről.

2. A statisztikai szervezet alkalmazottai a kinevezésükkel egyidejűleg titoktartási nyilatkozatot írnak alá.

3. A statisztikai adatvédelem bárminemű szándékos megsértése büntetést kell maga után vonnia.

4. Az adatok előállításának és közzétételének teljes folyamatában követendő adatvédelmi előírások és irányelvek állnak az alkalmazottak rendelkezésére. Az adatvédelmi politika mindenki számára nyilvános.

5. Fizikai, technikai és szervezeti óvintézkedések alkalmazása a statisztikai adatbázisok biztonságának és integrálásának védelmére.

Az ismérvek közül az első ismérvre már utaltam, annak megvalósulását uniós szinten a 223/2009/EK-rendelet, tagállami szinten pedig a nemzeti statisztikai törvények biztosítják. Tekintettel arra, hogy a jogszabályok nem tartalmazhatnak részletes, akár eljárási jellegű rendelkezéseket, ezért a Gyakorlati Kódex szerint elvárás (4. ismerv), hogy a statisztikai tevékenység valamennyi lépését, folyamatszakasát belső, eljárási jellegű szabályok kísérik, melyek gyakorlati útmutatásul szolgálnak a munkatársak számára. Ezek közül a legfontosabb az adatvédelmi politika, amely az adatvédelemmel, adatkezeléssel kapcsolatos alapvető elveket rögzíti az adatszolgáltatók, felhasználók számára megismerhető módon.

Az Eurostat esetében ezeknek a legfontosabb információknak a rögzítése az Eurostat honlapján történik meg.⁶ Az adatvédelmi politika mellett minden statisztikával foglalkozó szervezet, így az Eurostat is rendelkezik adatvédelmi szabállyal, emellett léteznek iránymutatások az egyes statisztikák felfedés elleni védelmével kapcsolatos intézkedésekkel, módszerekkel kapcsolatban.

Az Eurostat esetében érdemes kiemelni, hogy a tudományos célú hozzáférést szolgáló, anonimizált mikroadat-állományok anonimizálási módszertanát a tagállamok véleményének kikérésével alakítják ki. A jogszabályok és egyéb, belső szabályozó eszközök mellett a Gyakorlati Kódex 5. elvének második ismérve titoktartási nyilatkozat aláírását is megköveteli, melyek a jogszabályokban foglalt absztrakt szabályok tényleges adatkezelést végző személyek által történő megismerését és alkalmazását rögzítik formális módon, ezáltal közvetlen személyes felelősséget és kötelezettséget keletkeztetve. A titoktartási nyilatkozat aláírása nem csupán az Eurostat és a tagállami statisztikai hivatalok munkatársaitól elvárás, hanem az intézményekkel szerződéses kapcsolatban álló vállalkozók, egyéb szervezetek számára is kötelező. Az Eurostat belső szabályzata alapján minden dolgozó köteles ilyen titoktartási nyilatkozat aláírására, amelyben foglaltak megtartása jogviszonyuk megszűnését követően is kötelezi őket. A Gyakorlati Kódex 5. elvének 3. ismérve arra az esetre vonatkozik, amikor valaki megsérti a jogszabályban és egyéb szabályzatokban

⁶ <http://ec.europa.eu/eurostat/web/research-methodology/statistical-confidentiality>

foglalt kötelezettségeit. Ezekben az esetekben rendelkezésre kell állnia egy megfelelő szankciórendszernek, amely alapján büntethető a vétkes személy.

A szankciók meglétére vonatkozóan a 223/2009/EK-rendelet is tartalmaz előírást (26. cikk), amely azt is kimondja, hogy az előírt szankcióknak hatékonyaknak, arányosaknak és visszatartó erejűeknek kell lenniük. A megfelelő szankciók meglétét a tagállamoknak kell saját jogalkotásuk útján biztosítani. Az Eurostat esetében ezt a luxemburgi jog teszi meg, annak a szabályai alkalmazandók az Eurostat munkatársai által elkövetett esetleges jogsértések esetére. Ezek a jogi szankciók eltérő jellegűek lehetnek, elsőként a fegyelmi eljárás alkalmazandó, amely az uniós tisztviselők esetében a Bizottság alkalmazottaira vonatkozó egységes személyügyi szabályok szerint történik, de súlyos esetekben büntetőjogi szankció alkalmazására is sor kerülhet. Hangsúlyozni kell ugyanakkor, hogy a titoktartási nyilatkozatok aláírása nem jelenti azt, hogy bármely dolgozó bármilyen adathoz hozzáférhet. A hozzáférés kizárólag munkakörhöz kötve történhet, a megfelelő informatikai biztonsági intézkedések megtartása mellett, biztonságos környezetben.

Ehhez kapcsolódik a Gyakorlati Kódex 5. elvének ismérése, amely megköveteli a megfelelő informatikai biztonsági infrastruktúra és intézkedések meglétét az ESR tagjainál.

Az informatikai biztonságnak egyaránt biztosítani kell a szabályozási, személyi oldalát, azaz a megfelelő belső szabályzat és informatikai biztonsági felügyelő meglétét, illetve a fizikai oldalát, az adatbázisok megfelelő védelmét. A hozzáférés célhoz kötöttsége kapcsán informatikai biztonsági kérdés a szervezeten belül nyomon követhető, dokumentált hozzáférési jogosultságkezelés, ami biztosítja, hogy az egyes munkatársak vagy szerződéses partnerek csak a munkakörükben, feladatuk ellátásának céljából, a szerződésükben meghatározott adatokhoz férjenek hozzá. Az ismertetett szabályok és garanciák érvényesítése elengedhetetlen az adatszolgáltatók bizalmának a megszerzése és megőrzése érdekében. A jogszabályban és a Gyakorlati Kódexben megfogalmazott egységes elvárások emellett azt szolgálják, hogy az adatvédelem szintje minden ESR-tag esetében egyformán magas legyen, valamint az adathozzáférés gyakorlata szempontjából se legyenek nagy eltérések az Eurostat és az egyes tagállamok között.

7. A tudományos célú hozzáférés szabályozása az Európai Unióban

Az adathozzáférés vonatkozásában ki kell emelni, hogy az európai uniós jogi szabályozás is elismeri és hangsúlyozza, hogy a statisztikának az adatszolgáltatók

védelme mellett elsősorban mégiscsak a felhasználók érdekeit kell szolgálnia és a lehető legszélesebb hozzáférést kell biztosítani az ESR-en belül előállított adatokhoz. A 223/2009/EK-rendelet már a preambulum bekezdései között kimondja, hogy ebből a célból közös elveket és iránymutatásokat kell megállapítani az európai statisztikák előállításához használt adatok bizalmas kezelésének és az említett bizalmas adatokhoz való hozzáférésnek a biztosítására, megfelelően figyelembe véve a technikai fejlődést és a felhasználói igényeket egy demokratikus társadalomban. Az ESR igényei szerinti bizalmas adatok hozzáférhetősége különösen fontos az adatokból származó előnyök maximalizálása, az európai statisztikák minőségének növelése és az újonnan felmerülő közösségi statisztikai igényekre adott rugalmas válaszadás érdekében. Az európai tudományos haladás érdekében a kutatói közösségnek elemzés céljából szélesebb körű hozzáférést kell élveznie az európai statisztikák fejlesztéséhez, előállításához és közzétételéhez használt bizalmas adatokhoz. Ezért fejleszteni kell a kutatók bizalmas adatokhoz való tudományos célú hozzáférését, anélkül, hogy a bizalmas statisztikai adatok által igényelt magas szintű védelem veszélybe kerülne.⁷ Annak érdekében tehát, hogy a kutatói közösség a részletesebb adatok felhasználásával pontosabb következtetéseket tudjon levonni a gazdaság és a társadalom jelenségeiről, a statisztikai közösségnek át kellett térnie a kockázatkizárásra épülő adatkezelési gyakorlatról a kockázatkezelésre alapozott adathozzáférési csatornák kialakítására.

A 223/2009/EK-rendelet elvárását támasztja alá a Gyakorlati Kódex 5. elvének 5., illetve 15. elvének 4. ismérve is, mely megköveteli az ESR tagjaitól, hogy megfelelő biztosítékok mellett nyújtsanak hozzáférést tudományos célból kutatók számára az általuk kezelt bizalmas adatokhoz. A tudományos célú kutatói hozzáférés részleteinek szabályozása az egyes tagállamok esetében a nemzeti jogalkotó feladata, ezzel kapcsolatban uniós jogszabály rendelkezéseket nem fogalmaz meg. Az uniós szintű, Eurostat által biztosított adathozzáférés esetében ugyanakkor a 223/2009/EK-rendelet 23. cikkelye felhatalmazást ad a Bizottság részére, hogy végrehajtási jogszabályban, rendeletben szabályozza a tudományos célú hozzáférés részletes szabályait. Az ezzel kapcsolatos uniós jogi szabályozás már többéves múltra tekint vissza, ugyanakkor a technikai fejlődésre és a megnövekedett felhasználó igényekre tekintettel a korábban alkalmazott eljárást felül kellett vizsgálni abból a célból, hogy gyorsabb, rugalmasabb és az informatika által kínált új lehetőségeknek megfelelő legyen. Mindezekre tekintettel 2013-ban megszületett az a bizottsági rendelet, amely újraszabályozta a bizalmas adatokhoz történő tudományos célú hozzáférés szabályait és hatályon kívül helyezte az addig hatályos 831/2002/EK bizottsági rendeletet.

Az európai statisztikákról szóló 223/2009/EK európai parlamenti és tanácsi rendeletnek a bizalmas adatokhoz való tudományos célú hozzáférés végrehajtásáról és a

⁷ 223/2009/EK-rendelet (24)–(26) preambulum bekezdése.

831/2002/EK bizottsági rendelet hatályon kívül helyezéséről szóló 557/2013/EU-bizottsági rendelet 2013. július 8-án lépett hatályba. A rendelet megállapítja azokat a feltételeket, amelyek mellett a statisztikai elemzések és tudományos felhasználás céljából engedélyezni lehet a hozzáférést az Eurostat részére továbbított bizalmas adatokhoz, és a nemzeti statisztikai hatóságok között e hozzáférés megkönnyítése érdekében folytatott együttműködés szabályait.

A rendelet tárgyi hatálya a tudományos célból felhasznált bizalmas adatokra terjed ki, amelyek a bizottsági rendelet 2. cikkelyében foglalt fogalom-meghatározás szerint olyan adatok, amelyek kizárólag a statisztikai egység közvetett azonosítását teszik lehetővé. Ezzel kapcsolatban fel kell arra hívni a figyelmet, hogy még tudományos célból sem megengedett a közvetlen azonosítókkal ellátott adatállományok megismerése, vizsgálata.

A bizottsági rendelet alkalmazása során, a hozzáférés lehetséges körülményei szerint a bizalmas adatoknak két fajtája különíthető el. Az egyik csoport a biztonsági célú fájlok köre. Ezek olyan, tudományos célból felhasznált, az egyes statisztikai egységekre vonatkozó rekordok (mikroadatok), amelyek esetében nem alkalmazták kiterjedt felfedés elleni védelmet, azokról csupán a közvetlen azonosítókat távolították el. A tudományos célból megismerhető bizalmas adatok másik fajtája a tudományos célú fájlok köre. Ezeknek a mikroadat-állományoknak az esetében már nemcsak a közvetlen azonosítók kerültek törlésre, hanem azokat megfelelően anonimizálták is, azaz az állományokon további felfedés elleni védelmet is alkalmaztak annak érdekében, hogy – összhangban a mindenkor legjobb gyakorlattal – a statisztikai egységekről szóló információk felfedésének kockázatát megfelelő szintre csökkentsék.

Tekintettel arra, hogy a biztonsági célú fájlok esetében a felfedés kockázata nagy, ezért a bizottsági rendelet értelmében ezek csak megfelelően biztonságos környezetben ismerhetők meg. Ez a biztonságos környezet jelenleg az Eurostat luxemburgi kutatószobája. Ugyanakkor a bizottsági rendelet lehetőséget ad arra, hogy a jövőben – az előírt feltételeknek megfelelő, szigorú szakmai és biztonsági szempontokra egyaránt kiterítő akkreditáción átesett – egyéb adat-hozzáférési helyeken is lehessen ilyen állományokon kutatást végezni. Az európai statisztikákhoz hozzáférést nyújtó kutatószobákra vonatkozó elvárások, követelmények, infrastrukturális feltételek keretrendszerét már kidolgozták, de a részletszabályok megalkotása még folyamatban van. Az egységes feltételrendszer megszületését követően, elsőként a nemzeti statisztikai hivataloknak lesz alkalmuk arra, hogy sikeres engedélyezési eljárást követően saját kutatószobáikban nyújtsanak hozzáférést európai uniós mikroadat-állományokhoz. A kutatószobai hozzáférés legfontosabb jellemzője, hogy az adathozzáférés fizikai és adminisztratív intézkedésekkel védett környezetben történik, ahol nincs lehetőség az adatállomány elvételére, kijuttatására. A kutató a létesítmény szakemberei által felfedési kockázat szempontjából ellenőrzött kutatási eredményeket viheti el, így az adatszolgáltatók érdekei nem sérülnek, ugyanakkor a kutatási cél is megfelelően megvalósítható.

A tudományos célú fájlok esetében, az alkalmazott anonimizálási módszerek következtében a felfedési kockázat minimálisra csökken, így az ilyen állományok esetében nem szükséges a biztonságos környezet használatának előírása. A fájlokat úgy adják át a kutatóknak, hogy a kutatási cél megvalósulását követően megsemmisítik azokat.

A hozzáférés elvben bármilyen, az Eurostat kezelésében levő adatállományra kiterjed, a gyakorlatban ugyanakkor nem minden állomány hozzáférhető. Ennek oka, hogy hozzáférés csak végleges állományokhoz nyújtható, illetve az anonimizálási módszerek kidolgozása, egyeztetése a tagállamokkal hosszú időt vesz igénybe. Jelenleg tizenegy nagy felvétel mikroadat-állománya kutatható, de ezek köre folyamatosan bővül. A hozzáférhető állományok listája megtalálható az Eurostat honlapján (<http://ec.europa.eu/eurostat/web/microdata/overview>).

A hozzáférhető állományok mellett a bizottsági rendelet a hozzáférési eljárás tekintetében is számos szabályt, garanciát fogalmaz meg mind az adatszolgáltatók, mind az adatot az Eurostat részére továbbító nemzeti statisztikai hivatalok és egyéb nemzeti hatóságok számára. A bizottsági rendeletben meghatározottak alapján a hozzáférést kezdeményező kutatóhelyeknek kétlépcsős akkreditációs folyamaton kell átesniük.

Első lépésként magát a kutatóhelyet vizsgálja az Eurostat, mely vizsgálatnak a célja annak megállapítása, hogy az adott kutatóhely esetében biztosított-e minden feltétel az adatok megfelelő, biztonságos kezeléséhez. A bizottsági rendelet 4. cikkelye alapján az Eurostat a következő szempontokat mérlegeli az engedélyezés során:

- a)* a kutatási egység célja; az egység céljának értékelését az egység alapszabálya, küldetése vagy más szándéknyilatkozata alapján kell végezni; az egység céljai között szerepelnie kell a kutatási célnak is;
- b)* az egység hírneve vagy tekintélye a minőségi kutatás terén, amelynek eredményeit nyilvánosan hozzáférhetővé teszi; az egység kutatási projektekből szerzett tapasztalatának értékelését többek között az egység publikációiról rendelkezésre álló listák és azon kutatási projektek alapján végzik, amelyekben az egység részt vett;
- c)* a kutatást érintő belső szervezeti rendelkezések; a kutatási egység a kutatásra koncentrálni független, jogi személyiséggel rendelkező szervezet vagy egy szervezeten belüli kutatási osztály; a kutatási egységnek függetlennek, a tudományos következtetések megfogalmazásakor önállóan kell lennie, és el kell különülnie azon testület politikai területeitől, amelyhez tartozik;
- d)* az adatbiztonság érdekében rendelkezésre álló biztosítékok; a kutatási egység teljesíti az adatbiztonságot szavatoló műszaki és infrastrukturális követelményeket.

A mérlegelés a kutatóhely által kitöltött, nyilvánosan hozzáférhető kérdőív alapján történik, amelyhez csatolni kell a kutatóhely képviselője által aláírt titoktartási nyilatkozatot, melyben a kutatóhely elfogadja az adathozzáférés szabályait és kötelezi magát azok megtartására, valamint az adatok fizikai biztonságának biztosítására. A nyilatkozat valamennyi, az adatokhoz később hozzáférő kutatóra is kiterjed és meghatározza a hozzáférés feltételeit, a kutatók kötelezettségeit, a statisztikai adatok bizalmosságának tiszteletben tartása érdekében hozott intézkedéseket és e kötelezettségek megsértése esetében alkalmazandó szankciókat.

Amennyiben több kutatóhely végez közös kutatást, mindegyik szervezetnek külön akkreditáltatnia kell magát. A kutatóhely akkreditációs eljárása általában négy hetet vesz igénybe, és az elismert kutatóhelyek listáját az Eurostat a honlapján hozza nyilvánosságra. A kutatóhelynek haladéktalanul közölnie kell az Eurostattal, amennyiben a kérdőíven megjelölt adataiban változás következik be, ugyanakkor az Eurostat is rendszeresen felülvizsgálja az értékeléseket és dönthet úgy, hogy visszavonja a kutatóhely akkreditációját. Tényleges adathozzáférés iránti kérelmet csak az előbbieket szerint akkreditált kutatóhelyek munkatársai nyújthatnak be az Eurostathoz. A kérelem elbírálása az engedélyezési folyamat második lépcsője, melynek során a kutatási projektet, javaslatot értékeli az Eurostat, illetve a kutatási projekt által igényelt adatállományok oldaláról érintett tagországok. Adathozzáférési kérelmet csak a kutatóhellyel munkaviszonyban vagy egyéb szerződéses viszonyban álló kutatók vagy a kutatóhely által alkalmazott szakember által felügyelt egyetemi hallgató nyújthat be.

A kérelmet a kutatóhely akkreditációja során megjelölt kapcsolattartó személynek alá kell írnia, és ahhoz mellékelnie kell az adatokhoz hozzáférő kutatók által aláírt egyéni titoktartási nyilatkozatokat is.

A bizottsági rendelet 5. cikkelye meghatározza azokat a főbb információkat, melyet a kérelemnek, kutatási javaslatnak tartalmaznia kell:

- a)* a kutatás törvényes célja;
- b)* magyarázat arra vonatkozóan, hogy ez a cél nem bizalmas adatok felhasználásával miért nem teljesíthető;
- c)* a hozzáférést kérelmező egység;
- d)* az egyes kutatók, akik hozzáférnek az adatokhoz;
- e)* az alkalmazni kívánt, hozzáférést biztosító létesítmények;
- f)* azon adatkészletek és elemzésük módszerei, amelyekhez hozzáférést kérnek, valamint
- g)* a megjelentetésre vagy egyéb közzétételre szánt kutatás tervezett eredménye.

A kutatási projektek maximális hossza öt év lehet.

A kutatási kérelmek, javaslatok elsődleges vizsgálatát maga az Eurostat végzi el, majd azokat megküldi a kutatni kívánt adatot szolgáltató nemzeti statisztikai hivatal, vagy más statisztikai hatóság részére. Ez az eljárás jelentős garanciát biztosít a tagállami adatgazdák számára, hogy általuk nem támogatott célból vagy nem elfogadott személy, szervezet részére ne legyenek hozzáférhetők a tagállami bizalmas adatok.

A konzultáció időtartama általában négy hét, ugyanakkor van arra lehetőség, hogy ezt az időt a tagállamok és az Eurostat között kötött egyedi megállapodások útján lerövidítsék, vagy akár külön megállapodásban lefektetett feltételek fennállása esetén teljesen el is tekintsenek tőle.

Amennyiben a tagállami szervezet tiltakozik az adathozzáférés ellen, úgy az adott ország adatállományát az Eurostat a kutatásra előkészített mikroadat-állományból eltávolítja, és az nem lesz hozzáférhető a kutató számára. Az egyes kutatási projektek kapcsán nem kerül sor a kutatóhely és az Eurostat közötti szerződéskötésre, az Eurostat csupán az adatállomány felhasználási feltételeit adja át a kutatóknak. A megjelölt kapcsolattartó felelőssége, hogy ezeket az adatokhoz hozzáférő kutatókkal megismertesse.

Ezeknek az akkreditációs eljárásoknak részletes értékelési szempontjait az Eurostat az ESR többi tagjával közösen dolgozza ki, illetve közösen történik az alkalmazandó felfedés elleni védelmi módszerek megvitatása és véglegesítése is.

8. Jövőkép

A költséghatékony megoldások kialakítása érdekében a jövőben egyre szorosabb együttműködésre és a feladatok, felelőségek megosztására fog sor kerülni az Eurostat és a tagállamok között, ami javítja az előállított statisztikák minőségét és rugalmasabb, gyorsabb hozzáférést biztosít az európai statisztikákhoz anélkül, hogy veszélyeztetné a bizalmas adatok biztonságos kezelését.

A nemzeti szabályozásoknak tekintettel kell lenniük a statisztikai adatok iránti nemzetközi kutatói igényekre, és meg kell teremteniük annak lehetőségét, hogy kutatók külföldön is hozzáférhessenek egy adott ország adatállományaihoz, illetve, hogy az európai adatállományok valamennyi tagállamban elérhetők legyenek.

A technikai fejlődés és a kutatók igényei arra is ösztönzik a statisztikai hivatalokat, hogy olyan biztonságos távoli hozzáférési megoldásokat találjanak, amelyek nem követelik meg a kutatók fizikai jelenlétét a kutatószobákban. Mindezek során – ugyanakkor – is elsődleges célnak kell lennie az adatvédelem megtartásának és a folyamatos egyensúlykeresésnek az adathozzáférés és a bizalmas adatkezelés szempontjai között.

Summary

The article briefly summarizes the main rules concerning statistical confidentiality in the European legislation with special regard to the Eurostat's practice.

It describes the general regulations of the European Data Protection Directive, presents the principles of the European Statistics Code of Practice, the obligations deriving from the Regulation No 223/2009 on European statistics.

The author not only explains the basic definitions of statistical confidentiality (the difference between identification and disclosure, etc.) and the exceptions from confidentiality (such as passive confidentiality and the transmission of confidential data within the European Statistical System) but also provides a detailed description of access to microdata for scientific purposes.

Az információszabadság szerepe, jelentősége, aktuális kihívásai

Révész Balázs,

a Nemzeti Adatvédelmi
és Információszabadság
Hatóság főosztályvezetője

E-mail: revesz.balazs@naih.hu

Buzás Péter,

a Nemzeti Adatvédelmi
és Információszabadság
Hatóság vizsgálója

E-mail: buzas.peter@naih.hu

Szabad információáramlás, a közadatokhoz való korlátlan hozzáférés és az adatok újrahasznosítása a piaci szereplők által – laikus megközelítésben ez a lényegi tartalma az információszabadság fogalmának. A szerzők tanulmányunkban összefoglalják az információszabadság jelentőségét, sokrétű szerepét, kihívásait, ismertetnek néhány, napjainkban meghatározó ügycsoportot a Nemzeti Adatvédelmi és Információszabadság Hatóság gyakorlatából, és összefoglalják a jogterületet érintő aktuális jogszabályi módosításokat.

TÁRGYSZÓ:

Információszabadság.
Közadatok újrahasznosítása.
Közérdekű adatigénylés.

Gondoskodó vagy laissez-faire állam, központosító vagy decentralizáció törekvések, mindezek a kormányzati attitűdök is erősen befolyásolják az információszabadság helyzetét, jövőjét. Napjainkban megjelent a „Jó Állam” fogalma, „ami a közjó és a közszolgálati etika normáira támaszkodva szorosan összekapcsolódik a jó kormányzás és a jó közigazgatás fogalmával. Az állam és a kormányzat egyre növekvő felelőssége, valamint az egyre sokrétűbbé váló, egymással gyakran átfedő feladatok ellátásához szükséges integrált megközelítés gyakorlata egyre jelentősebb kapacitásokat, intézményi és adminisztratív képességeket igényel, amelyek megteremtése, karbantartása és folyamatos fejlesztése a mindennapi kormányzati gyakorlat szerves részének tekinthető.” (Kaiser [2015] 1. old.)

Az állam komplex feladatainak ellátása számos adat előállítását teszi szükségessé. Az információkat az adófizetők pénzéből hozzák létre, a nyilvántartásokat közpénzből és a közjó érdekében működtetik. Egyrészt az állam rendelkezésére álló adatok a nemzeti vagyon részét képezik, és ily módon modern felfogásban közkincsnek tekinthetők, a pontos naprakész adatbázisok a hatékony állam működéséhez elengedhetetlenek. Másrészt az állam nem mindenható, és nem is eshet kívül a társadalmi kontroll határain, tevékenységét – a politikai céloknak megfelelően – a választópolgárok felhatalmazása alapján határozza meg az aktuális kormányzat, működtetése közpénzből történik.

Az ideális jogállam garantálja magánszféránk nyugalma, személyes adataink védelmét, és ezek megsértése esetén az üzleti résztvevőket, sőt saját szerveit is szankcióval sújtja, ugyanakkor nyitott kapukkal várja az érdeklődő polgárokat, megosztja velük a működésével, gazdálkodásával kapcsolatos információkat, és végső soron csak a polgárai védelme, gyarapodása érdekében tart titokban információkat.

Kívánatos, hogy az állam számára a polgárok kevésbé, viszont számukra az állam teljes valóságában megismerhető legyen. Magánszféra és közszféra, magánjó és közjó, adatvédelem és adatnyilvánosság, ezek mögött az ellentétes fogalmak mögött egymástól eltérő érdekek jelennek meg. Ugyanakkor törekedni kell arra, hogy megtaláljuk azt az arany középutat, amely mindazon személyek elégedettségét szolgálja, akik e fogalmak kapcsán érintettek (Révész–Somogyvári [2012] 180–182. old.).

Állandóan változó világunkban a polgárok elvárása is folyvást alakul az állammal és annak szerveivel szemben, ami a kormányzati, önkormányzati prioritásokat szintén befolyásolja. Napjainkban egyre erősebb az emberek biztonság iránti vágya, amit a kormányzatok szerte a világban kihasználva központosító, adatgyűjtési célokra is felhasználnak. Mindez odavezethet, hogy az állam átláthatósága, illetőleg a polgárok magánszférájának védelme átértékelődik, és idealisztikus vágy marad.

1. Az információszabadság jelentősége és szerepe a korrupció elleni harcban

A világ számos országában a korrupció a mindennapi élet része, a lakosok és az országba látogató turisták, külföldi üzletemberek a kapcsolatépítés szükségszerű elemeként élük meg a csúszópénzek – sokszor leplezetlen – átadását. A korrupció a demokratikus államokban is jelen van, legfeljebb „intellektuálisabb” formában, kifinomultabb módszerekben ölt testet. Hosszú távon erkölcsileg és gazdasági hatásait illetően is súlyosan káros következményekkel jár, így az ellene való fellépésnek nemcsak utólagos büntetőjogi szankciókat kell magával vonnia, de a megelőzésére is nagy hangsúlyt szükséges fordítani.

A közérdekű adatok nyilvánossága hathatós eszköz a korrupció elleni harcban. Az ellene való fellépés sikeréhez hozzátartozik az elszámoltathatóság a jogalkotásban, a végrehajtásban és az igazságszolgáltatásban. Ehhez nyújt segítséget a nyilvános hozzáférés biztosítása a közinformációkhoz, amely jogintézmény alapvetően két bástyára épül: az elektronikus közzététel kötelezettségére és a közérdekű adatigénylés lehetőségére.

Az információszabadság, mint a közadatokhoz való hozzáférés alkotmányos biztosítéka, a magyar jogban ugyanazon törvényben – a személyes adatok védelméhez való jogunk garanciáival – nyert szabályozást. Sokszor szembesülünk azzal, hogy a korrupció elleni harcban, az információszabadság érvényesítésének jelszavát is zászlóra tűzve, az adatvédelmet korlátozó intézkedéseket szeretnének hozni, az átláthatóságot kiterjesztenék a polgár magánszférájára is. Fontos azonban megjegyezni, hogy bár a korrupció elleni küzdelemnek kétségtelenül elsőbbséget kell élveznie, e tény és az érdekében fogantatott intézkedések, bevezetendő szabályok nem lehetnek öncélúak, nem vezethetnek az alkotmányos alapjogok szükségtelen, valamint aránytalan korlátozásához.

2. Történeti előzmények, nemzetállami szabályozások

Először az 1766-ben elfogadott svéd sajtótörvény deklaráta, hogy az állampolgároknak joga van a hivatalos iratokba betekinteni (*Révész–Somogyvári* [2012] 180. old.).

A francia Emberi és Polgári Jogok Nyilatkozatának (1789) XIV. cikke pedig az állami kiadások ellenőrzését tette lehetővé a polgároknak, amikor kimondta, hogy saját személyükben vagy képviselő útján joguk van a karhatalom és a közigazgatás költségeinek fedezéséhez nélkülözhetetlen hozzájárulások felhasználását nyomon követni.¹ Komoly előrelépésnek számított ez akkoriban, hiszen a hadi költségek a kormányzati kiadások jelentős részét képezték.

¹ <http://mek.oszk.hu/00000/00056/html/228.htm>

Az Egyesült Nemzetek Szervezete Emberi Jogok Egyetemes Nyilatkozatának (1948) 19. cikke kimondta, hogy „Minden személynek joga van a vélemény és a kifejezés szabadságához, amely magában foglalja azt a jogot, hogy véleménye miatt ne szenvedjen zaklatást és hogy határokra való tekintet nélkül kutathasson, átvihesen és terjeszthessen híreket és eszméket bármilyen kifejezési módon”.²

Az információszabadság törvényi szabályozása azonban csak a XX. század második felében, az 1960-as évektől kezdődően terjedt el igazán világszerte. Kiemelkedő jelentőségű az Egyesült Államok által 1966-ban elfogadott Freedom of Information Act (az információszabadság törvénye). A jogszabály hatálya alá a szövetségi kormány által alakított vagy irányított „kormányügynökségek” tartoznak (*Bártfai et al.* [2005] 17. old.). E szervezetek egyrészt proaktív közzétételi kötelezettség terheli, másrészt bizonyos adatok tekintetében kötelesek – közzététel nélküli – folyamatos hozzáférhetőségét biztosítani, végezetül pedig egyes információkat csak kérelemre kell megismerhetővé tenniük. A törvény hatálya alá tartozó szervek díjat kérhetnek az adatok megismerhetővé tételéért, amelynek meghatározása során az igénylés céljára kell figyelemmel lenniük. 1996-ban a szövetségi törvényhozás módosította a jogszabályt, megteremtve az elektronikus információszabadság jogszabályi feltételeit az Egyesült Államokban (*Szilágyi–Jóri–Szabó* [2008] 2–4. old.).

Az Egyesült Királyságban 2000 novemberében fogadták el az információszabadságról szóló törvényt, amely több mint 100 ezer szerv adataihoz teszi lehetővé a hozzáférést. Az adatigénylést húsz munkanapon belül kell teljesítenie a megkeresett szervezetnek, azonban – méltányolható közérdekből – e határidő meghosszabbítható. A jogszabály érdekessége, hogy rendkívül hosszasan sorolja azokat az információkat, amelyek esetében nem, vagy csak korlátozottan érvényesül az információszabadság. Ezek az információk három nagy csoportra oszlanak (*Banisar* [2006] 155. old.). Az abszolút megismerhetetlen adatok kategóriájába tartoznak például a titkoszolgálatokkal kapcsolatos bírósági iratok. Az ún. minősített osztály kivétel (qualified class exemption) alá esnek például a kormányzati döntések meghozatalával, a nemzetbiztonsággal vagy a királyi kommunikációval kapcsolatos információk. Ezeket csak akkor lehet visszatartani, ha az azok titokban tartásához fűződő érdekek erősebbek a megismerésükhöz kapcsolódó közérdeknél. A harmadik kategóriába azon – például bűnmegelőzéssel, gazdasági kapcsolatokkal összefüggő – információk tartoznak, amelyek nyilvánosságra hozatala kárt okozhat. A kár veszélyének bizonyítása az adatot kezelő szervezetet terheli (*Bártfai et al.* [2005] 35–47. old.).

Németországban meglehetősen későn, 2005-ben fogadták el a szövetségi kormányzat adataihoz való hozzáférésről szóló törvényt. Azt megelőzően sokáig csak tartományi szinten, illetőleg bizonyos adatfajták – például a környezeti adatok – tekintetében szabályozták az információszabadságot (*Kerekes* [2004] 116. és 118. old.). A jogszabály lehetővé teszi bárki számára, hogy a szövetségi kormány szervei vagy a közfel-

² http://www.menszt.hu/informaciok/emberi_jogok/az_emberi_jogok_egyetemes_nyilatkozata

adatot ellátó szervek, illetve személyek kezelésében levő adatokat megismerhesse. Az adatigényléseket 30 napon belül kell teljesíteni. A törvény azonban számos kivételt is meghatároz, így például nem kell a döntések megalapozását szolgáló adatok vagy a minősített adatok megismerhetőségét biztosítani (*Banisar* [2006] 77. old.).

3. Információszabadság az Európai Unióban és az Emberi Jogok Európai Bíróságának gyakorlatában

Az információszabadság EU³-s szinten tulajdonképpen kettős értelemben érvényesül. Egyrészt Alapjogi Chartájának (2000) 11. cikke – hasonlóan a korábban említett nemzetközi emberi jogi dokumentumokhoz – a véleménynyilvánításhoz való jog keretében deklarálja az információszabadságot; továbbá 42. cikkében leszögezi, hogy „Bármely uniós polgár, valamint valamely tagállamban lakóhellyel, illetve létesítő okirat szerinti székhellyel rendelkező természetes vagy jogi személy jogosult hozzáférni az Unió intézményeinek, szerveinek és hivatalainak dokumentumaihoz, függetlenül azok megjelenési formájától”. Szemben tehát az Amszterdami Szerződéssel, a Charta az iratok megismerhetőségét már minden uniós intézményre kiterjesztette.

Másrészt az EU külön is lehetővé teszi a polgárai számára azt, hogy megismerhesse az uniós szervek kezelésében levő iratokat. Ebben a vonatkozásban kiemelkedő az Európai Unióról szóló szerződés és az Európai Unió működéséről szóló szerződés⁴ 15. cikke (az Európai Közösséget létrehozó szerződés korábbi 255. cikke), amely biztosítja az Európai Parlament, a Tanács és a Bizottság dokumentumaihoz való hozzáférés jogát. A kapcsolódó részletszabályokat az 1049/2001/EK rendelet⁵ határozta meg.

Az információszabadságnak – az EJEB⁶ ítélkezési gyakorlatának tükrében – hagyományosan a véleménynyilvánítás szabadságával van szoros kapcsolata. A bíróság számos esetben, így a Kenedi kontra Magyarország ügyben⁷ is azt állapította meg, hogy az adatok eltitkolása a kérelmezőt a véleménynyilvánítás szabadságának gyakorlásában akadályozta. *Kenedi János* a Belügyminisztérium felügyelete alatt működő Nemzetbiztonsági Szolgálat 1960-as években folytatott tevékenységéről kívánt tanulmányt írni. Ehhez kérte a minisztériumtól, hogy engedélyezzék az ott tárolt iratok egy részéhez a hozzáférést. A minisztérium megtagadta a kérelem teljesítését, ami miatt a kérelmező pert indított. A magyar bíróság az adatigénylőnek adott igazat, azonban a minisztérium többszöri bírság kiszabása ellenére sem biztosított hozzáfé-

³ EU: Európai Unió.

⁴ <http://eur-lex.europa.eu/legal-content/HU/TXT/HTML/?uri=CELEX:12012E/TXT&from=HU>.

⁵ <http://eur-lex.europa.eu/legal-content/HU/ALL/?uri=CELEX:32001R1049>.

⁶ EJEB: Emberi Jogok Európai Bírósága.

⁷ http://www.menschenrechte.ac.at/orig/09_3/Kenedi.pdf

rét a kért iratokhoz. Az EJEB az ügyben azt állapította meg, hogy az iratokhoz való hozzáférés megtagadása nem biztosította a kérelmező számára az Emberi Jogok Európai Egyezményének 10. cikk szerinti, a véleménynyilvánítás szabadságának jogát. Kifejtette továbbá, hogy az eredeti iratokhoz való hozzáférés korlátozásának ebből fakadóan meg kell felelnie a 10. cikk 2. bekezdésében foglalt feltételeknek. A strasbourgi testület figyelembe vette továbbá, hogy a Belügyminisztérium a magyar joggal ellentétesen járt el, amikor – a hazai bíróságok döntéseivel szembe menve – vonakodott átadni a kért iratokat. Az EJEB szerint ezért a minisztérium visszaélt a hatalmával akkor, amikor megakadályozta a kérelmezőt a véleménynyilvánítás szabadságának gyakorlásában. A strasbourgi bíróság következőképpen megállapította az Egyezmény 10. cikkének megsértését.

4. Az információszabadság „hivatalos óre”

A magyar adatvédelmi hatóság a személyes adatok védelme mellett az információszabadság érvényesüléséről is gondoskodni hivatott, követve az Alaptörvény rendszerét, amely az alapjogok sorában egyenrangúként ismerve el e két jog primátusát, mindkettőt egyazon cikkben szabályozza.⁸

Nem minden uniós tagállam követi ezt a gyakorlatot. Egyes országokban a független ombudsman/adatvédelmi hivatal, másutt különálló hivatal felügyeli a transzparencia ügyét.

Önmagában egyik intézményi megoldásra sem lehet azt mondani, hogy jobb vagy hatékonyabb mint bármely másik, annyi azonban bizonyos, hogy a két jog (a személyes adatok védelméhez és a közérdekű és közérdekből nyilvános adatok megismeréséhez fűződő jog) összeütközésének szakmai vitája itthon „házon belül” dől el.

Más területen azonban az adatvédelem és az információszabadság érvényesülésének esetleges konfliktusa kódolva van, és bekövetkezése szükségszerű. Az adatvédelmi hatóság kötelessége a közérdekű és közérdekből nyilvános adatok megismerhetőségének biztosítása, amely nem minden esetben találkozik a közpénzekből működő szervezetek vagy a közszférával szerződéses kapcsolatot létesítő üzleti szereplők spontán érdekeivel. Ők nem feltétlenül szeretnék annyi információt és ügy megosztani a közvéleménnyel, ahogyan azt az adatvédelmi törvény megköveteli. A jelenség nem magyar sajátosság, más jogállamban is előfordul, hogy az ügy résztvevői üzleti érdekeiket előbbre helyeznék az átláthatóság társadalmi ügyénél, és ilyenkor a független hatóságnak, illetve bíróságnak kell kimondania a végső szót.

⁸ Alaptörvény VI. cikk (2): „mindenkinek joga van személyes adatai védelméhez, valamint a közérdekű adatok megismeréséhez és terjesztéséhez”. http://net.jogtar.hu/jr/gen/hjegy_doc.cgi?docid=A1100425.ATV

5. Az információszabadság a 2010-es alkotmányreformot követően

Kodifikációs szempontból az információszabadság szabályozása a 2010-es alkotmányos reformnak köszönhetően nem hogy gyengült, hanem erősödött.

Az Alaptörvény bástyáján nyugodva az új adatvédelmi törvény⁹ (továbbiakban Infotv.) némely praktikus szabállyal kiegészülve segíti a közfeladatot ellátó szervezet, nyitva hagyva a nyilvánosság kapuját, emellett a jogalkotó a nemzeti vagyon nyilvánosságát szigorúbban deklarálja.¹⁰

A hatóság gyakorlatát az elmúlt években alapvetően a közpénzek átláthatóságának alaptörvényi szintre emelése, valamint a 2013-as évben megszületett alkotmánybírószági határozat¹¹ alakította, amely – a Nemzeti hitvallás alapvetéseire is hivatkozva – az információszabadságot gyakorlatilag a közélet tisztaságának elengedhetetlen feltételeként határozta meg.

Az állam a közérdekű adatok megismerhetőségét két módon biztosítja: egyrészt közzététel útján, másrészt azzal, hogy a közfeladatot ellátó szervezetek egyedi kérelemre adatot szolgáltatnak.

A közfeladatot ellátó szervezetek (minisztériumok, önkormányzatok, állami, önkormányzati cégek) a tevékenységükre, működésükre vonatkozó közérdekű adatokat ún. közzétételi listában meghatározott tartalommal hozzák nyilvánosságra a világhálón.

A közérdekű adatokhoz egyedi adatigénylés keretében is hozzá lehet jutni, azokat bárki, érdekeltsége igazolása nélkül megismerheti.

Az adatigénylésnek fő szabály szerint 15 napon belül közérthető formában és – amennyiben az aránytalan költséggel nem jár – az igénylő által kívánt technikai eszközzel, illetve módon kell eleget tenni. Amennyiben a kérelmező elektronikus úton igényelte az adatokat, úgy a választ is e-mailen keresztül kell teljesíteni. Ha a közérdekű adatot tartalmazó dokumentum az igénylő által meg nem ismerhető adatot is tartalmaz, azt a másolaton felismerhetetlenné kell tenni.

A közfeladatot ellátó szervezeteknek azonban nem minden közérdekű adatot kell nyilvánosságra hozniuk, kivételt képeznek a minősített adat védelméről szóló törvény¹² szerint tipizáltak. A korábbi állam- és szolgálati titok szerint osztályozó, némileg rutin-szerűnek is tekinthető minősítési eljárás helyett ma már differenciáltabb, négy szintű minősítési rendszert alkalmaznak az adatkezelők. A minősítés alkalmával figyelembe

⁹ Az információs önrendelkezési jogról és az információszabadságról szóló 2011. évi CXII. törvény. http://net.jogtar.hu/jr/gen/hjegy_doc.cgi?docid=A1100112.TV

¹⁰ A nemzeti vagyorról szóló 2011. évi CXCVI. törvény. <http://net.jogtar.hu/jr/gen/getdoc2.cgi?docid=A1100196.TV>

¹¹ Az Alkotmánybíróság 21/2013. (VII. 19.) számú határozata.

¹² A minősített adat védelméről szóló 2009. évi CLV. törvény. http://net.jogtar.hu/jr/gen/hjegy_doc.cgi?docid=A0900155.TV

kell venniük, hogy az adatok esetleges nyilvánosságra kerülése milyen súlyú károkat okozna. A dokumentumok, illetve az adatok a feltételezeten bekövetkező „rendkívül súlyos”, „súlyos” és egyéb kisebb súlyú hátrányokozás szerinti „szigorúan titkos”, „titkos”, „bizalmas”, illetve „korlátozott terjesztésű” minősítést kaphatnak.

Az adatfajták meghatározásával egyéb törvények is korlátozhatják az érintettek közérdekű és közérdekből nyilvános adatok megismeréséhez fűződő jogát, feltéve, hogy az honvédelmi, nemzetbiztonsági, bűnüldözési, környezet- és természetvédelmi, központi pénzügyi vagy devizapolitikai érdekből, külügyi, nemzetközi kapcsolatokra, bírósági vagy közigazgatási hatósági eljárásra tekintettel, illetve a szellemi tulajdon tiszteletben tartása jegyében indokolt, és az alkotmányossági kritériumokra is figyelemmel nem eredményez aránytalan jogkorlátozást.

Szintén bizalmasan kezelhető az olyan információ, amely a közfeladatot ellátó szerv döntésének megalapozását szolgálja (döntést megalapozó adat), és nyilvánosságra kerülése esetlegesen az adott szerv külső befolyásolástól mentes működését veszélyeztetné.

A nyilvánosság alól, korlátozott körben szintén kivételt képezhet az üzleti titok körébe tartozó információ. Üzleti titok a gazdasági tevékenységhez kapcsolódó minden olyan tény, információ, megoldás vagy adat, amelynek nyilvánosságra hozatala, illetéktelenek által történő megszerzése vagy felhasználása a jogosult – ide nem értve a Magyar Államot – jogszerű pénzügyi, gazdasági vagy piaci érdekeit sértené, illetve veszélyeztetné, és amelynek titokban tartása érdekében a jogosult a szükséges intézkedéseket megtette.

Azok a piaci szereplők azonban (egyéni vállalkozók, gazdálkodó szervezetek), akik az államháztartás valamely alrendszerének szereplőjével üzleti kapcsolatba kerülnek, kötelesek gazdasági tevékenységük adatainak nyilvánosságra kerülését eltérni, mégpedig olyan mértékig, hogy a közvagyonnal való gazdálkodás, a közpénzek felhasználása ellenőrizhető legyen.

6. A közérdekű adatok piaci szemszögből – a közadatok¹³ újrahasznosítása

A közadatok újrahasznosításában sok gyümölcsöző lehetőség rejlik. Az állami, önkormányzati szervek óriási mennyiségű értékes adattal rendelkeznek, de ez az adatvagyon csak részben hasznosul megfelelően. Az állami szervek ugyanis szűk keresztmetszetben, célzottan használják fel az adatokat, az államberendezkedés mű-

¹³ A közadatok gyűjtőfogalma a közérdekű és közérdekből nyilvános adatokat öleli fel (a közadatok újrahasznosításáról szóló 2012. évi LXIII. törvény 4. § 2. pont). <http://www.complex.hu/kzldat/t1200063.htm/t1200063.htm>

ködtetéséhez elengedhetetlenül nem, vagy kevéssé szükséges adatfeldolgozásban nem érdekeltek és ez nem is alapfeladatuk. A privátszektor résztvevői ellenben profitorientált szempontoktól vezérelve, a civilszféra pedig részben karitatív szerepét betöltendő, innovatív megoldásokban gondolkodik, olyan új ötletek megvalósításában, amelyek piacképesek. Míg az állami meteorológiai szolgálat, mely az időjárással, éghajlattal kapcsolatos adatok sokaságát „termeli”, honlapján a tájékoztatás jegyében ugyan megjeleníti az alapinformációkat, de ehhez nem teszi hozzá azokat a bővülő tartalmakat, amelyek a mindennapi életben még a honlap látogatóit érdekelheti. Más piaci résztvevő viszont, amely honlapja látogatottságának növelésében érdekelt, az alapadatokból új információkat hoz létre, más adatbázisokkal bővíti a tartalmat, a hasznosítás sokféle verzióját teremti meg, így a polgárok sokszínű információhoz jutnak, az internetes oldal „gazdája” pedig reklámbevételekre tesz szert.

Az Európai Bizottság régóta szorgalmazza a közadatok újrahasonosítását és ezt a célt prioritásként kezeli. A közzféra információinak további felhasználásáról szóló 2003/98/EK európai parlamenti és tanácsi irányelv (a továbbiakban: PSI¹⁴-irányelv) lehetőséget biztosít arra, hogy a közzféra által előállított egyes adatokat piaci-üzleti célból újrahasonosítsák. Az irányelv azonban számos európai országban sokáig nem hozott áttörést, így 2011-ben az Európai Bizottság Európai nyílt adatkezelési stratégiát alkotott, propagálva az újrahasonosításban rejlő lehetőségeket. *Neelie Kroes*, az Európai Bizottság alelnöke nyilatkozatában kifejtette, hogy az információk gyűjtése, előállítása az adófizetők pénzéből történik, így méltányolható elvárás, hogy az adatokat olyanok rendelkezésére bocsássuk, akik újszerűen, az emberek támogatására, a munkahelyteremtésre és a növekedés ösztönzésére alkalmas módokon használják tovább azokat.

Az Európai Bizottság megbízásából elkészített tanulmányok azt mutatták, hogy az uniós polgárok még mindig nehézségekkel szembesülnek a közzféra adatainak fellelése és további felhasználása során. Azaz a közadatok kezelése még mindig gyerekcipőben jár Európa nagy részén.

A geoinformatikai ágazat vonatkozásában a bizottság felmérésében résztvevők mintegy 80 százaléka úgy nyilatkozott, hogy akadályokba ütközik a közszervek által kezelt információk átfogó használata. Ennek okai az adatok további felhasználásával összefüggő magas díjakban, az átláthatatlan szabályozásban és gyakorlatokban keresendők, de nehézséget okoz az is, hogy nem világos, pontosan ki, milyen típusú adatokat kezel. Említhetők továbbá a kizárólagos licenzmegállapodások is, melyek a verseny aláaknázását okozhatják.

A bizottság által kezdeményezett európai digitális menetrendben a közzféra információinak további felhasználását – a nagy sebességű és szupergyors internet-

¹⁴ PSI (public sector information): közadat.

hozzáférés mellett – az egységes digitális piac megteremtésének egyik kulcsfontosságú tényezőjeként határozták meg.¹⁵

A PSI-irányelv végrehajtása jegyében megalkotott, a közadatok újrahasonosításáról szóló 2012. évi LXIII. törvény hosszas előkészületek után született meg, és sajnos a gyakorlati alkalmazása elég nehézkesnek bizonyul, ennek egyik oka lehet, hogy nehéz megtalálni az egyensúlyt a közérdekű adatokhoz való – alapvetően ingyenes – hozzáférés alapjogi biztosítása és az adatok díjazás ellenében, újrahasonosításra történő kiszolgáltatásának gyakorlata között. A közfeladatot ellátó szervek ugyanis kizárólag akkor járhatnak el a törvény rendelkezései szerint, ha a kérelmet benyújtó igénylő úgy nyilatkozott, hogy a közadatot újrahasonosítás céljából igényli. Nyilatkozat hiányában az Infotv.-re tekintettel a közfeladatot ellátó szerv az igénylés célját, illetve azt, hogy egy adott igénylés ténylegesen a közadat újrahasonosítására vonatkozik-e, nem vizsgálhatja. A két adatigénylésnek a célja alapvetően különböző, elhatárolásuk a gyakorlatban mégis fejtörést okoz a szervezeteknek: míg a közérdekű adatokhoz való alkotmányos alapjogunk biztosítása alapvetően azt teszi lehetővé, hogy bárki (polgár, civilszervezet, oknyomozó újságíró, politikus) két választás között is kontrollálhassa a közpénzek elköltését, a közszféra törvényes működését, addig a közadat újrahasonosítása érdekében foganatosított adatigénylések nagy mennyiségű adatra vonatkoznak és általában visszatérő rendszerességgel jelentkeznek, felhasználásuk nyereségorientált.

7. Meghatározó ügycsoportok a Nemzeti Adatvédelmi és Információszabadság Hatóság gyakorlatában

Számos alkalommal merül fel jogértelmezési dilemmaként az, hogy egy szerv közfeladatot ellátónak tekinthető-e, vagy sem. A tradicionális állami, önkormányzati hivatalok közfeladatot ellátóként való minősítése egyértelmű. Azon állami, önkormányzati cégek esetében azonban, akik piaci szereplőként is jelen vannak, kérdéses lehet, hogy meddig tart a nyilvánosság határa, és hol kezdődik az üzleti érdeké. E kérdések tisztázása érdekében a NAIH¹⁶ még 2012-ben ajánlást bocsátott ki.¹⁷

A jelenlegi jogszabályi környezet nemcsak arra helyezi a hangsúlyt, hogy valamely szerv, személy jogszabályban meghatározott tényleges közfeladatot lát-e el, hanem a nemzeti vagyonnal rendelkezés és gazdálkodás tényére is. Ezen rendelkezések és kö-

¹⁵ http://europa.eu/rapid/press-release_IP-11-1524_en.htm

¹⁶ NAIH: Nemzeti Adatvédelmi és Információszabadság Hatóság.

¹⁷ Ajánlás a nemzeti vagyonnal gazdálkodó vagy azzal rendelkező, piaci szereplő gazdasági társaságok üzleti adatainak nyilvánosságáról (http://www.naih.hu/files/Uzleti-titok_kontra_nyilvanossag_AJANLAS.pdf).

vetkeztetések figyelembe vételével tehát irrelevánssá vált az állami, illetve önkormányzati tulajdonban álló gazdasági társaságok azon „védekezése” a közérdekű adatigénylés teljesítésének kötelezettségével szemben, amely szerint ők nem közfeladatot ellátó szervnek minősülnek, mert nincs jogszabályban meghatározott feladatkörük.

Az állami tulajdonban levő cégek közzétételi kötelezettségével a TI¹⁸ Magyarország is foglalkozott egy kutatásában.¹⁹ A TI a cégek széles körében készített felmérést az interneten eddig közzétett adatokból, a cégek vezetőivel interjúkat készített, hogy felmérje a nyilvánosság melletti hajlandóság pozitívumait, illetve a nyilvánosság megteremtése során tapasztalt kihívásokat, a visszatartó erők körét, jelentőségét.

Mind a NAIH, mind a TI azt tapasztalta, hogy a nyilvánossággal egyáltalán foglalkozó állami, önkormányzati cégek még azt az alapvető elvárást is vitatják, amely szerint a közpénzzel való gazdálkodás egyenértékű a minimálisan korlátozott nyilvánossággal. Az előremutató alaptörvényi szabályozás követése a várt, remélt változásokat egyelőre nem igazolta. Míg a kizárólag versenyszférában tevékenykedő, magántulajdonban levő cégek a korrupció megelőzése, leküzdése mellett teszik le voksukat, és az internetes közzétételt az üzleti irányelv részévé teszik, az állami, önkormányzati cégek visszafogottan követik ezt a példát. A nyilvánosság kiterjesztése tényleges megvalósításának tehát elsődleges feltétele az, hogy a közpénzből gazdálkodó, valamint az abból alapított gazdasági társaságok elfogadják, hogy ők az információszabadság, mint alapjog kötelezettjei.

A technológiai eljárásokra, a műszaki megoldásokra, a gyártási folyamatokra, a munkaszervezési és logisztikai módszerekre, továbbá a know-how-ra vonatkozó adatokat leszámítva – esetről-esetre meghatározva – mérlegelni kell, az információk megismeréséhez és a megismerhetőség kizárásához fűződő közérdek súlyát. A nemzeti vagyonnal gazdálkodó cégek – amelyek piaci szereplők is egyben – „jogszerű pénzügyi, gazdasági vagy piaci érdeke” kerülhet szembe az információszabadsággal. Csak azok az adatok minősíthetők üzleti titoknak, amelyek illetéktelen nyilvánosságra hozatala vagy felhasználása a titok birtokosának érdek-, illetve jogsérelmet okoz. Egy adat üzleti titokká minősítésének az is elengedhetetlen feltétele, hogy „a titokbirtokos jogszerű érdekében álljon az adat megőrzése” (165/2011. (XII. 20.) AB-határozat). (*Eötvös Károly Közpolitikai Intézet* [2015])

A NAIH – többnyire az önkormányzatok kapcsán – szembesült azzal, hogy az egyedi közérdekű adatigényléseknek sok esetben úgy próbáltak gátat szabni, hogy a valóban felmerülő másolási költségek helyett indokolatlanul magas, több száz forin-

¹⁸ TI: Transparency International.

¹⁹ A Transparency International korrupcióérzékelési indexét (corruption perception index – CPI) 1995-ben alakították ki, amely a közszférában a korrupció érzékelt mértéke alapján rangsorolja a világ országait. A CPI egy összetett mutató, amely különböző független és elismert intézetek által készített szakértői felmérések korrupcióval kapcsolatos adataira épít, és a világon a legszélesebb körben használt. Lásd: http://www.transparency.hu/Korrupcio_Erzekelesi_Index.

tos másolási költséget határoztak meg. A hatóság ezeket a törekvéseket elítélte, és minden esetben kimondta azok jogellenességét, valamint felszólította a jogsértő önkormányzatot, hogy a költségeket reális mértékűre mérsékelje. Ugyanígy sarkalatos kérdés a terjedelem megítélése, hiszen az az adatigénylés teljesítésére álló határidő meghosszabbítására adhat alapot.

Más ügyekben a közfeladatot ellátó szervek az adatigénylő kilétét firtatnák, de az alkotmányos alapjog esetében a kérelmező bárki lehet. Ugyanígy nem vonható górcső alá a kérelmező adatigénylésének célja, ami tetszőleges és nem befolyásolhatja az adatigénylés sikerét. Az adatigényléssel kapcsolatos azon kritikákra pedig, hogy az adatigénylő gyakorlatilag nem a közérdekű adatokat kívánja megismerni, hanem csupán joggal való visszaélést valósít meg, a NAIH elnöke egy állásfoglalásban²⁰ akként reflektált, hogy az alapvető jog érvényesülési határainak meghatározása, korlátainak kijelölése, valamint a joggal való visszaélés megállapítása nem a hatóság feladata. Mindentől függetlenül a közérdekű adatok közérdekűek maradnak akkor is, ha egy valaki vagy többen és többször igénylik, illetve ha sok vagy jelentős terjedelmű adat az adatigénylés tárgya. A joggal való visszaélés megállapítása az eset összes körülményének mérlegelése után a bíróság hatáskörébe tartozhat, azonban figyelembe kell venni, hogy az információszabadság nem polgári jogi igény, hanem alapvető jog. Az Infotv. a joggal való visszaélésre tekintettel alapjog-korlátozást lehetővé tevő rendelkezést nem tartalmaz. A közérdekű adatigénylő adatai csak annyiban kezelhetők, amennyiben az az igény teljesítéséhez és a másolatkészítésért megállapított költségterítés megfizetéséhez szükséges. Az igénylő személyével kapcsolatos egyéb adatok, azokból levonható következtetések nem kezelhetők, azokat a közfeladatot ellátó szerv az adatigénylés teljesítése során nem veheti figyelembe. Természetesen a polgármesteri hivatalnak jogában áll az adatigénylővel az általa megadott elérhetőségeken egyeztetni, egyrészt felkérheti az igénye pontosítására akkor, ha az nem egyértelmű,²¹ illetve felhívhatja a figyelmét arra, hogy az adatigénylést másolatkészítést nem igénylő módon is teljesíteni tudja.²²

Ha az adatigénylés olyan új adat előállítására irányul, amely egyszerű számítástechnikai műveleteken, leválogatásokon, adatgyűjtéseken, összegzéseken túlmutató, bonyolult programozást igénylő módon teljesíthető, akkor az adatigénylést el lehet utasítani arra való hivatkozással, hogy az új adat előállítására az adatkezelő nem köteles, mivel az Infotv. meghatározása alapján a kezelésében levő közérdekű vagy közérdekből nyilvános adat lehet az adatigénylés tárgya.

Az információszabadság szempontjából és a nyilvánosság eredeti jogalkotói céljával ellentétes tendencia rajzolódott ki azokból a közérdekű adatigénylésekből, melyek egyes közfeladatot ellátó szervek, sőt szakmák teljes állományára tételes adatké-

²⁰ NAIH-1361-2/2014/V számú ügy.

²¹ Infotv. 28. § (3) bekezdés.

²² Infotv. 29. § (4) bekezdés

réseket fogalmaztak meg. A Klebelsberg Intézményfenntartó Központnak címzett, mintegy 140 ezer foglalkoztatottat érintő adatigénylés kapcsán született hatósági állásfoglalás²³ emlékeztetett arra, hogy a személyes adatok nyilvánossága csak a közfeladat ellátása körében és alkotmányosan csak olyan mértékig indokolt, amely e cél eléréséhez szükséges. Az Infotv. szellemiségével nem az ilyen, az Alkotmánybíróság által „készletező adatkezelésnek” minősülő személyes adatok megismerése áll összhangban. Az adott szerv gazdálkodását az állampolgárok anélkül is képesek ellenőrizni, hogy a közalkalmazottakat megalázó helyzetbe hoznák, és belőlük ellenérzést váltsanak ki az ilyesfajta listázásokkal. A cél megvalósítható a lista név nélküli kiadásával, illetőleg összesített adatok szolgáltatásával is.

A pedagógusok magánszférájának ilyen mértékű korlátozása, egy teljes szakma listázása, az érintettek szakmai életének, bérezésének effajta feltérképezése egyenlőtlen helyzetbe hozhatja őket a munkaerőpiacon is. A közszférában foglalkoztatottak ilyen listázása lehetőséget teremthet az ún. zsákmányrendszer kialakítására is, amely a szakmai közigazgatással szemben a politikai elkötelezettség látszatát sugallja.

8. Az Infotv. 2015. évi módosítása és annak várható hatásai

Az Országgyűlés a 2015. július 6-i ülésnapján fogadta el az információs önrendelkezési jogról és az információszabadságról szóló 2011. évi CXII. törvényt, továbbá egyes más törvények módosításáról szóló 2015. évi CXXIX. törvényt. A módosítás oka, hogy „egyes esetekben az Infotv. rendelkezései nem kellően rugalmasak”, amelyek így a közfeladatot ellátó szervek számára sokszor többletterheket rónak. Ez pedig „mind alapfeladataik ellátását, mind a közérdekű adatok megismerése iránti igények maradéktalan teljesítését akadályozhatják”.²⁴ A jogszabály ezért jelentős és érdemi újításokat vezetett be a közérdekű és közérdekből nyilvános adatok megismerése iránti igényekhez kapcsolódó eljárási szabályok, így különösen az igénylés teljesítésére vonatkozó előírások tekintetében. A jogalkotó az új előírások révén kívánta továbbfejleszteni a hatályos törvényi szabályokat, amelyek egyszerre erősíthetik az információszabadság védelmének garanciáit, ugyanakkor figyelembe veszik a közfeladatot ellátó szervek vagy személyek érdekeit is.

Noha az új előírások csak 2015. október 1-jén léptek hatályba, egyes rendelkezések máris heves vitát váltottak ki. A szakmai és civil szervezetek számos fórumon kifejtet-

²³ NAIH-2774-2013-V. számú ügy.

²⁴ Lásd az információs önrendelkezési jogról és az információszabadságról szóló 2011. évi CXII. törvény, továbbá egyes más törvények módosításáról szóló T/5404. számú törvényjavaslat indokolását. <http://www.parlament.hu/irom40/05404/05404.pdf>

ték fenntartásaikat az adatigénylés új szabályaival kapcsolatban, amelyek szerintük az információszabadság alapjogának aláásását célozzák.²⁵ A hatóság által a témában kibocsátott állásfoglalások, valamint Péterfalvi Attila nyilatkozatai ugyanakkor arra engednek következtetni, hogy a gyakorlat továbbra is megszorító, azaz elsődlegesen az állampolgárok jogait és érdeket fogja szem előtt tartani.²⁶ A módosuló szabályok elemzése során kitérünk a jogalkotó, a civil szervezetek és a hatóság álláspontjára.

Új szabályként a törvény kimondja, hogy a közfeladatot ellátó szerv az adatigénylésnek – a beérkezését követő legrövidebb idő alatt, legfeljebb azonban 15 napon belül – köteles eleget tenni.²⁷ Korábban a válaszadási határidő kezdete az érintett szerv vagy személy általi tudomásszerzés volt, amely „logikus és méltányos”, hiszen így például a hivatali időn túl beadott kérelmek teljesítésére nyitva álló intervallum csak a következő munkanaptól kezdődött.²⁸ E módosítás azonban objektív időpontot határoz meg a teljesítésre nyitva álló időintervallum kezdeteként, amely jelentős módon befolyásolhatja az állampolgár jogorvoslati lehetőségeit.²⁹

A technikai jellegű módosítások mellett érdemi változások is születtek az új szabályok szerint az adatigénylések teljesítésével összefüggésben. Ezek alapvetően az adatok megismerésének korlátait, illetőleg a teljesítésért kérhető költségtérítést érintik.

Jelentős újítás, hogy az adatigénylést nem kell teljesíteni abban az esetben, ha azt az igénylő, egy éven belül, azonos adatkör tekintetében nyújtja be.³⁰ E szabályt azonban csak akkor lehet alkalmazni, ha az adatokban nem állt be változás. Az említett módosítással kapcsolatban számos kritika fogalmazódott meg a civil szervezetek részéről.³¹ Lehetséges problémaként merült fel az az eset, amikor az adott közfeladatot ellátó szerv korábban semmilyen formában nem teljesítette az állampolgár által benyújtott adatigénylést. Ha ugyanis az érintett ismételt kérelmet nyújt be, akkor – pusztán az Infotv. szövegéből kiindulva – az adatkezelő az igény teljesítése elutasításának indokaként hivatkozhatna e rendelkezésre.³² Hasonló aggályok merülhetnek fel akkor is, ha a szerv megtagadta a teljesítést.

A hatóság egy újságíró kérdéseire válaszul kifejtette álláspontját a kérdéses szabállyal kapcsolatban. Eszerint nem jogsértő az éves limit bevezetése az adatigénylések teljesítése tekintetében.³³ Az említett rendelkezést azonban csak akkor lehet alkalmazni, ha az állampolgár adatigénylését egyszer már teljesítették. Ilyenkor – az

²⁵ <http://blog.atlatszo.hu/2015/06/tiltakozunk-az-informacioszabadsag-ujabb-korlatozasa-ellen>

²⁶ <http://nepszava.hu/cikk/1066012-peterfalvi-nem-kegyet-gyakorolnak-a-nyilvanossaggal>

²⁷ Infotv. 29. § (1) bekezdés.

²⁸ <http://atlatszo.hu/2015/07/16/matol-hatalyos-az-infotorveny-modositasa-osszeszedtuk-hogy-mi-valtozott>

²⁹ Vö. Infotv. 31. §.

³⁰ Az információ önrendelkezési jogról és az információszabadságról szóló 2011. évi CXII. törvény, továbbá egyes más törvények módosításáról szóló 2015. évi CXXIX. törvény 5. § (2) bekezdés.

³¹ <http://atlatszo.hu/2015/06/28/adatigenyles-ellenes-torvenymodositások-ujabb-pofon-az-informacioszabadsagnak>

³² <http://blog.atlatszo.hu/2015/06/tiltakozunk-az-informacioszabadsag-ujabb-korlatozasa-ellen>

³³ NAIH/2015/4710/2/V. állásfoglalás.

adminisztratív terhek csökkentése érdekében – az ismételten benyújtott kérelemre válaszul csupán fel kell hívni az érintett figyelmét e tényre. Amikor viszont a közfeladatot ellátó szerv „hallgatása” vagy az igénylés teljesítésének elutasítása miatt nem került sor a közérdekű, illetve közérdekből nyilvános adatok megismerésére, az új rendelkezésre nem lehet jogszerűen hivatkozni. Ellenkező esetben az adott szerv nyilvánvalóan jogsértő módon fog eljárni.³⁴ A kifogásolt rendelkezés „visszaélészerű” alkalmazására így nem lesz lehetőség a jövőben.

A közfeladatot ellátó szervek részéről számos esetben fel szokott merülni az a gond, hogy teljesíteni kell-e egy olyan adatigénylést, amit névtelenül, például elektronikus levélben adtak be. Ilyenkor a legváltozatosabb érveket szokták felhozni annak alátámasztására, hogy milyen kockázatokat rejt magában az anonim igénylés intézménye. A hatóság eddigi gyakorlata szerint ugyanakkor – figyelemmel az Infotv. 28. § (2) bekezdésére – csak olyan információkat lehetett felhasználni az igénylővel kapcsolatban, amelyek közvetlenül kapcsolódnak az adatigénylés teljesítéséhez. Ilyennek minősült például a név és lakcím papíralapú adatigénylés esetén. Az elektronikusan benyújtott kérelemnél pedig elegendő volt csupán az érintett e-mail címének kezelése.

Az Infotv. módosítása ebben a tekintetben is változtat az eddigi szabályokon. Az új rendelkezés szerint ugyanis nem kell teljesíteni az adatigénylést, ha az igénylő nem adja meg nevét, valamint azt az elérhetőséget, amelyen számára az adatigényléssel kapcsolatos bármely tájékoztatás és értesítés megadható.³⁵ A módosító törvény indokolása szerint az említett „minimális adatok” ismeretének hiányában az adatigényléssel kapcsolatban bizonyos eljárási cselekmények – például pontosításra való esetleges felhívás vagy a határidő meghosszabbításáról szóló tájékoztatás –, illetve maga az igénylés vagy annak elutasítása sem lennének teljesíthetők.

A hatóság – a már említett állásfoglalásában – ezzel összefüggésben leszögezte, hogy az érintettek személyes adatainak kezelése tekintetében is restriktív gyakorlatot fog követni. Az elektronikus úton, csak egy név és e-mail cím feltüntetésével benyújtott adatigényléseket ugyanis minden további nélkül meg kell válaszolnia az adott közfeladatot ellátó szervnek. Ilyenkor további adatok megadására az állampolgár nem kötelezhető, ami a gyakorlatban azt jelenti, hogy anonim adatigénylést is be lehet nyújtani, például egy kitalált név és egy absztrakt elektronikus levelezési cím segítségével. Továbbra is tilos lesz viszont az érintettek személyazonosságának ellenőrzése, azaz a személyi okmányok másolatának benyújtása, vagy ebből a célból a személyes megjelenés a jövőben sem kérhető az állampolgároktól.³⁶

³⁴ A hatóság ugyanakkor minden esetben arra buzdítja a közfeladatot ellátó szerveket, hogy ha egy adott adatkörre sorozatosan több adatigénylés is érkezik, akkor – figyelemmel az Infotv. 30. § (2) bekezdésére – azokat a honlapjukon tegyék közzé.

³⁵ Az információs önrendelkezési jogról és az információszabadságról szóló 2011. évi CXII. törvény, továbbá egyes más törvények módosításáról szóló 2015. évi CXXIX. törvény 5. § (3) bekezdés.

³⁶ NAIH/2015/4710/2/V. állásfoglalás.

Az Infotv. módosítás talán legnagyobb vitát kiváltó újítása az adatigénylés teljesítéséért kérhető költségtérítésre vonatkozó rendelkezés. A civil szervezetek hevesen támadták e rendelkezést, mivel álláspontjuk szerint a törvénybe olyan bizonytalan tartalmú jogfogalmak kerültek, amelyek lehetővé tennék az önkényes jogértelmezést.³⁷ Ez pedig végső soron az információszabadság mint alapjog aránytalan korlátozását jelentené.

Bizonyos kiadások adatigénylőre terhelésére már eddig is lehetőséget biztosított a törvény. Vonatkozó részletszabályok hiányában azonban alapvetően a bírói és a hatósági gyakorlat érvényesült e tekintetben. A hatóság eddigi álláspontja szerint a közfeladatot ellátó szervek az adatigénylések teljesítésekor nem szolgáltatást nyújtanak, hanem az Alaptörvényben meghatározott alapvető jogból eredő kötelezettségeiket teljesítik. E szervek nem gazdasági tevékenységük körében, üzletszerűen értékesítik a másolatokat, hanem lehetőségük van a felmerült anyagköltségek megtérítését kérni az adatigénylőtől. Mivel a közfeladatot ellátó szerv köteles a közérdekű adatok kiadása előtt az abban szereplő személyes adatokat felismerhetetlenné tenni, a dokumentumok anonimizálása a másolással kapcsolatban felmerült költségnek minősül. A másolás költsége azonban csak alacsonyabb lehet, mint a piaci ár, mivel sem általános forgalmi adót, sem munkadíjat, energiaköltséget, amortizációs elemeket nem tartalmazhat.³⁸

Az Infotv. módosítása a költségtérítés lehetőségét jelentősen átalakítja, kibővíti. Az új előírás szerint a közfeladatot ellátó szerv a költségtérítés mértékének meghatározása során figyelembe veheti – az anyagköltségen túl – a kézbesítés, illetőleg a munkaerő-ráfordítás mértékét is, amennyiben az igénylés teljesítése a közfeladatot ellátó szerv alaptevékenységének ellátásához szükséges munkaerőforrás aránytalan mértékű igénybevételeivel jár.³⁹ A módosító törvényhez fűzött indoklás szerint bizonyos esetekben már annak megállapítása is érdemi munkaerőforrás ráfordítását igényli, hogy „az igényelt adatok rendelkezésre állnak-e, és azok közérdekű vagy közérdekből nyilvános adatként megismerhetők-e, de azon igénylések is idetartoznak, amelyek megfelelő teljesítéséhez nagyszámú adathordozó vagy adatbázis áttekintése és a kért adatok leválogatása szükséges”.⁴⁰ Ezért az ilyen esetekhez a törvény jogkövetkezményként a költségtérítés megfizetését fűzi. A költségelemek megállapítható mértékét pedig kormányrendelet határozza meg.

Szintén sokat vitatott rendelkezés az, amelynek értelmében a döntés megalapozását szolgáló adat megismerésére irányuló igény a döntés meghozatalát követően

³⁷ <http://blog.atlatszo.hu/2015/06/tiltakozunk-az-informacioszabadsag-ujabb-korlatozasa-ellen>

³⁸ www.naih.hu/files/Infoszab_jelentes-NAIH-2015-2141-5-V-2015-07-13.pdf

³⁹ Az információs önrendelkezési jogról és az információszabadságról szóló 2011. évi CXII. törvény, továbbá egyes más törvények módosításáról szóló 2015. évi CXXIX. törvény 5. § (8) bekezdés.

⁴⁰ Lásd az információs önrendelkezési jogról és az információszabadságról szóló 2011. évi CXII. törvény, továbbá egyes más törvények módosításáról szóló T/5404. számú törvényjavaslat indokolását.

akkor is elutasítható, ha az adat további jövőbeli döntés megalapozását is szolgálja.⁴¹ Amennyiben tehát a kérdéses információ olyan adatösszesség része, amelynek egyes elemei tekintetében már döntés született, de más elemei még további döntés tárgyai lesznek, ez utóbbi adatok tekintetében az első döntés meghozatala után nem érvényesül a nyilvánosság. Annak eldöntése, hogy melyik adat kapcsolódhat adott esetben egy további döntéshez, nem teljesen egyértelmű. Egy információ gyakorlatilag számtalan további tevékenység vagy döntés alapjául szolgálhat, az Infotv. szövegének ilyen tág értelmezése azonban szintén aránytalanul jogkorlátozó lenne.⁴² Ugyanakkor a hatóság állásfoglalásaival hozzájárulhat a megfelelő gyakorlat kialakulásához.

Végezetül kiemelendő, hogy az Infotv.-nyel párhuzamosan módosulni fog a szerzői jogról szóló 1999. évi LXXVI. törvény, amelynek új rendelkezése értelmében a szerzői jogi védelem alatt álló művet csak megtekintés révén lehet közérdekű vagy közérdekből nyilvános adatként megismerni.⁴³ Az új szabály célja a személyhez fűződő jogok védelme, amelynek érdekében az érintett szerv – az adatigénylő által kívánt forma és mód helyett – lehetővé teheti a hivatali helyiségében az adott műbe történő betekintést és az arról való feljegyzés készítését. Ezen előírás tehát tulajdonképpen nem érinti az információszabadság tartalmát, csupán az adatigénylés módjának teljesítését határozza meg szűkebben.

Összességében elmondható, hogy az Infotv. módosítása jelentős változásokat hoz a közérdekű adatigénylések terén. Az új rendelkezések ugyanakkor önmagukban nem üresítik ki az információszabadságot mint alapjogot, közérdekű adatigényléssel így ezután is bárki a közfeladatot ellátó szervekhez fordulhat.⁴⁴ A vonatkozó gyakorlat kialakításában pedig kiemelkedő szerep jut majd a NAIH-nak, amely az igénylők és a közfeladatot ellátó szervek között egyfajta mediátorként, illetőleg – bizonyos esetekben – döntőbíróként továbbra is kiemelt szerepet fog játszani.

9. Összegzés

Amint jelen tanulmányból is kitűnik, az információszabadság társadalmi jelentősége vitathatatlan. Hatása, az adatvédelmi tárgyú ügyekkel összevetve, ha a médiában is megjelenő, és közvéleményt formáló meghatározó ügyeket is értékeljük, egyre jelentősebb. Határait ugyan folyvást változó világunkban néha óvatosan igazítani

⁴¹ Az információs önrendelkezési jogról és az információszabadságról szóló 2011. évi CXII. törvény, továbbá egyes más törvények módosításáról szóló 2015. évi CXXIX. törvény 17. § (4) bekezdés.

⁴² <http://atlatszo.hu/2015/06/28/adatigenyles-ellenes-torvenymodositások-ujabb-pofon-az-informacioszabadsagna>

⁴³ Az információs önrendelkezési jogról és az információszabadságról szóló 2011. évi CXII. törvény, továbbá egyes más törvények módosításáról szóló 2015. évi CXXIX. törvény 23. §.

⁴⁴ <http://atlatszo.hu/2015/07/16/matol-hatalyos-az-infotv-modositasa-osszeszedtuk-hogy-mi-valtozott>

kell, de alkotmányos elvi alapjait érintetlenül kell hagyni ahhoz, hogy a rendszerváltáskor kivívott szabadságunk értékeit megőrizhessük.

Irodalom

- BANISAR, D. [2006]: *Freedom of Information Around the World 2006: A Global Survey of Access to Government Information Laws*. Privacy International. <http://ssrn.com/abstract=1707336>
- BÁRTFAI ZS. – JÓRI A. – KERÉKES ZS. – MAJTÉNYI L. – MIKLÓSI Z. – POLYÁK G. – SZABÓ M. D. – SZOBOSZLAI J. [2005]: Az elektronikus információszabadság nálunk és más nemzeteknél. In: *Majtényi L. – Molnár P. – Petri L. Á. – Szabó M. D. (szerk.): Az elektronikus információszabadság*. Eötvös Károly Intézet. Budapest. 13–56. old.
- BUZÁS P. – RÉVÉSZ B. [2012]: *Közérdekű adatok (e-tananyag)*. Nemzeti Közszerződési Egyetem. Budapest. http://korrupciomegelozes.kormany.hu/download/b/be/b0000/Buz%C3%A1s%20P%C3%A9ter%20-%20R%C3%A9v%C3%A9sz%20Bal%C3%A1zs_K%C3%B6z%C3%A9rdek%C5%B1%20adatok%20%28e-tananyag%29.pdf
- EÖTVÖS KÁROLY KÖZPOLITIKAI INTÉZET [2015]: Hozzáférés a közzsféra adataihoz, 2005–2006. ekint.org/ekint_files/File/tanulmányok/hozzaferesakozszferaadataihoz.pdf
- KAISER T. (szerk.) [2015]: *Jó Állam Jelentés 2015*. Nemzeti Közszerződési Egyetem. Budapest. http://uni-nke.hu/uploads/media_items/jo-allam-jelentes.original.pdf
- KERÉKES ZS. [2004]: Információszabadság az Európai Unióban. *Fundamentum*. 8. évf. 4. sz. 115–119. old.
- NAIH (NEMZETI ADATVÉDELMI ÉS INFORMÁCIÓSZABADSÁG HATÓSÁG) [2012]: *Ajánlás a nemzeti vagyonnal gazdálkodó vagy azzal rendelkező, piaci szereplő gazdasági társaságok üzleti adatainak nyilvánosságáról*. Budapest. http://www.naih.hu/files/Uzleti-titok_kontra_nyilvanossag_AJANLAS.pdf
- RÉVÉSZ B. – SOMOGYVÁRI K. [2012]: Az információszabadság jelentése, tartalma. In: *Péterfalvi A. (szerk.): Adatvédelem és információszabadság a mindennapokban*. HVG-Orac. Budapest.
- RÉVÉSZ B. [2015]: Az adatvédelem általános alapjai In: *Christián L.: Információvédelem*. Nemzeti Közszerződési Egyetem Rendészettudományi Kar. Budapest. 13–29. old.
- RÉVÉSZ B. [2015]: Kettős küldetésben. *Ügyvédek Lapja Online*. 1. sz. <http://ugyvedek-lapja.hu/lapszam/2015-1/kettos-kuldetesben>
- SZILÁGYI K. – JÓRI A. – SZABÓ M. D. [2008]: *Az információszabadság elektronikus kézikönyve*. Educatio Társadalmi Szolgáltató Nonprofit Kft. Budapest. <http://www.tankonyvtar.hu/hu/tartalom/tkt/informacioszabadsag/adatok.html>

Summary

The performance of public duties and governmental tasks requires the production and processing of a great amount of data. On one hand, these information constitute national assets. On the other hand, the publicity of such data may contribute to the control and transparency of the state. A democratic state therefore must ensure access to information concerning its functioning

and its economy for the citizens. In addition, the emergence of the other information right – the right to the protection of personal data – shall also be taken into account since it is a primary constitutional obligation of the state as well. On many occasions, these two fundamental rights may conflict with each other, particularly in the case of personal data of public interest. Therefore, a fair balance must be struck between the transparency of the state and the protection of citizens' privacy.

The aim of this article is to present the current Hungarian regulation of the freedom of information in the historical and international context. Chapter 1 deals with the importance of this right as one of the most important tools within the fight against corruption. Chapter 2 focuses on the historical context of the freedom of information, and outlines the development of the most important national regulations. Chapter 3 presents the relevant provisions adopted by the European Union that guarantee transparency, as well as the practice of the European Court of Human Rights. Chapter 4 briefly outlines the institutional models for the freedom of information supervisory bodies. Chapter 5 analyzes the Hungarian regulation ensuring freedom of information. This part primarily focuses on the rules concerning access to public information upon request, and the possible limitations of the freedom of information. Chapter 6 introduces the legal background of the reuse of public data. Chapter 7 presents the practice of the National Authority for Data Protection and Freedom of Information in respect of frequent categories of cases. Finally, Chapter 8 outlines the most important amendments of the Freedom of Information Act recently introduced by the government.

The study concludes that the social importance of the freedom of information is indisputable. Its effect is becoming more significant. Although its boundaries shall be adjusted in our ever changing world, its constitutional foundations shall – however – remain intact to ensure the freedom achieved during the democratic transition of 1989.

Az adat-hozzáférési igények és az adatvédelmi kötelezettségek egyensúlyáról

Vereczkei Zoltán,
a KSH főosztályvezető-
helyettese
E-mail: Zoltan.Vereczkei@ksh.hu

A tanulmány arra tesz kísérletet, hogy egyrészt bemutassa az adat-hozzáférési igények és adatvédelmi kötelezettségek főbb elemeit, együttes kezelésük lehetőségeit, másrészt adat-hozzáférési/adatvédelmi szempontból ismertesse a KSH elmúlt egy évtizedes tapasztalatai alapján kialakított adat-hozzáférési csatornáit, melyek a két, bizonyos tekintetben egymás ellen ható igény egyensúlyának megteremtésére is irányulnak.

TÁRGYSZÓ:
Adathozzáférés.
Adatvédelem.
Adat-hozzáférési csatornák.

A hivatalos statisztikák fejlesztéséért, előállításáért és közzétételéért felelős intézmények működését leíró adat-előállítási folyamatokban kulcsszerepe van a tájékoztatásnak, általánosságban a statisztikai célra kezelt adatok hozzáférhetővé tételének, mely a statisztikai információk felhasználóinak igényeire épül. Ezek az intézmények az adat-előállítás során létrehozott statisztikai termékek mellett számos egyéb, a folyamatok során létre nem hozott statisztikai adat iránti igényt szolgálnak ki. Ebben a tekintetben az adat-hozzáférési igények a hivatalos statisztikai adat-előállítási folyamatok mozgatórugóját jelentik. Ezen igényeken belül különleges szerepe van a tudományos célú, kiemelten a statisztikai mikroadatokat iránti hozzáférésnek.

Ugyanilyen fontos elv a statisztikai célra kezelt adatokra épülő adat-előállítási folyamatban az adatszolgáltatók statisztikai célra szolgáltatott adatainak védelme, azaz az adatvédelmi kötelezettség. Ennek a kötelezettségnek az adatvédelem több területének együttműködésében jogi, módszertani, tájékoztatási és informatikai eszközökkel tudnak az érintett statisztikai intézmények eleget tenni.

A KSH¹ gyakorlatában az elmúlt bő tíz évben indított fejlesztések eredményeinek hatására az adathozzáférés gyakorlatában ma már az adathozzáférés és az adatvédelem igényeinek közös, azonos súlyú kezeléséről beszélünk. Mindez egy olyan környezetben folyik, ahol a tudományos célú hozzáférési igény a mikroadat-állományokhoz beépült a nemzeti statisztikai hivatalok működésének rendjébe, hazai és európai viszonylatban egyaránt.

1. Adatvédelmi tevékenység a KSH működésében

A hivatalos statisztika célja: jó minőségű statisztikai információk előállítása és hozzáférhetővé tétele minél szélesebb körben. A társadalmat, gazdaságot, környezetet jellemző statisztikai információk közzététele – a tájékoztatási tevékenység – a statisztikai adat-előállítási folyamat integrált része, melynek az Európai Statisztikai Rendszerben minden tagországban jellemzően több csatornája érhető el a felhasználók számára. Ezzel összhangban a statisztikai célokra kezelt adatok két kiemelt jellemzőjeként beszélhetünk a hozzáférhetővé tételről, illetve a hozzáférhetővé tett statisztikai információk jó minőségéről.

¹ KSH: Központi Statisztikai Hivatal.

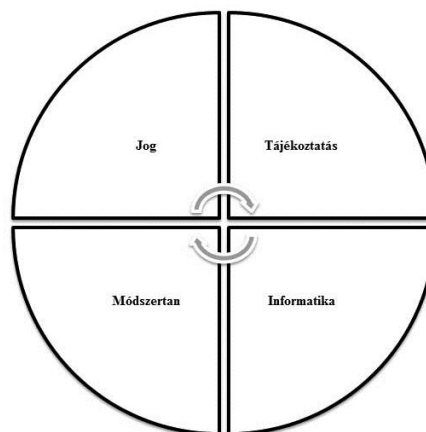
Mind a statisztikai adat-előállítási folyamat, mind a statisztikai adatokhoz való hozzáférés tekintetében fontos tényező az adatvédelem, ami a hivatalos statisztikában – általánosságban szólva – az adatszolgáltatók egyedi adatainak védelmét, azaz az azonosítás, illetve felfedés kockázatának kezelését jelenti. Az egyedi adatok védelmében fontos alapvetés, hogy – egyes jogszabály által meghatározott esetektől eltekintve (Nagy E. [2015]) – tilos azoknak a statisztikai egységeknek (egyedeknek) a beazonosítása, amelyekre a statisztikai célra kezelt adatok vonatkoznak. A hivatalos statisztikában ezt az eseményt, azaz a statisztikai egységek egyértelmű meghatározását hívjuk *azonosításnak*. A statisztikai célra kezelt adatok terén egy másik fontos jelenség az ún. *felfedés*, amely a statisztikai egységekre vonatkozó eddig nem ismert információ megismerését jelenti. Ezzel összhangban az adatvédelem tekintetében a hivatalos statisztika célja annak megakadályozása, hogy az egyes statisztikai egységekről olyan információ váljon megismerhetővé, amely egyébként ne lenne ismert. Ennek a célnak megfelelően az adatvédelem módszertani eszköztárát és megoldásait felfedés elleni védelemnek hívjuk.

Az egyedi adatok védelmére a hivatalos statisztikában több megoldás is alkalmazható, illetve több terület együttműködését indokolja. Ennek megértéséhez fontos a KSH adatvédelmi tevékenységének bemutatása.

1.1. A KSH adatvédelmi tevékenységének területei

A KSH-ban az adatvédelmi kötelezettség teljesítését alapvetően négy funkcionális terület (jogi, módszertani, informatikai és tájékoztatási) közös felelősségi körébe tartozó tevékenységeivel jellemezzük. (Lásd az 1. ábrát.)

1. ábra. A KSH adatvédelmi tevékenységének kerete



A *jogi adatvédelem* körébe tartozónak tekintünk minden olyan adatvédelmi kérdést, amelyet a hivatali működésre, a statisztikai folyamatokra, illetve statisztikai termékekre vonatkozóan külső és belső joganyagok útján szabályozunk. A hivatalos statisztikai célból kezelt adatok védelmére vonatkozóan európai és hazai jogszabályok határozzák meg az adatok védelmének kötelezettségét. Az adatvédelmi előírások az adatszolgáltatók, illetve a statisztikai adatok által jellemzett egységek adatait védő jogot hivatottak érvényre juttatni.

Ezek közül legrelevánsabb irányadó jogszabályok az európai statisztikákról szóló 223/2009/EK európai parlamenti és tanácsi rendelet és ehhez kapcsolódóan az Európai Statisztika Gyakorlati Kódexe; hazai szinten pedig a statisztikáról szóló 1993. évi XLVI. törvény, valamint annak végrehajtási rendelete (70/1993. (XII. 3.) Korm. rendelet a statisztikáról szóló 1993. évi XLVI. törvény végrehajtásáról), illetve az információs önrendelkezési jogról és az információszabadságról szóló 2011. évi CXII. törvény. A tanulmánynak nem célja, hogy a vonatkozó európai és hazai jogi adatvédelmi környezetet bemutassa, ezt a kapcsolódó tanulmányok részletesen megteszik (Nagy E. [2015], Lakatos [2015]).

A *módszertani adatvédelem* a jogi mellett a hivatali adatvédelmi tevékenység másik nagy meghatározó területe, mely jellemzően olyan logikai, matematikai, statisztikai megoldások fejlesztését és alkalmazását jelenti, amelyek azt hivatottak biztosítani, hogy egy adott adatállományon keresztül egyedi statisztikai információk, illetve a közzétett adatokból statisztikai egység(ek)re vonatkozóan új információ megismerésére ne kerüljön sor.

Az *informatikai adatvédelem* az informatikai rendszerek működésében szabályozható, technikai, illetve biztonsági megoldások fejlesztését, alkalmazását jelenti. Idetartoznak többek között az informatikai rendszerek által biztosított kommunikációs csatornák megfelelő védelmi eljárásai, a KSH belső hálózatán kezelt statisztikai célú adatok hozzáférhetőségének technikai, jogosultsági feltételei és általánosságban minden olyan megoldás, amellyel a működés irányadó informatikai biztonság elvárt szintje biztosítható.

A *tájékoztatási adatvédelem* a tájékoztatási célra előállítandó statisztikai termékek szabta feltételeket (például a táblák struktúráját, részletezettségét), illetve a tájékoztatási célra üzemeltetett rendszerek működési kereteit jelenti (például felhasználó által egy táblázatban lekérdezhető ismérvek maximális számának meghatározása ún. dimenziókorlátozás révén a KSH Tájékoztatási Adatbázisában). A tájékoztatási adatvédelem az adatvédelmi tevékenység egyik meghatározó eleme, hiszen a felhasználói igények kiszolgálására kialakított termékek (például a megjelentetendő statisztikai táblák tartalma, szerkezete) aktívan befolyásolják többek között a módszertani adatvédelmi megoldások szükségességét és alkalmazhatóságát.

Az adatvédelem négy kulcsterületének megfelelően a KSH-ban adatvédelemnek tekintünk minden olyan jogi, módszertani, tájékoztatási, informatikai, fizikai és

egyéb eljárást, módszert, tevékenységet, amelyet annak érdekében alkalmazunk, hogy a rendelkezésre álló adatokat a jogosulatlan, céltól eltérő felhasználástól megóvjuk, és így a felfedést megakadályozzuk. A céltól való eltérést az adathozzáférés keretében a lehető legáltalánosabban értelmezzük. A hivatal által gyűjtött adatok kezelése kivétel nélkül statisztikai célból történik. Az egyedi adatok gyűjtése a KSH tevékenységének eszköze és nem célja (KSH [2014a]). Emellett idetartozónak értjük azt is, hogy az egyes adat-hozzáférési csatornáknál alkalmazott jogi adatvédelmi eszközökkel (például szerződésekkel) szabályozott felhasználási céltól való eltérés az adott adathozzáférés keretében nem megengedett.

Önmagában az „adatvédelmi oldal” tekintve kimondhatjuk, hogy a KSH adatvédelmi tevékenységének célja az irányadó adatvédelmi szabályok teljes körű érvényre juttatása különféle jogi, módszertani, informatikai, tájékoztatási megoldások alkalmazásával. Az adatvédelmi tevékenység mindig az adatok felhasználásának korlátozását jelenti, ami jellemző módon valamilyen jogi vagy a hozzáférésre biztosított állományok lehetséges információtartalmának korlátozásában, végső soron az adatok minőségének romlásában (információvesztésben) jelentkezik. Az adatvédelmi oldal tehát általánosságban a hozzáférés jogi körülményeinek szabályozása, illetve a felfedés kockázatának kezelése kapcsán a közzeendő tartalom korlátozása felé mozdul el. Ez az elvárás adja az adat-hozzáférési/adatvédelmi mérleg egyik oldalát.

1.2. Minőség az adat-hozzáférési igények, adatvédelmi előírások kapcsolatában

A statisztikai célokra kezelt adatok további kulcsjellemzője az adatok jó minősége. A *minőség* általános közmegállapodás szerint felhasználásra való alkalmasságot jelent, amelynek több tényezőjét ismerjük és mérjük a statisztikai folyamatokban, valamint a statisztikai termékekkel kapcsolatban is. Ezeket a tényezőket nevezzük *minőségdimenzióknak* (KSH [2014e]). A relevancia, pontosság, időszerűség, időbeli pontosság, hozzáférhetőség, érthetőség, összehasonlíthatóság és koherencia minőségdimenziók olyan általános, a hivatalos statisztikai folyamatokra, illetve azok eredménytermékeire kiterjedő szempontokat jelentenek, amelyek a hivatalos statisztikai tevékenység minden folyamatában megjelennek. A minőségdimenziók egymástól nem függetlenek, azaz sokszor a statisztikai adatok egyes dimenziók szerinti javítása csak más dimenzió(k) szerinti megfelelés „rovására” történhet.

A minőség iránti elkötelezettség a nemzeti statisztikai szolgálatok fontos alapjellemezője, melynek egyik kiemelt része a minőségdimenziók alapján mért minőség. Ezeket a minőségdimenziókat a nemzeti statisztikai szolgálatok jellemzően külön politikákként deklarálják, és ennek megfelelően az érintett folyamatokkal, termékekkel kapcsolatos általános elvárásoknak tekintik. Így a KSH minőségpolitikája dekla-

rálja a KSH minőség iránti elköteleződését, és megnevezi a főbb minőségdimenziókat. A minőségi irányelvek pedig a statisztikai adat-előállítási folyamat minden rész-folyamatára vonatkozóan határozzák meg azokat a követendő minőségi elvárásokat, amelyeknek minden hivatalos statisztikának meg kell felelnie (*KSH* [2014c]).

A mai napig számos szakmai vita alapját képezik azok a nézetek, amelyek a minőség összetevőt tekintik a felhasználók számára a „legrelevánsabbnak”. Többször hallani például arról, hogy a modern, technikailag fejlett világban az időszerűség szempontja a döntő (a piac a gyors információt értékeli igazán, amelynek „áraként” a pontosság terén „elszenvedett” veszteséget hajlandó tolerálni), holott a hivatalos statisztikák egyik legfontosabb jellemzőjeként továbbra is a pontosságot emlegetik, ami viszont jellemzően az időszerűség ellen ható minőségdimenzió. Erre a dilemmára jó példa a negyedéves vagy havi bruttó hazai termék adatok becslése, melyek az éves, sokkal részletesebb nemzeti számlák elkészültével mindig változnak, de nem vitás, hogy mindkét adatállomány fontos felhasználói igényeket szolgál ki.

A jó minőséget – és általában a minőséget – a hivatalos statisztikát fejlesztők, előállítók, közlétezők és felhasználók jellemzően tehát másként érzékelhetik; a minőség különféle dimenziót tarthatják hangsúlyosabbnak (mondjuk úgy, fontosabbnak), míg más felhasználói csoportoknál jellemzően más dimenziók lehetnek kiemelten fontosak. A minőség dimenzióinak eltérő rangsorolása az adathozzáférésekkel kapcsolatban olyan megoldások kidolgozását igényli, amelyek a felhasználók eltérő igényeinek megfelelően kellő rugalmasságot képesek biztosítani.

A felhasználók számára a tájékoztatási tevékenység az, aminek keretében kapcsolatba kerülnek a hivatalos statisztika eredménytermékeivel; itt történik a statisztikai eredmények eljuttatása a felhasználókhöz, a megfelelő tájékoztatási csatornák és közlési formák alkalmazásával. A statisztikai adat-előállítási folyamatban a tájékoztatási tevékenység ennek megfelelően kulcsfontosságú szakasz. Az ESTFM², mely a GSBPM³ magyar adaptálásának tekinthető, szintén ezen elgondolás mellett szerepelteti a tájékoztatási tevékenységet.

Az ESTFM-modell megvizsgálva látható (lásd a 2. ábrát), hogy a felfedés elleni védelemnek meghatározott „helye” van a folyamatmodellben (külön tevékenységként kerül feltüntetésre) az Adatelemzés részfolyamat alatt. Ez nem jelenti azt, hogy az adatvédelmi tevékenység csupán a tájékoztatási tevékenységet megelőzendően van jelen az adat-előállítási folyamatban. A statisztikai célú adatkezelés megköveteli, hogy a teljes folyamatban érvényesüljenek az adatvédelmi szabályok. A vonatkozó legfőbb előírások a KSH nyilvánosan is hozzáférhető adatvédelmi szabályzatában olvashatók (*KSH* [2013]).

² ESTFM: egységes statisztikai folyamatmodell.

³ GSBPM (generic statistical business process model): egységes statisztikai adat-előállítási folyamatmodell. Lásd: <http://www1.unece.org/stat/platform/display/metis/The+Generic+Statistical+Business+Process+Model>

2. ábra. Az egységes statisztikai folyamatmodell

Értékelés, visszacsatolás/Minőségmenedzsmnt/Metaadatok menedzsmenje									
I. Igényspecifikáció	II. Tervezés	III. Szervezés, fejlesztés	IV. Adatok gyűjtése	V. Adatelő-készítés	VI. Feldolgozás	VII. Adat-elemzés	VIII. Tájékoztató-adatmeg-jelenítés	IX. Archiválás	
I.1. Információ-igények meghatározása	II.1. Adatforrások feltérképezése, elemzése II.2. Adat-előállítási folyamat oupujának specifikációja	III.1. Adat-gyűjtési eszközök tesztelése, fejlesztése	IV.1. Felvételi keret, adatszolgáltatási kör, minta kiválasztása	V.1. Adatbevitel, adagyűjtés lezárása	VI.1. Adat-integrálás	VII.1. Adat-előállítási folyamat tervezésének elkészítése	VIII.1. Adatok átöltése adattárházba	IX.1. Archiválási szabályok definiálása	
I.2. Igények egyeztetése, jóváhagyása	II.3. Fogalmak, változók definiálása	III.2. Informantikai eszközök teszteléséből	IV.2. Adagyűjtés szervezése, oktatása	V.2. Kódolás	VI.2. Mutató-képzés	VII.2. Szemzonalis kiigazítás	VIII.2. Tájékoztatósi termékek előállítás	IX.2. Archiválási fájlok kezelése (menedzsmenje)	
I.3. Adat-előállítási folyamat oupujának meghatározása	II.4. Adagyűjtési módszer tervezése	III.3. Informantikai támogató eszközök tesztelése	IV.3. Adagyűjtés	V.3. Ebnálás, makrovalidálás, adatok pótlása	VI.3. Súlyképzés, teljeskörűsítés	VII.3. Eredmények validálása	VIII.3. Közvetített menedzsmenje	IX.3. Archiválás	
I.4. Fogalmak meghatározása	II.5. Felvételi keret meghatározása	III.4. Proba-felvetel	IV.4. Érkeztetés, nyugomon követés		VI.4. Aggregálás	VII.4. Felfedés elleni védelem	VIII.4. Értékesítés, marketing	IX.4. Archivált fájlok karbantartása	
I.5. Adatfelvételi javaslatnrv készítése	II.6. Mintavétel megtervezése II.7. Adagyűjtési mérőeszköz, kérdőív tervezése	III.5. Adagyűjtés szervezése, oktatása, előkészítése			VI.5. Adat-állományok véglegesítése	VII.5. Adat-előállítási folyamat oupujának véglegesítése	VIII.5. Felhasználatok támogatása, információ-szolgálat		
	II.8. Adatelő-készítés és -feldolgozási folyamat tervezése								
	II.9. Munkaterv összeállítása								
	II.10. Részletes adatfelvételi terv összeállítása								

Forrás: KSH [2014c] 11. old.

A statisztikai adat-előállítási folyamatnak biztosítania kell, hogy az egyes konkrét folyamatok során a megfelelő minőségi elvárások érvényesüljenek. Ezt a célt szolgálják a KSH minőségi irányelvei is, melyek közül több az adatvédelmi szabályok érvényesítését, illetve az adat-hozzáférési igényekkel való összhang megteremtését írja elő. Példaként említhetők a következők:

- „az adatkérési igényeket mindig az adatvédelmi és adat-hozzáférési szempontokat tekintve legalkalmasabb adat-hozzáférési csatormán kell teljesíteni. Az adatkérők nyilvános, bárki számára elérhető forrásból tájékozódhatnak az adat-hozzáférési csatornákról és azok működési sajátosságairól”, vagy
- „táblázatos adatok nyilvánosságra hozatala, illetve anonimizált mikroadat-kiadás esetén törekedni kell arra, hogy minél kisebb információvesztés mellett biztosítsuk az adatok kellő védelmét, azaz a felfedési kockázat minimális szinten tartását”. (KSH [2014c]).

2. Az adat-hozzáférési igényekről

Az elmúlt évtizedben jelentősen felerősödött a statisztikai adatok megismerése iránti igény (mind mennyiségben, mind igényelt tartalomban), és kiemelkedően megnőtt a tudományos célú hozzáférési igények, legfőképp a mikroadat iránti adatkérések száma. A hivatalos statisztika fejlődése során hagyományosan az indikátorok, adatok nyomtatott, táblázatos megjelenítésére volt igény, de a kommunikációs eszközök fejlődésével egyre inkább előtérbe kerültek az elektronikus megoldások, az adatbázisok elérhetősége stb. Ezzel párhuzamosan a felhasználók az informatikai, módszertani eszközök fejlődésével egyre inkább képessé váltak a feldolgozás, elemzés egyes részeit saját maguk elvégezni. Ezek miatt a hivatalos statisztika világában a mikroadatszintű hozzáférés, azaz a megfigyelési egységekre vonatkozó adatokat tartalmazó rekordszintű adatok tudományos célú felhasználása kiemelt célkitűzéssé vált. Ez a megváltozott tendencia az elmúlt tíz év gyakorlatát tekintve hazai és európai szinten egyaránt megfigyelhető.

A mikroadat-hozzáféréssel járó nagyobb adatvédelmi kockázat a táblázatos adatok kezelésénél alkalmazott gyakorlatokhoz képest más, illetve új adatvédelmi megoldások kidolgozását tette szükségessé. A mikroadat-hozzáférés a KSH gyakorlatban tudományos célhoz kötött. A hivatalos statisztikában (a legtöbb európai tagország gyakorlatával összhangban) annak megállapítását, hogy adott adatkérés tudományos célúnak minősíthető-e vagy sem, az ún. kutatói akkreditációs vizsgálat hivatott bizto-

sítani, ami 2014 óta a KSH adat-hozzáférési gyakorlatának is integrált része.⁴ A tudományos célú hozzáférési igényeket vizsgálva a következő alap megállapítások tehetők:

- a tudományos célú hozzáférési igények száma évről-évre folyamatosan növekszik;
- egyre több statisztikai témához tartozó mikroadat-állományhoz kérnek tudományos célú hozzáférést;
- jelentős arányt képviselnek az egyedi kérésre összeállított mikroadat-állományok, azaz olyan összekapcsolt vagy összekapcsolásra előkészített mikroadatok iránti hozzáférési igények, amelyek nem szerepelnek a KSH adat-előállítási folyamata során előálló, kutatásra előkészített mikroadatok listáján.⁵

A tudományos kutatási célok teljesülése keretében a kutatók abban érdekeltek, hogy a hozzáférésre kért állományok jó minőségűek, a lehető legrészletesebbek legyenek, hogy az abból levont kutatási eredményeket ne torzítsák a felfedés elleni védelmi intézkedések. Az adat-hozzáférési oldal tehát általánosságban a közléteendő tartalom maximalizálása felé mozdul el.

Az adatkérési igények eddigi ismerete kapcsán elmondható, hogy az adathozzáférés keretében minden olyan statisztikai célból kezelt információ iránt van, illetve lehet hozzáférési igény, amely a KSH kezelésében van. Ez az elvárás adja az adatvédelmi/adat-hozzáférési mérleg másik oldalát.

3. Áttekintés – adat-hozzáférési/adatvédelmi igények és tendenciák változása a KSH elmúlt tíz évében

Az adat-hozzáférési, adatvédelmi igények terén bekövetezett változásokat vizsgálva kijelenthető, hogy bizonyos tekintetben alapvetően más tendenciák érvényesülnek a mostani gyakorlatban, mint egy évtizeddel ezelőtt, ugyanakkor egyes területeken a már korábban is meglévő elemek és nézetek nem változtak, azok ugyanúgy érvényesek ma is, mint tíz évvel ezelőtt vagy akár már azt megelőzően.

A 2000-es évek elején is ugyanúgy ismertek és a gyakorlatban betartottak voltak az adatvédelem alapelvei, mint a hivatal fennállása óta bármikor (*Lakatos* [2015]). Az alapelvek ezekben az években a táblázatos adatok kiadására vonatkozó adatvé-

⁴ A kutatói akkreditáció részleteit lásd: http://www.ksh.hu/kutato_i_akkreditacio

⁵ A kutatásra előkészített mikroadat-állományok listáját lásd: http://www.ksh.hu/kutatoszobai_hozzaferes

delmi szabályokat tartalmazták; a mikroadatszintű hozzáférés és a kapcsolódó adatvédelem nem volt még szabályozva.

A KSH-t kutatói oldalról sokszor érte az a kritika, hogy túlzottan erős adatvédelmet alkalmaz, és gyakorlata nem szolgálja kellő mértékben a tudományos kutatás igényeit. Ennek a kritikának egyik oka részint az volt, hogy a hivatal adatvédelme nem volt ismert a felhasználók számára. Másik fő okként az említhető, hogy a már abban az időben is jelentősen előretörő tudományos célú hozzáférés – kiemelten a mikroadatok – szerepének erősödésével a KSH gyakorlata nem idomult olyan mértékben a megváltozott adat-hozzáférési igényekhez, mint azt a tudományos élet résztvevői elvárták volna.

A felfedés elleni védelem szabályozott módszertani központosításáról a KSH gyakorlatában 2003 óta beszélhetünk érdemben, bár a hivatal táblázatos adatok kiadására vonatkozó rendje 2003 előtt is szabályozott volt. A tájékoztatási adatbázis, illetve a mikroadatszintű hozzáférés térnyerésével egyidőben váltak igazán láthatóvá a módszertani harmonizációt igénylő adatvédelmi problémák; ekkor fogalmazódott meg az összehangolás iránti igény. Ebben az időben kezdett foglalkozni a KSH központi módszertani egysége az adatvédelem módszertani megoldásainak egységesítésével és az egységes gyakorlatok kialakításával. A felfedés elleni védelem kitüntetettebb szerepével egy időben lényegében a hivatal is felismerte, hogy a „KSH az egyre bővülő hazai és nemzetközi felhasználói igényeket úgy tudja kielégíteni, ha szélesíti az adat-hozzáférési, -kiadási csatornák körét, köztük kiemelten a kutatói hozzáférés lehetőségeit (tájékoztatási adatbázis, kutatószoba stb.)” (*Szép* [2012]). Ennek keretében számos olyan munka indult, amely az adatvédelem módszertanának és hivatali gyakorlatának egységesítésére és továbbfejlesztésére irányult, és amelynek eredményei napjainkban váltak jobban láthatóvá a felhasználók számára is.

Ezen fejlesztések eredményeként jelentkező egyik leglátványosabb változás a mostani és az évtizeddel korábbi gyakorlathoz képest, hogy mára a mikroadat-állományok kezelése és hozzáférésük biztosítása mindennaposnak számít. Tíz évvel ezelőtt nem volt általánosan elismert és szabályozott a tudományos célú kutatás azon méltányolható alapérdeke, hogy a mikroadatszintű statisztikai adatok hozzáférhetőek legyenek. Bár a hivatal is érzekelte ennek igényét a tudományos élet hazai képviselői és az Európai Statisztikai Rendszer részéről.

Az elmúlt évek során több európai és hazai szintű fórumon szóltak a tudományos élet és a hivatalos statisztikai intézmények képviselői az adathozzáféréssel, adatvédelemmel kapcsolatos elvárásaikról, az általuk érzékelt problémákról (*Harcza* [2012]; *EC* [2013], [2014]). Ezek a párbeszédok hozzájárultak ahhoz, hogy a KSH mai gyakorlatában az adatvédelmi tevékenységet párban kezeljük az adat-hozzáférési igényekkel. Bár a hivatalnak fennállása óta fő célja, hogy a társadalom, gazdaság és környezet állapotáról és annak változásairól statisztikákat hozzon nyilvánosságra, az adatvédelmi tevékenység adat-hozzáférési igényekkel történő közös kezelése csak az

utóbbi tíz évben vált gyakorlatának integrált részévé. Ez leginkább az adat-hozzáférési lehetőségek bővítésével, illetve a felfedés elleni védelmi módszerek szabályainak egységesebb alkalmazásával valósul meg.

A KSH kiemelkedő eredményeként tartjuk számon az új és megújított adat-hozzáférési csatornákat (azon belül is a kutatószoba kialakítását és újraszervezését), illetve a tudományos élet képviselőivel kialakított párbeszédet is szolgáló, a hivatal adatvédelmi gyakorlatáról szóló információk nyilvánosságra hozatalát, beleértve több tanulmányt és tájékoztatót, amely a felfedés elleni védelmi megoldások iránt érdeklődők számára szabadon hozzáférhető (lásd például a KSH adatigénylési lehetőségekre vonatkozó részletes honlaptartalmat,⁶ *KSH* [2014b], *Erdei-Horváth* [2004], továbbá jelen szám több cikke: *Kristóf* [2015], *Nagy B.* [2015]). Új módszertani kutatási eredmények is születtek (például *Faragó* [2013]). Ezeknek a fejlesztési irányoknak a fontosságát a tudományos élet képviselőivel folytatott párbeszéd és a témában rendezett fórumok is deklarálták (*Harcza* [2012]). A kutatók felismerték, hogy érdekeltek az adatvédelmi módszerek megválasztásában, hiszen a különböző eljárások eltérően befolyásolják az adatminőséget, a védett adatokon végzett elemzések relevanciáját (*Bartus* [2013]).

Az adathozzáférés, adatvédelem közös kezelésének igénye nem hazai, hanem európai sajátosság. Az elmúlt évek nagyobb európai fejlesztései is azt a képet mutatják, hogy az adat-hozzáférési, adatvédelmi kérdések különálló kezelése semmiképp nem lehet hatékony a rendszer egészének működését tekintve. A kialakítandó, továbbfejlesztendő gyakorlatoknak minden főbb érintett (kiemelten a hivatalos statisztikát fejlesztő, előállító és közvétező intézmények, adatarhívumok és kutatók) közös párbeszédén, illetve nemzetközi gyakorlattal összhangban levő megoldásokon kell alapulnia (*EC* [2013], [2014]), biztosítva az adatvédelmi alapelvek maradéktalan érvényesülését.

4. Egyensúly az adathozzáférés és az adatvédelem között

Az eddig elmondottak ismerete tükrében felteheti a kérdést az Olvasó, hogy hogyan biztosítható az adathozzáférés és az adatvédelem bizonyos tekintetben egymás ellen ható érdekei közötti egyensúly. Az adatvédelmi elvárások végiggondolása után megérthetjük, hogy a hivatalos statisztikai adatok kezelésének egyik alappillére az adatszolgáltatók bizalmának megnyerése és fenntartása azáltal, hogy a hivatalos statisztikai adatokat kezelő intézmény mindent megtesz a statisztikai egységekre

⁶ Lásd: <http://www.ksh.hu/adatigenyles>

vonatkozó egyedi információk védelme érdekében. Ezen a gondolati síkon arra a következtetésre jutunk, hogy pusztán az adatvédelmi előírásoknak való megfelelés a hozzáférhetővé tehető adatok korlátozását eredményezi.

Az adatkérők érdekeit, minőségi elvárásait szem előtt tartva azt láthatjuk, elvárás, hogy a lehető legtöbb, hivatalos statisztikai célból kezelt információt hozzáférhetővé tegyünk, azaz a hozzáférhető adatok köre a lehető legszélesebb legyen.

A cél a két egymás ellen ható szempont valamiféle egyensúlyának megteremtése. Nincs ez másképp az adathozzáférés-adatvédelem igényei esetében sem. A hivatalos statisztikák célja ebben a tekintetben egyfajta nyugalmi egyensúlyi állapot elérése. Kimondhatjuk tehát, hogy az adat-hozzáférési-adatvédelmi tevékenység egyfajta optimumkeresési feladatot jelent. Amennyiben a feladat megoldója mind az adathozzáférési, mind az adatvédelmi szempontokat ugyanolyan fontosnak tartja, úgy egyúttal azt is kimondhatjuk, hogy a feladatnak nem lehet jó megoldása az, ha csak az egyik vagy csak a másik célt tűzzük ki magunk elé. A hivatalos statisztikai tevékenység tehát nem lehet igazán hatékony, ha nem társul hozzá az adat-hozzáférési igények alapos ismerete és az azoknak megfelelő gyakorlatok megvalósítása. A hivatalos statisztika „öngyilkosságának” fogható fel az, ha pusztán a hozzáférési elvárások teljesítését tűzi maga elé célul, és nem törődik a jogos adatvédelmi elvárásokkal.

Hogyan biztosítható tehát a hivatalos statisztikai adatokat kezelők gyakorlatában ennek az egyensúlynak a megteremtése? Ezekre adnak választ az ún. adathozzáférési csatornák, illetve a statisztikai adat-előállítási folyamatban alkalmazott módszertani megfontolások és a már hivatkozott minőségi irányelvek.

4.1. Az adat-hozzáférési igények és az adatvédelmi előírások egyensúlyának garanciái a KSH stratégiai értékű dokumentumaiban

Az adathozzáférés, illetve adatvédelem területén a két elvárás közötti összhang megteremtésének fontosságára talán a legjobb példa a KSH 2020-ig tartó stratégiája. A stratégiai dokumentumban a hét fő stratégiai irány közül az egyik kiemelten ezzel a teljesítendő céllal foglalkozik, mely „az adatvédelem és az adathozzáférés egyensúlyának biztosításáról” szól. A stratégiai főirány meghatározza, hogy „a hivatalos statisztika alapvető célja az adatokhoz való lehető legszélesebb körű hozzáférés biztosítása. Emellett azonban törvényi kötelezettségünk (és gyakorlati érdekünk) az adatszolgáltatók egyedi adatainak védelme. Folyamatosan törekszünk olyan új technológiák, módszerek fejlesztésére és alkalmazására, melyek az adatvédelem szintjének sérülése nélkül segítik elő a statisztikai adatokhoz való biztonságos hozzáférést”. (KSH [2014d]). A stratégiai főirány korszerű adatvédelmi megoldások alkalmazásáról, az adathozzáférés integrált, szolgáltatás-orientált rendszerének kialakításáról, modern biztonsági megoldások alkalmazásáról, statisztikai célra kezelt állományok-

hoz való széles körű hozzáférés biztosításáról, illetve közös adat-hozzáférési-adatvédelmi megoldások kidolgozásáról szól.

A stratégia mellett az egyensúly biztosításának fontos elemei a korábban már említett minőségi irányelvek, illetve a KSH szintén nyilvánosan hozzáférhető adatvédelmi politikája is. Az adatvédelmi politika határozza meg a hivatal kilenc adatvédelmi alapelvét, köztük kiemelten a statisztikai célra kezelt adatokhoz való tudományos célú hozzáférés fontosságát (KSH [2014a]).

4.2. A KSH adat-hozzáférési csatornái

Az utóbbi évtizedek informatikai fejlesztései, valamint a felhasználók informatikai eszközellátottsága és tudása jelentősen hozzájárult az adat-hozzáférési igények növekedéséhez. A különböző felhasználók eltérő igényei miatt nem lehet egységesen kezelni az adathozzáférés rendszerét, ezért is szükséges több adat-hozzáférési csatornát üzemeltetni. A KSH mindezeknek megfelelően alakította ki a csatornáit,⁷ amelyek a felhasználók számára az adathozzáférés területén az egyik legfontosabb garanciát jelentik.

A hivatalos statisztikai információkhoz hozzáférő felhasználók körében jellemzően megkülönböztetett szerep a kutatóké, a tudományos célú elérést kérőké. Ahogy az Európai Statisztikai Rendszer egésze, úgy a KSH is kiemelt figyelmet fordít az általa kezelt hivatalos statisztikai adatok tudományos célú hasznosítására és az adat-hozzáférési és adatvédelem közötti egyensúly megteremtésére (KSH [2014d]). Ez természetesen nem jelenti azt, hogy a hivatalos statisztikai adatokat nem tudományos célból felhasználók nem férhetnek hozzá ezekhez az adatokhoz, illetve, hogy az ő igényeik kevésbé lennének fontosak a hivatalos statisztika számára. A kétféle elvárás az alkalmazandó garanciák miatt mégis szükséges elkülöníteni; a tudományos és nem tudományos célú hozzáférési lehetőségek az elérhető adat-hozzáférési csatornák révén ismerhetők meg a legérthetőbb módon. A KSH által üzemeltetett hat adat-hozzáférési csatorna adatvédelmi szempontú csoportosítását mutatja a 3. ábra.

Ezeket az adat-hozzáférési csatornákat a környezet típusától függően két fő csoportba soroljuk: az ún. biztonságos környezetben és a tetszőleges környezetben történő hozzáférésekre. Az adat-hozzáférési csatornák ilyen megkülönböztetése minőségi szempontból is különbséget jelent a felhasználók számára, mint ahogy azt a következőkben látni fogjuk.

⁷ Lásd: <http://www.ksh.hu/adatigenyles>

3. ábra. A KSH adat-hozzáférési csatornáinak adatvédelmi szempontú csoportosítása

<p>Táblázatos adatok kiadása</p> <p>Nyilvános adatok, adatbázisok a KSH honlapján, évkönyvekben, gyorstájékoztatókban</p> <p>Egyedi kérésre összeállított táblázatos adatok</p> <p>Felhasználási feltételek</p>	<p>Anonimizált mikroadat-kiadás</p> <p>Kutatásra előkészített „kész” állományok</p> <p>Egyedi kérésre összeállított anonimizált mikroadat-állományok</p> <p>Szerződés és titoktartási nyilatkozat aláírása szükséges</p>	<p>Nyilvános mikroadatfájlok elérése</p> <p>2014 márciusa óta elérhető (tesztfájl és oktatási fájl, a 2011. évi népszámlálásra alapozva)</p> <p>Felhasználási feltételek</p>
<p>Kutatósobai hozzáférés</p> <p>Kutatásra előkészített „kész” állományok kutatók számára (ingyenes)</p> <p>Egyedi kérésre összeállított adatállományok (térítési díj)</p> <p>Szerződés és titoktartási nyilatkozat aláírása szükséges</p>	<p>Távoli hozzáférés</p> <p>Kutatósobai környezet elérése nem csak Budapestről (2014 márciusa óta Szegedről is)</p> <p>Szerződés és titoktartási nyilatkozat aláírása szükséges</p>	<p>Távoli végrehajtás</p> <p>Kutató az adatállományt „nem látja”</p> <p>KSH állítja elő a kért kutatási eredményeket</p> <p>Szerződés és titoktartási nyilatkozat aláírása szükséges</p>

4.2.1. Statisztikai adatokhoz való hozzáférés tetszőleges környezetben

A KSH a legtöbb adatkérést táblázatos formában szolgálja ki. Ez a csatorna nem követeli meg a biztonságos környezet kialakítását. Mind a táblázatos, mind a nyilvános mikroadatfájlok elérésének együttes jellemzője, hogy az adatkérők ténylegesen rendelkeznek az adatok felett (esetenként a továbbadás feltételei jogilag korlátozottak), azokat saját munkakörnyezetükbe elérhető formában áttemelhetik (letölthetik). Ennek megfelelően az ezen állományok hozzáférhetővé tételével járó felfedési kockázat is mindig magasabb, mint a biztonságos környezetű hozzáférhetővé tétel esetén, hiszen a felfedéshez több eszköz és lehetőség áll a felhasználó rendelkezésére, ezért (az egyenkockázat elve érdekében) szigorúbb adatvédelemre van szükség.

A tetszőleges környezetben működő adat-hozzáférési csatornák közül hármat különböztetünk meg: a táblázatos adat, valamint az anonimizált mikroadat-kiadást, illetve a nyilvános mikroadatfájlok elérését. Ezen csatornák közül az anonimizált mikroadat-kiadás az, amely csak tudományos kutatási célból hozzáférhető, a többi csatorna pedig bármely adatkérő rendelkezésére áll.

Táblázatos adatok kiadása. Idetartozik minden olyan aggregált statisztikai információhoz való hozzáférés, amely például a KSH honlapján különféle formákban (adatbázisok, gyorstájékoztatók, háttértáblák stb.) közölt adatok útján vagy bármely más formában (például papíralapú kiadványok) jut el a felhasználóhoz. Ezen tájékoztatói ág mellett ebbe az adat-hozzáférési csatornába tartozik az is, amikor adatkérők igényelnek hozzáférést olyan táblázatos adatokhoz, amelyeket a hivatal az adat-előállítási folyamatokban rendszeresen nem állít elő. Ezeket hívjuk táblázatos ada-

tokra vonatkozó adatkérési igényeknek, amelyeket bármely felhasználó benyújthat; az adatkérések tudományos célját ebben az esetben a KSH nem vizsgálja.

A hivatal a táblázatos adatok kiadásánál is minden esetet köteles teljes körűen megvizsgálni, és szükség esetén felfedés elleni védelmi megoldásokat alkalmazni a felfedési kockázat minimalizálására. A cél kettős: biztosítani kell az érintett állományok védelmét úgy, hogy az az adatszolgáltatói adatok felfedését megakadályozza, ugyanakkor az adatvédelmi tevékenység ellátásnak eredményeként létrejött ún. védett állományok információtartalma a lehető leggazdagabb legyen (a minőségromlás a lehető legkisebb legyen).

A KSH anonimizálási gyakorlatának megértéséhez fontos szempont az adat-hozzáférési-adatvédelmi egyensúlyra törekvés mellett a módszertani összehangoltság is. Cél, hogy a hivatal egységes módszertani ajánlásokon alapuló gyakorlatot alkalmazzon az érintett állományok felfedés elleni védelme során.

A KSH a táblázatos adatok felfedés elleni védelmére a cellaelnyomás módszerét alkalmazza, mely lényegében azt jelenti, hogy a védendő cellákat töröljük az állományból és egyezményes jellel helyettesítjük. Az irányadó adatvédelmi szabályozás, illetve módszertani megfontolások alapján fontos, hogy minden egyes nyilvánosságra kerülő táblázatos adatállománynál a felfedési szempontú vizsgálat megtörténjen, és szükség esetén az elsődleges cellaelnyomási feladatot a hivatal elvégezze. Utóbbi esetben kötelező minden esetben vizsgálni a másodlagos cellaelnyomás⁸ szükségességét; beigazolódása esetén azt végre kell hajtani. Az adat-hozzáférési igényekre reagálva a KSH a következő szabályokat alkalmazza a cellaelnyomási feladatok végrehajtásakor:

- minden érzékenynek ítélt cella védelme biztosított legyen a védett táblázatban;
- a cellaelnyomás során a lehető legkisebb legyen az információvesztés, amit alapvetően a következők betartásával juttatunk érvényre:
 - cél a lehető legkevesebb számú cella elnyomása;
 - amennyiben több lehetőség is adódik a minimális számú cellaelnyomására, akkor azt a megoldást választjuk, amely során összességében a legkevesebb hozzájáruló adata kerül elnyomásra;
 - az összesen adatok elnyomását kerüljük, mert azok jellemzően nagy információtartalmúak.

További részletek a táblázatos adatok védelméről a nyilvánosan hozzáférhető KSH kutatói tájékoztatójában olvashatók (KSH [2014b]).

⁸ Táblázatos adatok védelmére alkalmazott eljárás: az elsődleges cellaelnyomás során az elnyomott cellákon felül további cellákat nyomunk el a táblázatos adat védettségének biztosítása érdekében (KSH [2013]).

Anonimizált mikroadat-kiadás. A kutatók tudományos kutatási célból férhetnek hozzá az anonimizált mikroadat-állományokhoz. Ezeket az állományokat rendszerint valamilyen adathordozón kapják kézhez vagy elektronikus csatornán érik el. Az anonimizált mikroadat-állományok használatára is irányadó jogi adatvédelmi előírások vannak érvényben, melyek közül több a felfedési kockázat minimalizálását szabályozza (például az igénybejelentő adatlapon az előre nem jelzett állományokkal történő összekapcsolás tilalma, harmadik személy részére a hozzáférhetővé tétel tilalma, igénybejelentő adatlapon jelzett tudományos cél elérése esetén az állomány megsemmisítésének kitétele stb.), ezen állományokat a kiadást megelőzően szintén teljes körűen felfedés elleni védelemmel kell ellátni, hogy a felfedési kockázat a lehetőségekhez mérten minimális legyen.

A felfedési kockázati megoldások alkalmazása mindig az érintett adatállományok valamilyen szintű minőségi romlásával jár együtt, amely általában az állomány részletezettségében bekövetkezett változtatásokon érhető tetten, és eredményképp az állományból levont következtetésekből (például kutatási célra készített becslések) beálló torzításban mutatkozhat meg. Emiatt kiemelten fontos a felfedés elleni védelmi megoldások helyes megválasztása. Ebben alapvetően két fő szempont az irányadó:

- az állomány felfedési kockázata minimális legyen;
- az anonimizált mikroadat-kiadás során létrehozott anonimizált állománnyal a kitűzött kutatási cél a lehető legjobb minőségben valósuljon meg.

Ennél az adat-hozzáférési csatornánál tehát élesen megmutatkozik az adat-hozzáférési, illetve adatvédelmi igények összehangolásának igénye.

Nyilvános mikroadatfájlok elérése. A KSH gyakorlatában relatíve újnak számító adat-hozzáférési csatorna a nyilvános mikroadatfájlok elérése. A nyilvános mikroadatfájloknak kiterjedt és több megközelítést alkalmazó szakirodalma van, amelynek bemutatására jelen tanulmány nem vállalkozik. A nyilvános mikroadatfájlok céljáról, formáiról, KSH-ból elérhető állományairól és azok előállításának módszertanáról jelen szám más cikkében tájékozódhat az Olvasó (*Kristóf* [2015]).

Az adathozzáférés-adatvédelem egyensúlya kapcsán fontos kiemelni, hogy a nyilvános mikroadatfájlok nem azt a célt szolgálják, hogy azok alapján jó minőségű következtetéseket, becsléseket lehessen levonni, előállítani az érintett statisztikai sokaságról. Cél elsősorban a biztonságos környezetben folyó kutatómunka előkészítésével, illetve az oktatási céllal kapcsolatos felhasználói igények kiszolgálása.

4.2.2. Statisztikai adatok hozzáférése biztonságos környezetben

A hozzáférés biztonságos környezetben alapvetően olyan körülmények létrehozását jelenti, amelyben az adatokhoz hozzáférő személy a hozzáférést valamilyen kont-

rollált környezetben teszi meg. Mivel a biztonságos környezet önmagában több (főként informatikai és jogi adatvédelmi) garanciát ad a hozzáférés kontrollálására, ennél fogva alapvetően preferált megoldást jelent az általában magasabb felfedési kockázatot jelentő mikroadatokhoz való hozzáférésre.

A biztonságos környezet adat-hozzáférési csatornái: kutatósobai hozzáférés, távoli hozzáférés, távoli végrehajtás. Utóbbi kettő alapvetően azonos hozzáférési feltételeket jelent a felhasználók számára, míg a távoli végrehajtás során más típusú kontrollált környezettel találkozhat a kutató.

Kutatósobai és távoli hozzáférés. A kutatósobai hozzáférés egyik fő jellemzője, hogy olyan informatikailag biztonságos környezetben történik a hozzáférés a mikroadat-állományokhoz, amelyek garantálják, hogy olyan információ nem hagyja el a biztonságos környezetet, amely felfedési kockázatot foglalna magában. Az irányadó szabályok értelmében a kutatók a biztonságos környezetben kutatásra előkészített állományokat nem vihetik ki a környezetből. A kutatósobai, távoli hozzáférés célja, hogy a kutatásra kért mikroadatállományokból a biztonságos környezeten belül állítson elő kutatási eredményeket a kutató, amelyeket felfedés elleni védelmi ellenőrzést követően kaphat kézhez (KSH [2014b]).

A biztonságos környezet ezen formája lehetővé teszi, hogy olyan állományokhoz, illetve állományok olyan részletezettségű változataihoz férjen hozzá a kutató, amelyre egyébként biztonságos környezeten kívül nem lenne lehetősége. A kutatói hozzáférés kockázatának kezelését a kutatói akkreditáció, illetve az adatvédelem jogi eszközei hivatottak garantálni. Ennek megfelelően a kutatósobai, illetve távoli hozzáférés formájában van lehetőség a legrészletesebb mikroadat-állományok elérésére, ebből következően az ilyen környezetben előállított kutatási eredmények a legjobb minőségűek lehetnek. Ennek oka, hogy a biztonságos környezet jellegéből adódóan nincs szükség a hozzáférésre engedélyezett állományokon erős elsődleges adatvédelmet (anonimizálást) végezni: az ilyen hozzáférések során a közvetlen azonosítók eltávolításán felül rendszerint nem történik egyéb felfedés elleni védelmi akció, bár indokolt esetben végrehajtható lehet (többek között, ha a közvetlen azonosítás lehetőségének megszüntetésével egyes statisztikai egységek továbbra is egyértelműen felismerhetők, például monopol vállalatok egyes tevékenységi szinteken).

A kutatósobai és a távoli hozzáférés külön hozzáférési csatornának tekinthető, bár működésileg mindkét csatorna azonos jellemzőkkel rendelkezik. A különbségtétel inkább szervezési okokra vezethető vissza; a hozzáférést kérő kutató saját preferenciáinak megfelelően kérhet hozzáférést a hozzá földrajzilag legközelebb elérhető vagy számára egyéb szempontból leginkább preferált akkreditált hozzáférési ponton. A távoli hozzáférés így lényegében a kutatósobai hozzáférés kiterjesztésének minősül, melynek igazi célja, hogy a kutatósoba budapesti elhelyezkedése révén ne legyenek kizárva a tudományos célú kutatásból olyan projektek, illetve kutatók, akik számára a földrajzi távolság jelenti a hozzáférés akadályát. Ennek megfelelően a KSH kutatósobai környezete nemcsak Budapesten, hanem 2014 óta Szegeden is elérhető.

Távoli végrehajtás. A távoli végrehajtás, bár technikailag biztonságos környezetbeli hozzáférésnek minősül, eltérő működési módot jelent, mint a kutatószobai vagy a távoli hozzáférés. A legnagyobb különbséget az adja, hogy távoli végrehajtás során a kutató az adatállományokat nem látja, illetve fizikailag nem kell felkeresnie a hozzáférési pontokat. A működés lényege, hogy a kutató specifikációját, illetve előre elkészített programkódját eljuttatja a KSH-hoz és a hivatal munkatársai azok alapján előállítják a kért kutatási eredményeket.⁹

A távoli végrehajtás működési környezete biztosítja, hogy a kutató a kutatószobai, illetve távoli hozzáférési környezettel egyenértékű minőségű állományokon kutathat. Az előállított kutatási eredményeket hozzáférhetővé tétel előtt felfedés elleni védelmi szempontból a KSH teljes körűen ellenőrzi.

4.2.3. KSH adat-hozzáférési csatornáinak összehasonlítása adatvédelmi, adat-hozzáférési és minőségi szempontból

A KSH hat adat-hozzáférési csatornájának minőség, illetve adatvédelmi, adat-hozzáférési vonatkozású összefoglalását mutatja a táblázat.

A KSH adat-hozzáférési csatornáinak összehasonlítása

Szempont	Táblázatos adatok kiadása	Anonimizált mikroadat-kiadás	Nyilvános mikroadatfájlok elérése	Kutatószobai hozzáférés	Távoli hozzáférés	Távoli végrehajtás
Előzetes adatvédelem erőssége	Erős	Közepes	Erős	Gyenge	Gyenge	Gyenge
Adat-hozzáférési lehetőség	Bárki számára	Kizárólag tudományos célra	Bárki számára	Kizárólag tudományos célra	Kizárólag tudományos célra	Kizárólag tudományos célra
Legfőbb előny	Rugalmas	Kutató az állományt „elviheti”, így megszokott környezetében dolgozhat	Szabadon letölthető felhasználási feltételek elfogadása esetén	Legrészletesebb állományokon való kutatás lehetősége	Legrészletesebb állományokon való kutatás lehetősége	Legrészletesebb állományokon való kutatás lehetősége
Legfontosabb korlátozó tényező	Anonimizálás torzító hatása	Anonimizálás torzító hatása; jogi kötelezettsége (célhoz kötöttség, megsemmisítés stb.)	Állományok erős felfedés elleni védelme miatt csak az állományok típusának megfelelő célra használhatók	Használatához a központi kutatószoba (Budapest) felkeresése szükséges	Használatához a távoli hozzáférési pont (jelenleg: Szeged) felkeresése szükséges	A kutató az állományt nem „látja”, többkörös egyeztetés-igény a programfájlok paraméterezéséhez
Minőség vonatkozás	Közepes	Közepes	Gyenge	Magas	Magas	Magas

⁹ Részleteket lásd: http://www.ksh.hu/tavoli_vegrehajtás

Az *előzetes adatvédelem erőssége* a módszertani adatvédelmi eljárások (felfedés elleni védelem) erősségére utal. Az „erős” jelző mutatja, hogy az adatállományokat minden esetben teljes körű felfedés elleni védelemnek kell alávetni. Mivel az érintett adat-hozzáférési csatornák bárki számára hozzáférhetők, ezért nincs érdemi lehetőség a felfedési forgatókönyvek mérlegelésére; az állományok minőségét (amit itt az állomány részletezettségével szinonim fogalomként használunk) korlátozza a felfedés elleni tevékenység. A „közepes” jelző szintén teljes körű felfedés elleni védelmi tevékenység ellátását takarja, azonban a célhoz kötöttség (adatkérő, illetve a felhasználási cél) ismeretében megfelelő felfedési forgatókönyvek határozhatók meg és a kutatói preferenciák ismeretében a felfedés elleni védelmi tevékenység kutatási eredményekre gyakorolt torzító hatása is jobban kezelhető (kulcsfontosságú változókat csak a legszükségesebb esetben védjük; ha az állomány védettsége más változók kezelésével is biztosítható). A „gyenge” szint arra utal, hogy nincs előzetes adatvédelem (anonimizálás), hanem a létrejövő kutatási eredmények adatvédelmi szempontú ellenőrzésére (ún. output checking-re) kerül sor, szintén teljes körűen.

A *minőség vonatkozás* ezzel összhangban azt mutatja, hogy a hozzáférhető állományok minősége (részletezettsége) milyen mértékű. Látható, hogy a legrészletesebb adatokkal a biztonságos környezetbe tartozó adat-hozzáférési csatornákon dolgozhatnak a tudományos célból hozzáférők. A legrészletesebb adatok felhasználásának lehetőségével viszont együtt jár, hogy ezek az adat-hozzáférési csatornák követelik meg a legtöbb előkészítést a kutató részéről (kutatási cél megfelelő indoklása, kifejtése, kutatói akkreditáció stb.).

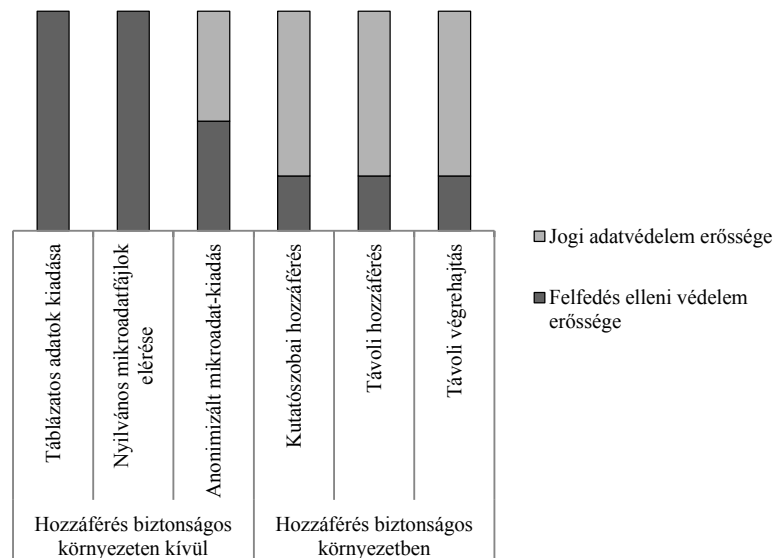
Összességében elmondható, hogy az adat-hozzáférési igények erőssége tükrében több adat-hozzáférési csatorna választási lehetősége biztosított a felhasználók, kiemelten a tudományos célból hozzáférni kívánó kutatók számára. Abban az esetben, ha a kutató a munkája során fontosnak tartja a megszokott munkakörnyezetet és hajlandó az állomány részletezettségében az anonimizálással járó „elszenvedett” minőségromlást elfogadni, úgy anonimizált mikroadat-kiadásra nyílik lehetősége. Amennyiben a legjobb minőségű (legnagyobb részletezettségű) állományokon kíván kutatni, úgy a kutatószobai, illetve a távoli hozzáférési, valamint a távoli végrehajtás adat-hozzáférési csatorna áll rendelkezésére.

4.2.4. Egyenkockázat elve

A KSH az elmúlt tíz év során deklarált célként tekintett az ún. egyenkockázat elvének érvényesítésére, mely a felfedési kockázat állandó alacsony szinten tartását jelenti, függetlenül attól, hogy a statisztikai adat milyen adat-hozzáférési csatornán jut el a felhasználóhoz. Ennek az elmúlt évtizedben követett gyakorlatnak is nagy szerepe volt az adat-hozzáférési csatornák meghatározásában és működési rendjük

kialakításában. A KSH hat adat-hozzáférési csatornája tükrében az egyenkockázat elvének érvényesülését mutatja a 4. ábra.

4. ábra. Az egyenkockázat elvének érvényesülése a KSH adat-hozzáférési csatornáin



Az egyenkockázat elve megköveteli, hogy a KSH adatállományaihoz a hozzáférés a különböző csatornákon felhasználói és adatvédelmi szempontból összehangolt legyen. Az összehangolt adatvédelem megteremtését a gyakorlatban a felfedés elleni védelmet biztosító eszközök és a jogi adatvédelmi eszközök együttes alkalmazásával érhetjük el. Általánosságban kijelenthető, hogy minél erősebb a jogi védelem, annál kevésbé szigorúan kell védeni az adatokat a felfedés ellen és fordítva.

Fontos cél, hogy minden esetben a lehető legnagyobb mértékben kihasználjuk a jogi védelem adta lehetőségeket, hogy ezáltal csak a szükséges mértékű felfedés elleni védelmet alkalmazzuk, mivel a módszertani adatvédelem a közzétett adatok körének korlátozásával, részletezettségének csökkentésével, azaz jellemzően minőségromlással járhat.

A 4. ábrán látható, hogy a biztonságos környezetű hozzáférés esetén lehetséges a legkisebb erősségű felfedés elleni védelmi megoldások alkalmazása. Ezeknél az adat-hozzáférési csatornáknál az adatvédelmet erős jogi (és kapcsolódó informatikai) garanciák erősítik, így a kontrollált környezetű hozzáférésnél a felfedés elleni védelem az érintett adatállomány közvetlen azonosíthatóságának megszüntetésével, illetve a készített kutatási eredmények utólagos adatvédelmi szempontú ellenőrzésével biztosított.

A másik „leglátványosabb” példa a bárki számára hozzáférhető adat-hozzáférési csatornáknál látható, ahol a jogi védelem erőssége alacsony, hiszen mind a táblázatos

adatok kiadása, mind a nyilvános mikroadatfájlok elérése esetén a felhasználási feltételek elfogadása az egyetlen érvényesíthető jogi adatvédelmi eszköz. Ennek megfelelően a szükséges adatvédelmet ezeknél az adat-hozzáférési csatornáknál erős felfedés elleni védelemmel kell biztosítani az adatvédelmi előírások érvényre juttatásához.

4.2.5. Az egyensúly biztosításának kulcstényezője: együttműködés és konszenzus

Az adat-hozzáférési és adatvédelmi egyensúly biztosításának eszköztára az említett stratégiai célkitűzéseken, módszertani irányelveken és diverzifikált tájékoztatási csatornákon felül további módokon is erősíthető és fejleszthető. Ez a kulcselem az adat-hozzáférési, adatvédelmi tevékenységben érintett szereplők párbeszéde, és az ezek eredményeként kialakítható konszenzus. Az Európai Statisztikai Rendszer ilyen irányú együttműködésének jó példái a nemzeti statisztikai szolgálatok, az adatarhívumok és kutatók együttműködésében képzelik el az adathozzáférés-adatvédelem jövőjét a hivatalos statisztikában (EC [2013], [2014]).

Bár az elmúlt tíz évben a módszertani tevékenységnek mindvégig kiemelt célja volt együttműködni a szakmai közösségekkel, az adathozzáférés-adatvédelem egyensúlyi problémája az adatvédelem területén is megköveteli az aktívabb és eredményesebb együttműködést a szakma résztvevőivel. Az adathozzáférés-adatvédelem eddig elért eredményeinek továbbfejlesztése, az egyensúlyi tényező jobb kiegyensúlyozása céljából a szakmai együttműködés ilyen formája elengedhetetlen.

5. Konklúzió

Az adatvédelmi elvárások teljesítése és az adat-hozzáférési igények lehetőségek szerinti maximális kiszolgálása a nemzeti statisztikai szolgálatok és így a KSH fontos célkitűzése. A két célkitűzés egyensúlyának biztosítása a KSH 2020-ig tartó stratégiájának fontos célja, melynek érvényre juttatását, gyakorlatban történő érvényesítését több eszköz is biztosítja.

Csak az utóbbi években merült fel az igény, és vált deklarált stratégiai céllá az adat-hozzáférési igény és az adatvédelmi kötelezettség egyensúlyának megteremtése. E törekvés megvalósítását szolgálják a minőségi irányelvek, illetve az adatvédelmi politika és az arra épülő megújult KSH adat-hozzáférési, adatvédelmi rendszerét modernizáló új belső utasítások. Az elmúlt évtized fejlesztéseinek egyik fő eredményként említhetjük az KSH hat adat-hozzáférési csatornáját. Az egyenkockázat elvének érvényre juttatásával az egyes adat-hozzáférési csatornák eltérő erősségű jogi, illetve módszertani adatvédelmi eszközök alkalmazását teszik szükségessé. Az

egyres adat-hozzáférési csatornákon az eltérő erősségű felfedés elleni védelmi megoldásokkal összhangban különböző minőségű (részletezettségű) állományokon nyílik lehetőség kutatásra. A stratégiai célkitűzésekkel összhangban pedig a tudományos célú kutatáshoz az egyes adat-hozzáférési csatornák közötti választás lehetősége is biztosított, a kutatók – a saját adat-hozzáférési igényeikhez mérten – több közül választhatják ki a számukra leginkább preferált adat-hozzáférési csatornát.

Az adat-hozzáférési-adatvédelmi tevékenység egyensúlyának biztosításához a KSH által a 2020-as évekig megfogalmazott stratégiai főirány céljai elengedhetetlenül fontosak, csakúgy, mint a felhasználókkal való aktív párbeszéd és közösen kialakított konszenzus, melyek az alkalmazott megoldások finomításának, illetve a folyamatban levő, valamint jövőbeli fejlesztési irányok meghatározásának, a közös elköteleződésnek elengedhetetlen feltételei.

Irodalom

- BARTUS T. [2013]: Adatcserével anonimizált mikroadatok használhatósága – Egy szimulációs vizsgálat tanulságai. *Statisztikai Szemle*. 91. évf. 5. sz. 465–497. old.
- EC (EUROPEAN COMMUNITY) [2013]: *Data without Boundaries*. Enlarging Cooperation: Conferences & Training Sessions. Deliverable D6.3: First Regional Workshop. 24–25 April. Ljubljana. http://www.dwbproject.org/export/sites/default/about/public_deliverables/dwb_d6-3_regional-workshop-report.pdf
- EC [2014]: *Data without Boundaries*. Enlarging Cooperation: Conferences & Training Sessions. Deliverable D6.6: Second Regional Workshop. 16–17 October. Athens. http://www.dwbproject.org/export/sites/default/about/public_deliverables/dwb_d6-6_regional-workshop2_report_final.pdf
- ERDEI V. – HORVÁTH R. [2004]: Az adatfelfedés elleni védelem statisztikai eszközei. *Statisztikai Szemle*. 82. évf. 8. sz. 705–727. old.
- FARAGÓ M. [2013]: Védett táblázatok morfológiája – Optimális másodlagos cellaelnyomás számítógép nélkül. *Statisztikai Szemle*. 91. évf. 10. sz. 947–970. old.
- HARCSA I. [2012]: „A mikroadatok hozzáféréssel és az adatok felfedés elleni védelmével kapcsolatos kérdésekről” címmel szervezett műhelykonferenciáról. *Statisztikai Szemle*. 90. évf. 11–12. sz. 1162–1164. old.
- KSH (KÖZPONTI STATISZTIKAI HIVATAL) [2013]: *27/2013. számú KSH utasítás a Központi Statisztikai Hivatal adatvédelmi szabályzatáról*. Budapest. http://www.ksh.hu/docs/szolgalattasok/adatigenyles/ksh_adatvedelmi_szabalyzat.pdf
- KSH [2014a]: *Adatvédelmi politika*. Budapest. http://www.ksh.hu/docs/bemutakozas/hun/adatvedelmi_politika_2014.pdf
- KSH [2014b]: *Kutatói tájékoztató – Útmutató a KSH kutatószobai környezetében folyó kutatómunkához*. Budapest. http://www.ksh.hu/docs/szolgalattasok/adatigenyles/ksh_kutatoj_tajekoztato.pdf
- KSH [2014c]: *Minőségi Irányelvek a Központi Statisztikai Hivatal folyamataira*. 3.1 verzió. Budapest. http://www.ksh.hu/docs/bemutakozas/hun/minosegi_iranyelvek_2014.pdf

- KSH [2014d]: *Stratégia 2020*. Budapest. http://www.ksh.hu/docs/bemutakozas/hun/strategia_2020.pdf
- KSH [2014e]: *A KSH minőségpolitikája, 2014*. Budapest. http://www.ksh.hu/docs/bemutakozas/hun/ksh_minosegpolitikaja_2014.pdf
- KRISTÓF P. [2015]: Nyilvános mikroadatfájlok összeállításának főbb jellemzői, különös tekintettel az adatvédelmi szempontokra. *Statisztikai Szemle*. 93. évf. 11–12. sz. 1112–1139. old.
- LAKATOS M. [2015]: Az adatvédelem és a statisztika kapcsolatának jogi szabályozása. *Statisztikai Szemle*. 93. évf. 11–12. sz. 1017–1050. old.
- NAGY B. [2015]: A célzott adatsere módszere a térstatisztikában. *Statisztikai Szemle*. 93. évf. 11–12. sz. 1152–1169. old.
- NAGY E. [2015]: A statisztikai adatvédelem és -hozzáférés szabályai az Európai Unióban. *Statisztikai Szemle*. 93. évf. 11–12. sz. 1051–1069. old.
- SZÉP K. [2012]: 2011 októberében volt 10 éve, hogy a KSH-ban létrejött egy központi módszertani egység. *Statisztikai Szemle*. 90. évf. 4. sz. 319–333. old.

Summary

Providing access to official statistical information is the ultimate goal of national statistical systems responsible for the development, production and dissemination of official statistics. The Hungarian Central Statistical Office is of no exception in this regard; it provides access to official statistical information both in large quantity and with great variety regarding the detail and modes of access. The datasets are the outputs of the statistical business processes, in which fulfilment of quality criteria and data protection regulations are two key objectives. The protection of confidential information provided for official statistics is a crucial factor in dissemination. This protection is ensured in practice by various legal, methodological and IT tools and solutions.

As an equally important factor, the needs for access to this vast amount of statistical information are realised and treated accordingly throughout the business processes, in which the role of microdata sets is more prominent than ever before.

The article is aimed at summarising the outcome and impact of the developments in the Hungarian Central Statistical Office of the last decade and presents the current state of the art regarding the practical solutions on how to ensure the balance between the data protection requirements and the needs for data access. This balance is realised by different data access channels in the practice of the Hungarian Central Statistical Office. Apart from the practical solutions, the article also discusses the drivers behind data access and data protection approaches.

Nyilvános mikroadatfájlok összeállításának főbb jellemzői, különös tekintettel az adatvédelmi szempontokra*

Kristóf Péter,
a KSH fogalmazója
E-mail: Peter.Kristof@ksh.hu

A tanulmány megírásával a szerző azt a célt tűzte ki, hogy mind elméleti, mind gyakorlati szempontból bevezető ismereteket nyújtson az Olvasó számára a magyar tudományos közéletben eddig kevésbé tárgyalt nyilvános mikroadatfájlokról. A cikk első két fejezete elméleti megközelítésűnek tekinthető, mivel a szerző ezekben a nyilvános mikroadatfájlok definícióját, alaptípusait és az előállítás során felmerülő kulcsproblémákat foglalja össze. A harmadik fejezetben a 2011-es magyarországi népszámláláson keresztül a gyakorlatban is illusztrálja, hogyan lehet nyilvános mikroadatfájlokat előállítani. Mindezzel nem titkolt célja az is, hogy a 2011-es magyarországi népszámlálásból készült nyilvános mikroadatfájlokat tovább népszerűsítse, és a minél szélesebb körű felhasználását ösztönözze.

TÁRGYSZÓ:
Anonimizálás.
Nyilvános mikroadatfájl.
Felfedési kockázat.

* A szerző köszönetét fejezi ki *Faragó Miklósnak, Kovács Marcellnek, Nagy Beátának, Vereczkei Zoltánnak* és *Waffenschmidt Jánosnének* a hasznos tanácsaikért, észrevételeikért.

Nyilvános mikroadatfájlokat (public use file) régóta publikálnak világszerte különböző célcsoportoknak más-más felhasználási célokra. A sokféle definíció és felhasználási cél a mikroadatfájlok anonimizálási módszereinek széles skáláját eredményezte. Bár a módszerek sokszínűek, a céljuk közös: biztosítani a nyilvános mikroadatfájlok megfelelő felfedés elleni védelmét. Az 1. fejezetben ismertetjük a nyilvános mikroadat-fájlok fő típusait és felhasználási lehetőségeit. A 2. fejezetben bemutatunk egy általános sémát a nyilvános mikroadatfájlok előállítására. Ez a séma több ponton is módosítható, bővíthető konkrét esetekben. Azokat a lépéseket és módszereket vettük számba, amelyeket a leggyakrabban alkalmaznak a gyakorlatban. Mindenesetre a megfelelő anonimizálási modell megtalálása mindig egy feltételes optimalizálási feladat: minimális felfedési kockázat biztosítása az adatkérő által még elfogadott információvesztés mellett.

A 3. fejezetben a 2011-es magyarországi népszámlálás példáján keresztül is illusztráljuk, hogyan készíthetők nyilvános mikroadatfájlok különböző célokra: programfájl tesztelésére, oktatási célokra, illetve kutatási célokra. A 2011-es népszámlálásból készült *tesztfájlok*, illetve *oktatási célú fájl* a KSH¹ első nyilvános mikroadatfájljai, melyeket 2014-ben illetve 2015-ben tett elérhetővé, ezzel megnyitva az adatok közzétételének egy új csatornáját.

1. A nyilvános mikroadatok fő típusai és felhasználási lehetőségei

A statisztikákat előállító állami szervezetek (jellemzően nemzeti statisztikai hivatalok) fő feladata, hogy a begyűjtött, illetve átvett adatokból jó minőségben, gyorsan, pontosan és hitelesen tájékoztasson a gazdasági és társadalmi folyamatokról. A tájékoztatás sokáig kizárólag kiadványokban megjelentetett aggregált statisztikák útján történt. Bár még ma is ez tekinthető a fő tájékoztatási formának, az 1950-es 1960-as évektől kezdve a „kész” statisztikák kiegészítéseként a mikroadatok² is egyre szélesebb és változatosabb formában váltak hozzáférhetővé. Az időpont nem véletlen: a mikroadatok elérését a számítástechnikai kapacitások rohamos bővülése és az infokommunikáció (internet) fejlődése tette lehetővé. A mikroadatok egy lehetséges csoportosítását mutatja az 1. táblázat.

¹ KSH: Központi Statisztikai Hivatal.

² Mikroadat: rekordok sorozatából álló állomány, amely megfigyelési egységek adatait tartalmazza.

1. táblázat

A mikroadatok csoportosítása

Tipus	Mire alkalmas?	Ki férhet hozzá?*	Az állomány kikerül-e az adatgazdától?
Anonimizált mikroadat	Kutatás	Intézményi háttérrel rendelkező kutatók	Igen
Biztonságos környezetben elérhető mikroadat	Kutatás	Kutatók	Nem
Nyilvános mikroadat	Oktatás/programfájl tesztelése/korlátozottan kutatás	Bárki	Igen

* A hozzáférés a jellemző nemzetközi (és a magyarországi) gyakorlatot tükrözi, de ebben országoként és nemzeti statisztikai hivatalonként lehetnek eltérések.

A mikroadatok közzétételének legnagyobb előnye, hogy a felhasználók tetszőleges elemzéseket készíthetnek, olyanokat is, amit az adott nemzeti statisztikai hivatal egyébként nem végezne el. A felhasználók (például kutatók) által publikált eredmények nem tekinthetők hivatalos statisztikának, de tovább bővítik a társadalmi és gazdasági folyamatokról való ismereteinket. Fontos megjegyezni, hogy a mikroadatok publikálása során sokkal nagyobb mértékben merülnek fel adatvédelmi problémák, mint a hagyományos aggregált adatok közzétételekor, mivel az adatok közzélése egyébként, az adatszolgáltatók szintjén történik. A mikroadatok anonimizálási módszereiről bővebben a 2. fejezetben lesz szó.

1.1. A nyilvános mikroadatok definíciója és típusai

Az első hivatalos, részben nyilvános mikroadatnak tekinthető állományokat az amerikai Social Security Administration's Office of Research and Statistics (Társadalombiztosítási Igazgatóság Kutatási és Statisztikai Hivatala) bocsátotta más szövetségi és állami hivatalok rendelkezésére 1962-ben. Ebben az időben a tárolás még szalagon történt, és az állományokat semmilyen adatvédelemmel nem látták el. Azért nevezhetők ezek az állományok részben nyilvános mikroadatnak, mert később nem kormányzati, független kutatók is hozzáférést kaphattak az állományokhoz. Az US Census Bureau (az Egyesült Államok Népszámlálási Hivatala) 1963-ban egyezrelekes mintát tett közzé az 1960-as népszámlálás és lakás-összeírásból (Census of Population and Housing). Az amerikai példát egyre több ország követi, a fejlett országokban található nemzeti statisztikai hivatalok többsége közzétesz nyilvános mikroadatfájlokat.

A nyilvános mikroadatfájlok három fő típusa különböztethető meg:

1. kutatási célú fájl (research file): PUMS, MF, PUMF, PUMSF;³
2. oktatási célú fájl (teaching (campus) file);
3. tesztfájl (test file).

A nyilvános mikroadatfájloknak nincs egységes elnevezése és definíciója. A nemzetközi gyakorlat sokszínűségét illusztrálja a fejezet végén található 2. táblázat is.

A kutatási célú fájlokra az *OECD*⁴ [2014] és az *ESR*⁵ (*Az Európai Unió Hivatalos Lapja* [2009]) terminológiáját emeljük ki, melyek nemzetközileg széles körben ismertek és elfogadottak.

a) Az OECD terminológiája szerint egy állomány kutatási célúnak tekinthető, ha:

- korlátozottan, de kutatási célra is alkalmas, a vizsgált jelenséget illetően érdemi következtetések vonhatóak le belőle;
- az adott szervezet által alkalmazott jogi és módszertani adatvédelmi előírásoknak megfelel;
- könnyen hozzáférhető (szabadon letölthető vagy regisztrációhoz/felhasználási feltételek elfogadásához kötött).

b) Az ESR a következő módon definiálja a kutatási célú fájlakat: „Egyedi statisztikai egységekre vonatkozó adatokat olyan anonimizált rekordokból álló, nyilvános használatú állomány formájában lehet közzétenni, amelyeket úgy készítenek elő, hogy a statisztikai egység ne legyen sem közvetlenül, sem pedig közvetve azonosítható, figyelembe véve minden olyan megfelelő eszközt, amelyet egy harmadik fél ésszerűen felhasználhat.”

Látható, hogy ez az Európai Unió rendelete által alkotott definíció az OECD terminológiával ellentétben kizárólag az állománytól elvárt adatvédelmi követelményt határozza meg módfelett általános módon, míg az állomány használhatóságának kritériumáról nem rendelkezik.

³ PUMS (public use microdata sample), MF (microdata files), PUMF (public use microdata file), PUMSF (public use micro stat files): a rövidítések a kutatási célú fájllok országspecifikus elnevezéseit takarják. További részleteket lásd a 2. táblázatban.

⁴ OECD (Organisation for Economic Co-operation and Development): Gazdasági Együttműködési és Fejlesztési Szervezet.

⁵ ESR: Európai Statisztikai Rendszer.

Az *oktatási célú fájlok* esetében már az elnevezés is sugallja, hogy elsősorban a közép- és felsőoktatási intézmények statisztikáival foglalkozó kurzusain lehet felhasználni az ilyen típusú állományokat, de hasznos lehet mindazok számára, akik a statisztikai módszerekkel szeretnének a gyakorlatban is megismerkedni. Általában csak korlátozott számú, a leginkább relevánsnak tartott változókat tartalmazza. Komolyabb kutatásra (szakdolgozat, disszertáció, tudományos cikk), általános érvényű, érdemi következtetések levonására nem alkalmasak, amit az állományokhoz mellékelt felhasználási feltételekben külön is ki szoktak hangsúlyozni.⁶ A fiktív adatokat tartalmazó állományokhoz képest az oktatási fájlok nagy előnye, hogy bár nem túl részletes, de valós (például népszámlálásból származó) adatok elemzését teszik lehetővé.

Az ún. *tesztfájlok* célja a kutatószobai vagy távoli hozzáférés során rendelkezésre bocsátott adatállományok struktúrájának előzetes megismerése, programkódok tesztelése. Ennek biztosítása érdekében ezeknek az állományoknak ugyanazokat a változókat kell tartalmazniuk, mint a biztonságos környezetben elérhető közvetlen azonosításra alkalmatlan mikroadatfájloknak. Mivel a tesztfájlok adatvédelmi szempontból a legerősebben⁷ védettek, általában még a legelemibb elemzési célokra sem használhatók.

A 2. táblázat a teljesség igénye nélkül tartalmaz néhány példát arra, hogy az egyes nemzeti statisztikai hivatalok milyen nyilvános mikroadatfájlokat tettek elérhetővé. A táblázat utolsó oszlopából azt is kiolvashatjuk, hogy bár a nyilvános mikroadatfájlok könnyen hozzáférhetők, de nem azonos módon. Van ahol előzetes regisztráció nélkül azonnal és díjmentesen letölthetjük az állományokat (jellemzően az Egyesült Államok gyakorlata ez), más esetben a letöltés regisztrációhoz kötött (például az Olasz Statisztikai Hivatalban), míg bizonyos hivatalok kizárólag egy meghatározott díj ellenében adják át az állományokat (például a Kanadai Statisztikai Hivatal). A KSH egyelőre egy *oktatási célú fájlt*,⁸ illetve *tesztfájlokat*⁹ tett elérhetővé a 2011-es népszámlálásból, melyeket a felhasználási feltételek elfogadásával bárki szabadon és díjmentesen elérhet a hivatal honlapján. A 2011-es népszámlálásból készült *kutatási célú fájlt*¹⁰ a KSH átadta a Minnesotai Egyetem által gondozott IPUMS¹¹ nemzetközi népszámlálási adatbázis projekt¹² részére. Ez várhatóan 2016 folyamán díjmentesen elérhető lesz a kutatók számára az IPUMS weboldalán.

⁶ Például: <http://www.forschungsdatenzentrum.de/en/campus-file.asp>

⁷ Például a 3.1. alfejezetben látni fogjuk, hogy egy tesztfájl egyes rekordjai nem is feltétlenül értelmezhetők, mivel a változók közötti kapcsolatok nem teljesülnek.

⁸ http://www.ksh.hu/nepszamlalas/oktatasi_mikroadat

⁹ <http://www.ksh.hu/nepszamlalas/tesztallomanyok>

¹⁰ A teljesség kedvéért megjegyezzük, hogy korábbi magyarországi népszámlálásokból (1970., 1980., 1990. és 2001.) is elérhetőek kutatási célra alkalmas fájlok az IPUMS oldalán, de ezekkel ebben a cikkben nem foglalkozunk.

¹¹ IPUMS (Integrated Public Use Microdata Series): egységesített nyilvános mikroadatfájlok adabázisa.

¹² <https://international.ipums.org/international/>

2. táblázat

Nemzetközi példák nyilvános mikroadatfájlokra

Ország	Felvétel	Szervezet	Fájl típusa	Év	Mire alkalmas?	Hozzáférhetőség
Egyesült Államok	Népszámlálás	Népszámlálási Hivatal	PUMS	1980, 1990, 2000	Oktatás, kutatás	Ingyenes, bárki számára letölthető 1 és 5 százalékos minta
	https://www.census.gov/main/www/pums.html					
	Időmérleg-felvétel	Munkaügyi Statisztikai Hivatal	MF	2003–2014	Oktatás, kutatás	Ingyenes, bárki számára letölthető
http://www.bls.gov/tus/data.htm						
Kanada	Nemzetközi utazási szokások	Statisztikai Hivatal	PUMF	2013	Oktatás, kutatás	Igénylés után díj ellenében adathordozón elvihető
	http://www.statcan.gc.ca/daily-quotidien/150528/dq150528i-eng.htm					
	Nemzeti háztartás felmérés	Statisztikai Hivatal	PUMF	2011	Oktatás, kutatás	Igénylés után díj ellenében adathordozón elvihető
http://www.statcan.gc.ca/daily-quotidien/141209/dq141209e-eng.htm						
Olaszország	Munkaerő-felmérés	Statisztikai Hivatal	PUMSF	2014	Oktatás, kutatás	Regisztráció után letölthető
	http://www.istat.it/en/archive/127804					
	PhD-fokozattal rendelkezők munkaerő-piaci integrációja	Statisztikai Hivatal	PUMSF	2009, 2014	Oktatás, kutatás	Regisztráció után letölthető
http://www.istat.it/en/archive/87789						
Németország	Mikrocenzus	Statisztikai Hivatal	Oktatási célú fájl	1976, 1996–1999 panel, 1998, 2002	Oktatás	Ingyenes, bárki számára letölthető 3,5, illetve 10 százalékos (1976) minta
	http://www.forschungsdatenzentrum.de/en/campus-file.asp					
	Adatok a munkáltatókról és az alkalmazottakról	Munkaerő- és Foglalkoztatási Kutatási Intézet	Tesztfájl	1993–2010	Programfájl tesztelése	Ingyenes, bárki számára letölthető
http://fdz.iab.de/en/Integrated_Establishment_and_Individual_Data/LIAB/Working_Tools.aspx						
Egyesült Királyság	Népszámlálás	Nemzeti Statisztikai Hivatal	Oktatási célú fájl	2011	Oktatás	Ingyenes, bárki számára letölthető 1 százalékos minta
	http://www.ons.gov.uk/ons/guide-method/census/2011/census-data/census-microdata/microdata-teaching-file/index.html					

Azt is érdemes megjegyezni, hogy a széles nyilvánosság számára közzétett mikroadatfájlok többsége lakossági felvételekből származik, bár van példa gazdaságstatisztikai felvételből készített nyilvános mikroadatfájllra is.¹³ Ezek azonban jellemzően erősebben védett¹⁴ tesztfájlok vagy oktatási célú fájlok. Ennek oka alapvetően az, hogy a gazdaságstatisztikai felvételekben sokkal magasabb a felfedési kockázat¹⁵ mértéke, amit a hagyományos anonimizálási módszerekkel¹⁶ nehéz jól kezelni.

2. A nyilvános mikroadatfájlok előállításának folyamata

Bár az első nyilvános mikroadatfájlok esetében még nem alkalmaztak semmilyen felfedés elleni védelmi módszert, manapság elképzelhetetlen lenne, hogy egy nyilvános mikroadatfájlt felfedés elleni védelem nélkül publikáljanak. A nyilvános mikroadatfájlok előállítási folyamatát lényegében az adatvédelmi szempontok vezérik. Az előállított nyilvános mikroadatfájlnak természetesen eleget kell tennie annak a célnak, amire készült: például statisztikai módszerek oktatása.

Nincs egységes és bevett gyakorlat arra, hogyan „kell” nyilvános mikroadatfájlokat előállítani. Kis túlzással azt lehet mondani, hogy ahány ország és nemzeti statisztikai hivatal, annyiféle módszertan és gyakorlat létezik. Az alkalmazott eljárásokat és módszereket nagymértékben az határozza meg, hogy milyen típusú nyilvános mikroadatfájlt készítenek és milyen felvételből.

A következőkben összefoglaljuk azokat a lépéseket, amelyeket minden esetben végig kell gondolni a nyilvános mikroadatfájlok elkészítése során. Az ismertetett folyamat alapvetően a lakossági felvételekből származó nyilvános mikroadatok előállítására vonatkozik, mivel a gyakorlatban jellemzően ilyen mikroadatfájlokkal találkozunk. Megjegyezzük, hogy az itt bemutatott folyamat viszonylag általános vezérfonalnak tekinthető, ami számos helyen módosítható, tovább finomítható konkrét esetekben. Részletesen itt nem fejtjük ki sem az egyes anonimizálási módszereket, sem a felfedési kockázat, illetve információvesztés mérését. Ezekről bővebben a szakirodalomban¹⁷ tájékozódhat az érdeklődő Olvasó.

2.1. Felfedési kockázat versus információvesztés

Az anonimizálás során a felhasználói igények (információvesztés minimalizálása) és az adatszolgáltatók adatvédelemhez fűződő jogos érdeke (felfedési kockázat

¹³ Például oktatási célú fájl a *Felmérés a keresetek szerkezetéről* (Német Statisztikai Hivatal).

¹⁴ Az erősebben védett azt jelentheti, hogy csak kevés változót publikálunk magas aggregáltsági szinten.

¹⁵ A felfedés fogalmát lásd a 2.1. alfejezetben.

¹⁶ Az anonimizálási módszerekről bővebben lásd a 2.2.4. szakaszt.

¹⁷ Jó összefoglalást ad a felfedés elleni védelem módszereiről például *Willenborg–de Waal* [2001].

minimalizálása) között kell megtalálni az egyensúlyt. Az információveszteség mérése azt jelenti, hogy a nyilvános mikroadatfájl egyes jellemzői milyen mértékben térnek el az eredeti mikroadat megfelelő jellemzőitől.¹⁸ Felfedés alatt azt szokták érteni, ha a támadó¹⁹ a közzétett mikroadat valamely adatszolgáltatóját képes beazonosítani, arról új információt kap vagy valamely jellemzőjére nagy valószínűséggel következtetni tud.

Az információveszteség és a felfedési kockázat közötti átváltást az 1. ábra szemlélteti. Először mindig meg kell határozni az adatvédő által tolerálható maximális felfedési kockázat szintjét (X), majd a felhasználó által tolerált maximális információveszteség szintjét (Y). A két pont által kijelölt A -val jelzett terület reprezentálja az ún. elfogadási tartományt. Az eredeti adatállomány az ábra szerint az $(1,0)$ pontban található: a felfedési kockázat mértéke 1, az információveszteség pedig 0. Világos, hogy ha az elfogadási tartomány egy pontja minél közelebb esik az origóhoz, az annál előnyösebb mind az adatvédő, mind a felhasználó számára.²⁰ A cél tehát az, hogy az anonimizálás eredményeként az elfogadási tartományon belül minél közelebb kerüljünk az origóhoz. Fontos megjegyezni azonban, hogy az elfogadási tartomány nem minden pontja feltétlenül meg is valósítható. Nem nehéz belátni, hogy a gyakorlatban nem létezik olyan anonimizálási módszertan, amely például egy olyan állományt eredményezne, amely zéró felfedési kockázattal és információveszteséggel lenne jellemezhető.²¹ A megvalósítható tartomány tehát az A -nak csak egy olyan részalmegegyesítője lehet, ami nem tartalmazza A -nak a tengelyekkel közös részét.

Feltételezhető, hogy ha egy anonimizálási modell különböző paraméterei mellett állítjuk elő a kiadásra szánt állományt, akkor a modell az 1. ábra által meghatározott koordinátarendszerben egy negatív meredekségű konvex görbével jellemezhető. Ez azt jelenti, hogy az információveszteség csökkentése csak nagyobb felfedési kockázat árán lehetséges és fordítva. Egy anonimizált állomány akkor kerülhet kiadásra a felhasználó részére, ha létezik olyan anonimizálási modell, amelyhez tartozó konvex görbének van legalább egy közös pontja A -val. Az 1. ábrán két anonimizálási modellt ($m1$ -t és $m2$ -t) tüntettünk fel. Könnyen belátható, hogy $m2$ preferált $m1$ -gyel szemben, mivel tetszőleges információveszteségi szint mellett alacsonyabb felfedési kockázat tartozik hozzá és fordítva. Bárhogy is kalibráljuk az $m1$ modellt, az adatvédő és a felhasználó sosem jut egyezségre: vagy az információveszteség mértéke vagy a felfedési kockázat mértéke lesz túlzottan nagy, illetve a kettő egyszerre. Más szóval az $m1$ szerint anonimizált állomány valamelyik félnek sosem elfogadható, így az állomány nem kerül kiadásra. Ezzel szemben, ha az $m2$ modellt úgy kalibráljuk,

¹⁸ Erről részletesebben lásd a 2.2.5. szakaszt.

¹⁹ Olyan felhasználó, aki megkísérli valamely, a mikroadatban szereplő válaszadó(ka)t beazonosítani, majd a róla nyert bizalmas információkat közzéteszi. A támadó motivációja lehet a statisztikai hivatalba vagy a kormányzat egészébe vetett bizalom csökkentése, illetve a kinyert információ által nyereség realizálása.

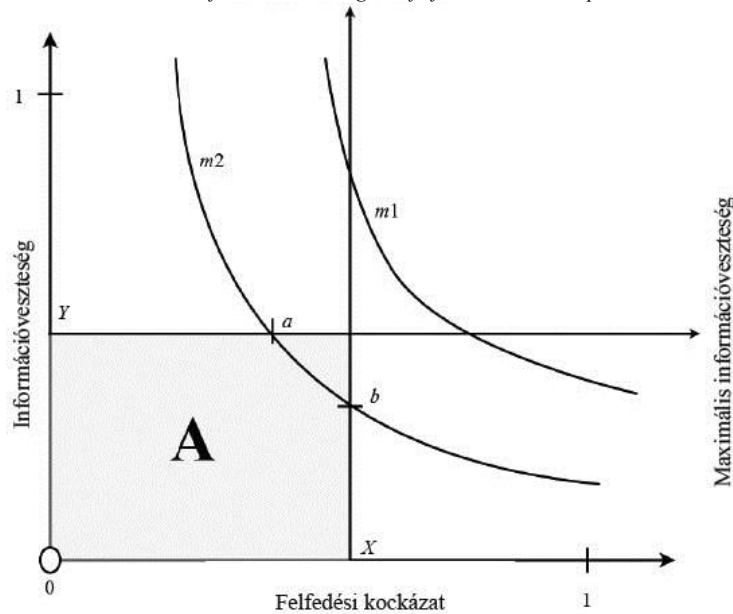
²⁰ Ez esetben mind a felfedési kockázat, mind az információveszteség mértéke alacsonyabb.

²¹ Az 1. ábrán ezt a pontot éppen az origó képviseli.

hogy éppen az ab szakaszon helyezkedjünk el, akkor mind az adatvédő, mind a felhasználó számára elfogadható eredmény születhet. Világos, hogy a felhasználó érdeke a b pont elérése, míg az adatvédő számára az a pont preferált. Viszont, ha az adatvédő számára az is hasznos, hogy a felhasználó a lehető legkisebb információvesztéssel szembesüljön, akkor hajlandó lesz „engedni”, s így végül valahol az a és a b pont között is megegyezhetnek.

Megjegyezzük, hogy az itt leírt fejtegetés csak egy meglehetősen absztrakt és leegyszerűsített elméleti konstrukciónak tekinthető. Ezzel a célunk mindössze annyi, hogy demonstráljuk azt, hogy az anonimizálás nem más, mint egy feltételes optimalizálási feladat. A gyakorlatban a maximális felfedési kockázat szintjét nem lehet egzaktan és teljes mértékig objektíven meghatározni, de a felhasználótól sem várható el, hogy pontosan definiálja a maximális információvesztés szintjét.

1. ábra. Az információvesztés és a felfedési kockázat kapcsolata



Forrás: Saját szerkesztés.

2.2. A nyilvános mikroadatfájl előállításának lépései

A következő szakaszokban bemutatjuk a nyilvános mikroadatfájlok előállításának során felmerülő legfontosabb kérdéseket, majd a fejezet végén felvázoljuk a nyilvános mikroadatfájlok egy lehetséges előállítási folyamatát.

2.2.1. A felfedési forgatókönyvek meghatározása

A nyilvános mikroadatfájl létrehozásának első lépése, hogy feltérképezzük a potenciális felhasználók körét, a lehetséges felhasználási célokat és igényeket, illetve az elérhető külső állományokat. Már ezekből is hasznos információkat nyerhetünk ahhoz, hogy meghatározzuk a felfedési forgatókönyveket, a nyilvános mikroadatfájl maximálisan tolerálható kockázati szintjét és a tipikus, felhasználó által tolerált maximális információveszteség szintjét. (Lásd az 1. ábrát.) Például egy szabadon letölthető oktatási célú fájl nyilvánosságából adódóan sokkal nagyobb adatvédelmi kockázatot jelent, mint egy korlátozott kutatói kör számára elérhető kutatási célú fájl. Az is világos, hogy a teszttállomány esetében jelentősen nagyobb információveszteség is elfogadható a felhasználó számára, mint a kutatási célú fájl esetében.

A leggyakoribb felfedési forgatókönyvek a következők:

– *Külső archívum*: a támadó olyan külső adatállománnyal rendelkezik, amit adott kulcs²² alapján hozzá tud rendelni a nyilvános mikroadatfájlhoz. Az állományok összekapcsolása a gyakorlatban legtöbbször valamilyen közvetlen azonosító²³ (például név, cím) alapján lehetséges.

– *Spontán beazonosítás*: a támadó valamely kulcskombináció értékeihez külső információt társítva beazonosítja az adatszolgáltatót (például egy híres személyt), illetve új információt tud meg róla.

– *Kíváncsi szomszéd*: a támadó valamely közeli ismerősét (szomszédját) fedi fel azáltal, hogy az általa ismert jellemzőkből képzett kulcskombinációkat megtalálja a nyilvános mikroadatfájlban.

2.2.2. Az állomány változóinak kategorizálása, kulcsok képzése

A következő lépés, hogy az anonimizálandó állomány(oka)t részletesen is áttanulmányozzuk, elsősorban felfedés elleni védelmi szempontból. Érdemes az állományban található változókat a következő kategóriákba sorolni:

I. Adatvédelmi szempontú kategorizálás

– *közvetlen azonosítók* (például társadalombiztosítási azonosító jel, vállalat/személy neve);

²² Kulcs: a mikroadatállomány néhány változójának vagy változó kategóriájának kombinációja.

²³ Közvetlen azonosító: a statisztikai egységhez rendelt egyedi azonosítási kód (függetlenül annak nyilvános hozzáférhetőségétől), illetve a statisztikai egység megnevezése/neve, a hozzá tartozó pontos címadat (lakcím, székhelycím, telephelycím stb.), valamint elérhetőségi adatai (e-mail cím, telefonszám stb.). Nem minősül közvetlen azonosítónak az anonimizálást végző által képzett technikai azonosító.

- *közvetett azonosítók* (például lakóhely települése, kor, nem, nemzetiség, iskolai végzettség, foglalkoztatási státusz, foglalkozás);
- *érzékeny változók* (például egészségi állapot, nemzetiség, valóság, jövedelem, politikai elkötelezettség);
- *kulcsok* (például kor-nem-nemzetiség-iskolai végzettség).

II. Információvesztés szempontú kategorizálás

- a felhasználó számára *fontos* változó,
- a felhasználó számára *irreleváns* változó.

A közvetlen azonosítókat minden esetben törölni kell, hogy a külső archívumként hivatkozott felfedési forráskönyvben felvázolt összekapcsolási kísérleteket meghiúsítsuk.

A közvetett azonosítók azok, amelyek alapján a leginkább valószínű, hogy felfedés történik. Érzékenyek tekinthetők azok a változók, melyek közzététele különösen súlyos hátránnyal járhat az adatszolgáltatóra nézve.²⁴ Az hogy mi számít érzékenynek, függ az adott ország kulturális, történelmi hagyományaitól és az ebből fakadó jogi szabályozástól is. Az érzékeny változók körét a felvétel szakértőjével közösen kell meghatározni. Fontos megjegyezni: attól, hogy egy változó érzékeny, nem feltétlenül azonosít is önmagában²⁵ és fordítva: az azonosításra alkalmas információ (például kor, nem) sem számít minden esetben érzékenynek.

Kulcsnak a közvetett azonosítók és az érzékeny változók köréből képzett változóhalmazt tekintjük. A változóhalmaz egy elemét kulcsváltozónak nevezzük.

Nyilvános mikroadatfájlok esetén bizonyos közvetett azonosítókat²⁶ vagy érzékeny²⁷ változókat is törölni szoktak, erre azonban nincs egységes gyakorlat. A kulcsváltozók meghatározásakor szintén érdemes a felvétel szakértőjével konzultálni. Megjegyezzük, hogy a kulcsváltozók kiválasztása szorosan összefügg azzal, hogy melyik felfedési forráskönyv(ek)et vesszük figyelembe. Ha feltételezzük, hogy bizonyos jellemzők, amelyek szerepelnek az anonimizálandó állományban, a támadó számára is ismertek (külső archívum forráskönyv), akkor célszerű ezeket a változókat bevonni a kulcsváltozók körébe. Általában azt szokták javasolni, hogy a változókból több különböző tartalmú és/vagy hosszúságú kulcsot is képezzünk, s hasonlítjuk össze a felfedési kockázatukat. A felfedés szempontjából legrelevánsabb kulcs kiválasztásában is érdemes tanácsot kérni az adott felvétel által érintett téma szakértőjétől. Amennyiben lehetséges, a potenciális felhasználók körében is ajánlott fel-

²⁴ Például a munkáltató megtudja, hogy az alkalmazottja valamilyen súlyos, krónikus betegségben szenved.

²⁵ Például, ha közlik, hogy 2014-ben mennyi volt a leggazdagabb magyar állampolgár jövedelme, az érzékeny információ, de az iparág és pontosabb földrajzi lehatárolás hiányában az adatszolgáltató nem feltétlenül azonosítható.

²⁶ Például településszintű területi bontás helyett csak országos adatok állnak rendelkezésre.

²⁷ Például Hollandia nem publikál jövedelmi adatokat nyilvános mikroadatfájlokban.

mérni, hogy számukra melyek a legrelevánsabb változók. Ha adatvédelmi szempontból lehetséges, törekedni kell arra, hogy a felhasználók számára fontos változók a lehető legkisebb információvesztéssel szerepeljenek a nyilvános mikroadatfájlból.

2.2.3. A felfedési kockázat mérése

Ha kiválasztottuk a kulcsváltozókból képzett kulcsokat, akkor a felfedési kockázatot kell megmérni.

A kockázat lehet *egyedi* (rekord szintű) vagy *globális* (az egész adatállományt jellemző) kockázat. Továbbá a módszerek megkülönböztethetők attól függően, hogy *kategorikus* kulcsváltozókra vagy *folytonos* kulcsváltozókra alkalmazzuk őket.

I. Az egyedi kockázatot mérő mutatók. A leggyakrabban használt kategorikus változókra vonatkozó egyedi kockázat mértékek a következők.

a) Egyedi kockázat számítása a mintasúlyok segítségével. Tegyük fel, hogy a „nem-életkor-foglalkozás-iskolai végzettség” kulcsot határoztuk meg. Az anonimizálandó mikroadatfájlból meg kell keresni a kulcs összes előforduló kombinációját (például 40 éves felsőfokú végzettségű férfi fogorvos). Minden egyes kombinációra ismerjük a mintabeli gyakoriságot, amiből a súlyok segítségével kiszámíthatjuk a várható sokasági előfordulások számát is. A módszer minden egyes rekordhoz hozzárendel egy egyedi értéket, ami az adott rekord kockázatát fejezi ki. Azok a rekordok lesznek kockázatosak, amelyeknél a vizsgált kulcs egy kombinációjára nemcsak a mintában, hanem a sokaságban is alacsony előfordulást tapasztalunk.

b) *k* anonimitás. Egy állomány akkor teljesíti a *k* anonimitás kritériumát, ha a vizsgált kulcs minden kombinációjára legalább *k* előfordulás van a mintában.

c) *l* diverzitás. Tegyük fel, hogy az *a)* pontban ismertetett kulcsra felírt kombinációt vizsgáljuk, továbbá a mintában három 40 éves felsőfokú végzettségű férfi fogorvos van. Tétélezzük fel azt is, hogy az állományban szerepel egy érzékeny változó (vallás), amit nem soroltunk a kulcsváltozók közé. Amennyiben a *k* 40 éves felsőfokú végzettségű férfi fogorvos mind református vallású, akkor – bár az adott kulcskombináció nem sérti meg a *k* anonimitást, ha *k=3*-as küszöböt választunk –, de megsérti az *l* diverzitást, mert mind a három egyed ugyanazt az értéket veszi fel az érzékeny változóra.

A folytonos kulcsváltozók esetében a felfedési kockázatot csak a már levédett változóval összehasonlítva, utólag szokták kalkulálni. Például egy folytonos jövede-

lem vagy árbevétel változóhoz hozzáadnak egy zajváltozót, majd kiszámítják az eredeti és a zaj hozzáadása utáni értékek közötti távolságot egy megfelelő távolságmérték (például Mahalanobis-távolság) felhasználásával. Amennyiben a perturbált érték kívül esik egy előre meghatározott intervallumon, akkor a felfedési kockázat nem jelentős. Outlierek esetében (például kiugró árbevétel) érdemes nagyobb intervallumot meghatározni, mivel ezek felfedési kockázata magasabb.

II. A globális kockázat mérése. A globális kockázat az állomány egészére vonatkozó kockázati mérőszám. Relatív és abszolút mérőszámot is alkalmazhatunk. A relatív globális kockázati mérőszámot benchmark alapúnak is szokás nevezni, mivel az egyedi kockázatok eloszlásának valamilyen kitüntetett értékéhez viszonyít. Például hány olyan megfigyelés van, ami nagyobb, mint az egyedi kockázatok átlaga vagy mediánja. A benchmarkalapú mérőszám az egyedi kockázatok eloszlásának kiugró értékeit keresi, és ezek arányát viszonyítja az eloszlás „jól viselkedő” vagy átlagos részéhez. Az abszolút mérőszám az egyedi kockázatokot összegzi, és ebből számítja ki az azonosítható rekordok várható értékét.

2.2.4. Anonimizálási módszerek

A felfedési kockázat felmérése után olyan anonimizálási módszereket kell alkalmazni, melyek elegendő mértékben csökkentik az állomány felfedési kockázatát. Visszautalva az 1. ábrára: az *A*-val jelzett elfogadási tartomány valamely, origóhoz minél közelebb eső pontjának elérése a cél. A leggyakrabban alkalmazott anonimizálási módszereket a 3. táblázat foglalja össze.

3. táblázat

A leggyakrabban használt anonimizálási módszerek

Változó	Determinisztikus	Sztocasztikus
	módszer	
Kategorikus	Globális átkódolás, cellaelnyomás	Sorrendcsere, PRAM ²⁸
Folytonos	Mikroaggregálás	Zaj hozzáadása, globális átkódolás

a) *Globális átkódolás.* Alkalmazható folytonos változókra is, de általában kategorikus változókat szoktak átkódolni. A cél, hogy az átkódolt változó kimeneteinek számát csökkentjük. Folytonos változóra gyakori példa a jövedelem vagy az árbevétel előre megadott kategóriákba sorolása, míg kategorikus változókra a kor korcsoportokba átkódolását lehet megemlíteni.

²⁸ PRAM (post randomization method): utólagos randomizációs módszer.

A globális átkódolás speciális esete az alsó/felső kódolás. Például a lakás alapterületet felső kódoljuk 300-as négyzetméter értéken, akkor az így képzett új kategória a 300 négyzetméteres vagy annál nagyobb lakásokat foglalja magába.

b) *Cellaelnyomás*. Kategorikus változókra alkalmazzák, ami nem jelent mást, mint egy adott cella értékének törlését, majd valamely egyezményes jellel (például három ponttal) történő helyettesítését. Leggyakrabban a globális átkódolás után alkalmazzák annak érdekében, hogy a nyilvános mikroadatfájl k anonimitását biztosítsák. A cellaelnyomás különböző algoritmusok szerint történhet:

– különböző súlyt lehet adni az egyes kulcsváltozóknak: például a nagyobb azonosító erővel rendelkező változókban relatíve több cellát nyomunk el;

– csak a minimálisan szükséges cellát nyomjuk el.

c) *Sorrendcsere*. Ez a módszer bármilyen numerikus változóra alkalmazható. Az eljárás első lépése, hogy minden csere alá vont változót növekvő sorba rendezünk egymástól függetlenül. Jelölje X_i és X_j az X változó i . és j . elemét. Minden sorba rendezett változó esetén adott i . értéknek keresünk egy j . párt, majd a két értéket egymással felcseréljük. A csere korlátozott abban az értelemben, hogy nem cserélhetünk egy adott értéket bármilyen másik értékre. Általában meg szokták adni a rekordok arányában kifejezett $p\%$ értéket, ami azt mutatja, hogy az i . elem legfeljebb hány rekord távolságra lehet a kicserélendő j . elemtől.

d) *PRAM*. Lényege, hogy a kategorikus változók által felvett értékeket előre megadott valószínűséggel egy másik értékre változtassunk meg. Az is lehetséges, hogy adott érték nem változik meg, mert saját magára „cseréljük ki”. Az egyes kategóriák másikká való átkódolásának valószínűségeit egy ún. átmenetmátrixban tároljuk.

e) *Mikroaggregálás*. Ennek során egy folytonos változót előre megadott csoportokra bontunk, majd a csoporton belül aggregáljuk. Az aggregálás leggyakrabban azt jelenti, hogy adott csoportba tartozó értékeket a csoportátlaggal helyettesítjük.

f) *Zaj hozzáadása*. Ezt a módszert folytonos változókra szokták alkalmazni. Főbb típusai:

– Korrelálatlan véletlen zaj: a hozzáadott zajváltozó arányos az eredeti változó szórásával. A zajváltozót jelölje ε , míg az eredeti változó legyen x . Mivel $E(\varepsilon) = 0$, ezért teljesül, hogy $E(x) = E(x + \varepsilon)$, vagyis az eredeti változó és a „zajosított” változó átlaga ugyanaz marad. Könnyen megmutatható viszont, hogy a zajos változó variációjára

és egy másik változóval vett kovarianciája nem egyezik meg az eredeti változóéval.

– Korrelált zaj: előnye, hogy a kovariancia-mátrixok konzisztensen becsülhetők.

2.2.5. Az információveszteség mérése

Miután a felfedési kockázatnak megfelelő adatvédelmi módszereket alkalmaztuk az állományon, fel kell mérni, hogy milyen hatással voltak ezek az állomány minőségére, amely ebben a kontextusban annak „használhatóságára” utal: a nyilvános mikroadatfájl elvégzett számítások, elemzések várhatóan milyen mértékű torzítást tartalmaznak ahhoz képest, mintha az eredeti, „nyers” mikroadatfájl állt volna rendelkezésre. Az információveszteség mértékéről azért fontos tájékoztatni a felhasználót, hogy tisztában legyen, az általa kitűzött kutatási célra egyáltalán alkalmas-e az adott nyilvános mikroadatfájl, és ha igen, akkor várhatóan milyen mértékű torzítással kell számolnia a kapott eredményeket illetően. Az információveszteségi mutatókat két csoportba lehet sorolni:

- a) általános mutatók;
- b) állomány-specifikus mutatók.

Az *általános mutatóra* lehet példa, ha a felhasználók által gyakran használt változók eloszlását/megoszlását hasonlítjuk össze. *Állomány-specifikus* mutató lehet például a MEF²⁹ esetében a nemenként és korcsoportonként kiszámított munkanélküliségi vagy aktivitási ráta. Az összehasonlítás történhet egyszerűen az abszolút relatív különbségek kiszámítása révén.

Drechsler [2009] modellalapú összehasonlítást is javasolt. Ennek lényege, hogy a felhasználók által gyakran kutatott jelenséget modellezzük, és a kapott eredményeket összevetjük. Például a MEF esetében a munkanélküliség valószínűségét modellezhetjük logisztikus regresszióval, majd a becsült együtthatók konfidencia-intervallumait összevetjük.

2.2.6 A nyilvános mikroadatfájlok előállításának főbb lépései

A 2. ábra a 2. fejezetben eddig elmondottakat foglalja össze.

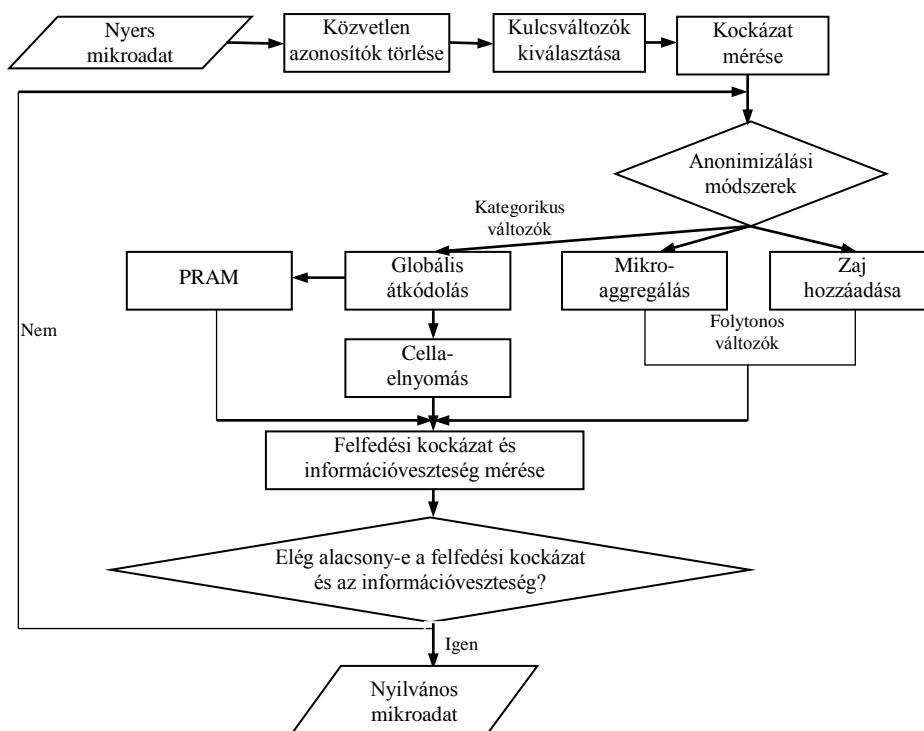
A bemutatott folyamatséma alapján az *R* sdcMicro³⁰ vagy a μ -Argus³¹ program-csomagokkal előállíthatók a nyilvános mikroadatfájlok.

²⁹ MEF: munkaerő-felmérés.

³⁰ <https://cran.r-project.org/web/packages/sdcMicro/index.html>

³¹ <http://neon.vb.cbs.nl/casc/mu.htm>

2. ábra. A nyilvános mikroadatfájl előállításának egy lehetséges folyamatsémája



Megjegyzés.

a) A kulcsváltozók kiválasztásánál és a felfedési forgatókönyv meghatározásánál (például elérhető külső állományok) érdemes a szakterület specialistájával konzultálni annak érdekében, hogy a reális felfedési forgatókönyv(ek)et vegyük figyelembe.

b) PRAM-ot akkor ajánlott használni, ha sok kulcsváltozónk van és sok egyedi előfordulást tapasztalunk az állományban.

c) Cellaelnyomást akkor érdemes a globális átkódolás után alkalmazni, ha relatíve kevés egyedi kulcskombináció marad az átkódolás után.

d) A 2. ábra nem tartalmazza a részben/teljesen szintetikus állomány generálását. Erről bővebben a szakirodalomban (például Drechsler [2011]) tájékozódhat az Olvasó.

Forrás: Saját szerkesztés Templ–Meindl–Kowarik [2014] nyomán.

3. Nyilvános mikroadatfájlok a KSH-ban

A mikroadatok iránti igény a 2000-es évek második felétől kezdve jelentkezett egyre erőteljesebben a KSH felé is. 2007-ben nyílt meg a hivatal kutatószobája,³²

³² http://www.ksh.hu/kutatooszobai_hozzaferes

ahol tudományos célból, közvetlen azonosításra alkalmatlanná tett mikroadat-állományokhoz férhetnek hozzá kutatók. A kutatószoba azonban meglehetősen szűk nyilvánosságot, a tudományos, akadémiai szféra igényeit szolgálja ki. A KSH azok számára is elérhetővé kívánja tenni a mikroadatfájlokat, akik nem feltétlenül számítanak a kutatószoba szűk célközönségébe. A hivatal a 2014-es év folyamán tesztfájlokat, 2015-ben pedig egy oktatási célú fájlt publikált a 2011-es népszámlálásból, továbbá a Minnesotai Egyetem számára kutatási célra alkalmas nyilvános mikroadatfájlokat adott át az IPUMS-projekt keretében. A következőkben ezt a három projektet mutatjuk be röviden. Az érdeklődő Olvasó bepillantást nyerhet abba, hogy milyen anonimizálási elvek és módszerek alkalmazásával álltak elő a megfelelő mikroadatfájlok.

3.1. Népszámlálási tesztfájlok

A népszámlálási tesztfájlok előállításának igénye még szorosan kötődik a kutatószobai használathoz. Ezeket a fájlokat ugyanis azok a kutatók tudják hasznosítani, akik a részletes népszámlálási mikroadatok kutatása előtt kívánják előzetesen megismerni a kutatószobai állomány felépítését, változóit. A tesztfájlok lehetőséget nyújtanak arra is, hogy a programkódokat akár otthon, előre megírják, tesztelhesék, mielőtt az éles kutatószobai állományokon futtatnák le azokat. Mivel a teszttállományok fiktív rekordokat tartalmaznak, ezért nem alkalmasak bárminemű elemzés elkészítésére, illetve érdemi következtetések levonására.

A kutatószobában elemezhető teljes népszámlálási mikroadatokat egymással összekapcsolható, négy (személyi, lakás-, háztartás- és család-) állomány formájában állnak rendelkezésre. A teszttállomány legfontosabb jellemzője, hogy a kutatószobai állomány struktúrájával a lehető legnagyobb mértékben megegyezik, ezért a teszttállományok is ugyanabból a négy állományból tevődnek össze, mint a kutatószobában elérhető mikroadatfájlok. Első lépésben a személyi állomány magánháztartásokat tartalmazó címlistájából kiválasztottunk ezer címet egyszerű véletlen módon, visszatevés nélkül. A mintavétel rétegzést nem tartalmazott. Ehhez az ezer címhez rendeltük hozzá a személyi, háztartás-, család- és lakásváltozókat. Az ezer címhez 2 437 személy tartozott, akik 1 059 háztartást és 646 családot alkottak. Az állományok anonimitását úgy biztosítottuk, hogy a célváltozók értékeit egymástól függetlenül megkevertük. Ez azt jelenti, hogy egy adott változó értékeit egy normális eloszlású véletlen változó szerint rendeztük növekvő sorrendbe. Így jöttek létre az előző bekezdésben említett fiktív rekordok. A továbbiakban egy példán keresztül szemléltetjük, hogyan változhat meg egy személyi rekord, miután a változók értékeit megkeverjük.

A példa változói a következők:

- MEGYE: KSH megyeazonosító (lakhely),
- TELTIPUS: település típusa (lakhely),
- REGIO: régió megnevezése (lakhely),
- LNCSOP: népesség-nagyságcsoport (fő),
- NEME: nem,
- SZEVI: születési év,
- HO: születési hónap,
- NAP: születési nap,
- CSPOT: családi állapot,
- EGYERUJ: élve született gyermekeinek száma,
- IRELSZ: az iskolarendszerben elvégzett legmagasabb szint,
- GAKT: gazdasági aktivitás.

4. táblázat

A népszámlálási tesztfájlból alkalmazott véletlen keverés illusztrálása

Változó	Fiktív	Valós
	kód	
MEGYE	Budapest (01)	Budapest (01)
TELTIPUS	Nagyközség, község	Budapest kerületei
REGIO	Észak-Magyarország	Közép-Magyarország
LNCSOP	5 000–9 999 fő	Budapest
NEME	Férfi	Nő
SZEVI	2001.	1985.
HO	1.	2.
NAP	10.	23.
CSPOT	Házas	Házas
EGYERUJ	5 fő	2 fő
IRELSZ	Érettségi vagy középfokú szakképesítés	Egyetemi oklevél
GAKT	Gyes, gyet, gyed	Gyes, gyet, gyed

A 4. táblázatból jól látszik, hogy a keverés eredményeként nem feltétlenül változik meg minden érték (lásd félkövérrel szedett változók). Mivel a személyi állomány jelenleg 101 változót tartalmaz, de még a legkisebb változókészlettel rendelkező háztartás állományban is 25 változó található, ezért nincs reális kockázata annak, ha egy rekordon belül esetleg néhány érték változatlan marad. Ráadásul a támadó sosem lehet biztos abban, hogy egy adott rekord mely értéke fiktív és melyik valós. Hang-

súlyozzuk, hogy e példa kitalált, még a valósan nevezett rekord sem a 2011-es népszámlálásból származik, csak egy olyan kitalált személy, akire a felsorolt változók közötti logikai kapcsolatok teljesülnek, így akár a valóságban is létezhet. Világos, hogy a fiktívnek elnevezett rekordhoz tartozó személy biztosan nem létezik, mert a változók közötti több logikai összefüggés sem teljesül (például, ha MEGYE Budapest, akkor a TELTIPUS nem lehet község).

3.2. Népszámlálásból készült oktatási célú fájl

Az oktatási célú fájl közzétételének fő motivációja, hogy a KSH nyisson a szélesebb felhasználói kör felé. Ahogy a bevezetőben is említettük, elsősorban a felsőoktatás statisztikai kurzusainak gyakorlatain használhatják a hallgatók az állományt, de természetesen bárki letöltheti, aki egy valós állományon keresztül szeretne statisztikai módszerekkel ismerkedni. Az oktatási célú fájl részletezettsége jelentősen alacsonyabb a KSH kutatószobájában biztonságos környezetben kutatható 10 százalékos személyi- és lakásmintához képest. Nagy előnye viszont az, hogy könnyen hozzáférhető: a felhasználási feltételek elfogadása után a hivatal honlapjáról bárki szabadon letöltheti.

A 10 ezer elemű személyi minta a teljes népszámlálási lakónépességből (9 937 628 főből) választott valószínűségi minta. A személyek kiválasztása rétegzés nélkül, egy lépcsőben, visszatevés nélküli egyszerű véletlen módon történt. E kiválasztásnak köszönhetően a mintából számított becslések pontossága könnyedén jellemezhető a jól ismert képlet szerinti szórás segítségével.

A változók száma jóval korlátozottabb, mint a tesztfájlok esetében, mindössze 13 személyi és lakásjellemző szerepel az állományban. (A változókat lásd az 5. táblázat első oszlopában.)

Az alacsony kiválasztási arány (körülbelül 1%) és a korlátozott számú változó bevonása már önmagában is jelentős felfedés elleni védelmi beavatkozásnak tekinthető. További felfedés elleni védelmet jelent, hogy ezek a változók módosításokon (alsó/felső kódolás, globális átkódolás) estek át ahhoz képest, ahogy a részletes kutatószobai állományban szerepelnek. Ez azt jelenti, hogy kevésbé részletes formában érhetők el, mint a teljes kutatószobai állomány esetében. Az 5. táblázatban láthatók az oktatási célú fájl változói és a kutatószobában elérhető állományban szereplő azonos tartalmú változók közötti eltérések.

Megjegyezzük, hogy a lakhely esetében a kutatószobai állomány is tartalmazza a REGIO változót (hét darab NUTS 2-es régió). Miközben az oktatási célú fájl esetében ez az egyetlen földrajzi bontás a lakóhelyre, a kutatószobában településen belüli címek szintjére is le van bontva ugyanez. A NEME változó mindkét állományban férfi/nő bontásban jelenik meg, míg a többi változó esetében globális átkódolást, illetve felső kódolást alkalmaztunk.

5. táblázat

Az oktatási célú fájlban szereplő változók részletezettsége a kutatászobai állománnyal összevetve

Változó*	Kutatászobai állomány	Oktatási célú fájl
Lakhely (REGIO)	Településen belüli pontos cím (Településkód-számlálókörzet- címsorszám)	7 db NUTS2 szintű régió
Nem (NEME)	Férfi/nő	Férfi/nő
Kor (KORCSOPORT)	Korév	9 korcsoport
Családi állapot (CSPOT)	7 kategória	4 kategória
Élve született gyermekeinek száma (EGYERUJ)	0,1,2,...	0/1/2/3/4/5/6-
Iskolába járás (TANUL)	Nem jár iskolába + 11 kategóriába besorolva, ha igen	Igen/nem
Iskolai végzettség (ISKOLA)	16 fokozatú skála	5 fokozatú skála
Gazdasági aktivitás (GAKT)	19 fokozatú skála	4 fokozatú skála
Jelenlegi foglalkozása, munkaköre (FOGLKOD)	2 jegyű FEOR '08	6 fokozatú skála
Lakásépítés éve (EPEV)	9 időintervallum	7 időintervallum
Lakás alapterülete (m ²) (LAT)	A lakás alapterülete m ² -ben	6 kategória
Családösszetétel (CSOSSZ)	13 kategória (ebből kettő nem családösszetétel)	4 kategória (ebből egy nem család- összetétel)
Lakásban lakó személyek száma (LLAKO)	0,1,2,...	0/1/2/3/4/5/6-

* A megnevezések az oktatási célú fájlban feleltethetők meg.

Hogy meggyőződjünk arról, hogy a kapott oktatási célú fájl felfedési kockázata elegendően alacsony-e ahhoz, hogy publikálható legyen, a 2. fejezet 2. ábrájának lépéseit követtük:

1. *Ellenőrzés.* Az állomány nem tartalmaz-e közvetlen azonosítókat.
2. *Felfedési forgatókönyvek.*

a) Van-e annak reális kockázata, hogy az oktatási célú fájl egy másik mikroadattal összekapcsolják, a benne szereplő személyeket beazonosítsák és róluk új információkat fedjenek fel? (Külső archívum, nyilvános adatok.)

Feltételezhető, hogy a népszámlálásból egyéb mikroadat nem áll rendelkezésre nyilvánosan,³³ mivel az adatok közzététele aggregált

³³ Kivéve a tesztfájlokat, de az abban szereplő rekordok fiktívek.

táblázatos adat formájában történik. Egyéb külső, személyi szintű állományokkal való összekapcsolás reális forgatókönyvként nem merült fel.

b) Van-e annak reális kockázata, hogy valaki beazonosítja egy közeli ismerősét (kíváncsi szomszéd)?

Ha egy felhasználó olyan rekordot talál az állományban, melynek – bizonyos jellemzői alapján – ő maga vagy egy általa ismert személy megfelel, akkor elhanyagolható a valószínűsége annak, hogy valóban önmagát vagy az általa ismert személyt azonosította be. Ennek oka az, hogy a beazonosítást lehetővé tevő kulcsokra legalább három előfordulás található a sokaságban. Másképp fogalmazva: a következő pontban meghatározott kulcsokra az állomány $k=3$ -ra teljesíti a k anonimitás kritériumát. Ezért egyáltalán nem biztos, hogy a beazonosítani vélt személy az, akire a felhasználó gondol, és nem egy másik, hasonló tulajdonságokkal rendelkező személy.

3. *Kulcsváltozók meghatározása.* Az állományban szereplő változók alacsony számossága miatt az összes lehetséges egy- és kétváltozós gyakoriságot megvizsgáltuk a mintában.

4. *Kulcsok képzése.* Ez azt jelenti, hogy 13 darab egyváltozós kombinációt, illetve 78 darab kétváltozós kombinációt kell áttekinteni. Kétváltozós kulcs például a REGIO-NEME, melynek összesen $7 \times 2 = 14$ darab elméleti kombinációja létezik. Ebben az esetben kulcskombináció például az Észak-Magyarországon élő férfi. Megvizsgáltunk néhány háromváltozós kulcsból képezhető kombinációt is, de a nagy számosság (286 kombináció) miatt kizárólag azokat,³⁴ melyek közvetett azonosítókat tartalmaznak, mert ezek alapján legnagyobb a valószínűsége a felfedésnek.

5. *Felfedési kockázat.* A következő lépés az, hogy meghatározzuk az egyes kulcskombinációk előfordulásait. Miután képeztük az összes lehetséges kulcsot, ki kell választani azokat a kombinációkat, melyek valóban elő is fordulnak az oktatási célú fájlban. Ezekre a kulcskombinációkra keresztábrákat készítettünk a sokaság (teljes népszámlálási állomány) felhasználásával. Ezek alapján arra a következtetésre jutottunk, hogy bizonyos esetekben (például EGYERUJ = 5,6 vagy CSPOT = 3) vannak alacsony előfordulások a mintában és a sokaságban is. Az alacsony mintabeli előfordulások azonban nem köszönnek vissza a sokaságban, míg az alacsony sokasági előfordulások kulcskombinációi nem kerültek bele a mintába.

³⁴ REGIO, NEME, KORCSOP, CSPOT, EGYERUJ, ISKOLA, GAKT, LAT.

6. *Információvesztés mérése.* Mivel az oktatási célú fájl korlátozott mértékben ugyan, de egyszerűbb elemzések elvégzésére, statisztikai módszerek begyakorlására is alkalmas, szükséges megadni néhány minőségindikátort. Ezek azt mérik, hogy az oktatási célú fájlból kapott eredmények milyen mértékben térnek el attól, mintha a sokaságból szármítottuk volna ki ugyanezeket az értékeket. Az eltérések nyilvánvalóan a mintavételi hibának tudhatók be. Illusztrációképpen kiszámítottuk az egyváltozós megoszlásokat az oktatási célú fájlra, és összevetettük a teljes népszámlálásban előforduló megoszlásokkal. Bár az oktatási célú fájlból számított megoszlások a legtöbb esetben elég jól visszaadták a valóságban tapasztalt arányokat, komolyabb kutatásra semmiképp nem ajánlható az oktatási célú fájlok használata. A kutatószobában megtalálható állományok jóval nagyobb mintát (10%) képviselnek, sokkal több és részletesebb változókészlettel rendelkeznek. Tudományos célra, publikációkhoz mindenképpen ezek használatát javasoljuk.

Összegezve azt lehet mondani, hogy az alacsony kiválasztási arány (körülbelül 1%), a viszonylag kevés (13) és erősen összevont kategóriákkal rendelkező változó és a NUTS 2 szintű területi bontás elegendő mértékben csökkentette a mintába került személyek beazonosításának és valamely jellemzőjének/jellemzőinek a felfedési kockázatát, így az állomány további beavatkozás nélkül publikálható.

3.3. Az IPUMS számára átadott kutatási célú fájl

A KSH 2015 elején kutatásra alkalmas nyilvános mikroadatfájlokat adott át a Minnesotai Egyetem számára az IPUMS-projekt keretében. Az állományok előzetes regisztráció³⁵ után kutatók számára ingyenesen hozzáférhetők lesznek 2016 folyamán. A regisztráció során beazonosítják a kutató személyét, megvizsgálják a kutató intézményi hátterét és elbírálják a kutatási tervet.

A mintavétel a sokasági lakásállományból indult ki. 5 százalékos egyszerű véletlen mintát vettünk, visszatevés nélkül. A mintavétel során rétegzés nem volt. A mintába került nem üres lakásokhoz hozzárendeltük a benne lakó személyeket, továbbá a személyeket kiegészítettük a hajléktalanokból és az intézeti lakókból vett 5 százalékos visszatevés nélküli egyszerű véletlen mintával. Az így létrejött személyi állomány körülbelül 500 ezer rekordot tartalmaz, míg a lakásállomány nagyságrendileg 220 ezret. A két állomány egy lakásazonosító alapján összekapcsolható. A személyi állomány 36 változót tartalmaz, a lakásállomány pedig 18-at.

³⁵ https://international.ipums.org/international-action/show_register?id=0

Egy adott változót vagy globálisan átkódoltunk vagy változatlan formában hagytuk. A kérdés az, hogy az átkódolások vajon elegendő mértékben csökkentették-e a felfedési kockázatot, vagy még további beavatkozásra is szükség van az állományok anonimitása érdekében. A végrehajtott lépések nagyon hasonlóak az oktatási fájlnál leírtakhoz:

1. *Ellenőrzés.* Meggyőződünk arról, hogy az állomány nem tartalmaz közvetlen azonosítókat.

2. *Felfedési forgatókönyvek.*

a) Van-e annak reális kockázata, hogy a kutatási célú nyilvános mikroadatfájl egy másik mikroadattal összekapcsolják, a benne szereplő személyeket beazonosítsák és róluk új információkat fedjenek fel? (Külső archívum, nyilvános adatok.)

Mivel a 2011-es Népszámlálásból nem áll rendelkezésre a Minnesotai Egyetemnek átadott állományhoz hasonló változókészlettel és megközelítőleg hasonló kiválasztási aránnyal rendelkező népszámlálásból származó vagy egyéb személyi szintű állomány, ezért a külső állománnyal való összekapcsolás lehetőségét elvetettük.

b) Van-e annak reális kockázata, hogy valaki beazonosítja egy közeli ismerősét (kíváncsi szomszéd)?

A kulcsváltozók kiválasztása során ezt a forgatókönyvet vettük figyelembe: a támadó az ismeretségbe tartozó egyedekről birtokában levő információkat (például kor, nem, gyermekek száma, foglalkozás stb.) veti össze az állományban található rekordokkal.

3. *Kulcsváltozók meghatározása.* A figyelembe vett felfedési forgatókönyv alapján a következő kulcsváltozókat határoztuk meg:

a) személyi állomány: családi állás, itt tartózkodás jogcíme, nem, korév, állampolgárság, családi állapot, családfőhöz való viszony, háztartásfőhöz való viszony, gyermekek száma, az iskolarendszerben elévített legmagasabb szint, gazdasági aktivitás, foglalkozása, munkáltatójának/vállalkozásának tevékenysége, lakhelyének régiója, születéskori lakhely települése;

b) lakásállomány: az épület típusa, a lakás építési éve, a lakóegység használata, tulajdon típusa, a lakás szobaszám szerinti nagysága, fürdőszobák száma, vízöblítéses WC-k száma, a lakás alapterülete.

4. *Kulcsok képzése.* Ezekből a kulcsváltozókból legfeljebb négy változóból álló kulcsokat képeztünk. Ezekre a kulcsokra ki kell számítani a felvehető értékek összes lehetséges kombinációjának a gyakoriságát az 5 százalékos mintában. Ez azt jelenti, hogy 30 darab egyváltozós, 435 darab kétváltozós, 4 060 darab háromváltozós és 27 405 darab négyváltozós kombináció van.

5. *Felfedési kockázat.* Ezek után a kulcskombinációk előfordulásait határoztuk meg. Látható, hogy a három- és négyváltozós kombinációk száma már olyan magas, amit egyenként áttekinteni nagyon időigényes feladat lenne. Ezért szükség van a vizsgálat alá vont esetek számának szűkítésére:

a) Az egy- és kétváltozós kombinációk keresztábráit egyenként áttekintettük. Már ezek alapján is elég jó közelítéssel ki lehet szűrni azokat a változókat, ahol alacsony előfordulások tapasztalhatók.

b) A hármas és négyes kulcsok számát úgy szűkítettük le, hogy csak azokat a kombinációkat vizsgáltuk meg, melyek a legnagyobb mértékben alkalmasak a beazonosításra és a felfedésre.

Ha megtaláltuk a minta alapján kockázatosnak ítélt kulcsokat, akkor ezekre a kulcsokra meg kell vizsgálni a sokasági előfordulásokat is.³⁶ Végül is csak azokon a változókon kell valamilyen felfedés elleni védelmi beavatkozást végrehajtani, amelyekre a sokaságban is alacsony előfordulásokat tapasztalunk. Ha egy változóra vagy kulcsra kizárólag a mintában tapasztalunk alacsony gyakoriságokat, de a sokaságban nem, akkor ezeket a változókat nem szükséges felfedés elleni védelmi okokból módosítani. Meg kell jegyezni, hogy azokat a kulcskombinációkhoz tartozó keresztábrákat, amelyek csak személyi jellemzőket tartalmaznak a személyi állományból, a csak lakásváltozókból állókat a lakásállományból, míg a vegyes jellemzőkkel rendelkezőket a személyi állományból³⁷ készítettük el.

Például a *korév* = 50, a *lakás alapterülete* = 70 kulcs esetében a kérdéses gyakoriság: hány 50 éves, 70 m²-es lakásban lakó személy található az 5 százalékos mintában?

6. *A nyers mikroadatfájlon végrehajtott módosítások.* A keresztábrák áttekintése után ezeket a következtetéseket vontuk le:

– A legtöbb vizsgált kombinációban a *korév* alacsony előfordulásokkal szerepelt a *KEV* > 90 esetében.

– A legtöbb vizsgált kombinációban az *élve született gyermekek száma* alacsony előfordulásokkal szerepelt az *EGYERUJ* > 10 esetében.

– A legtöbb vizsgált kombinációban a *fürdőszobák száma* és a *víz-öblítéses WC-k száma* alacsony előfordulásokkal szerepelt a *FURDOUJ* > 2 és a *WCUJ* > 2 esetében.

– A *lakás alapterülete* a *LAT* < 12 és a *LAT* > 300 esetében a legtöbb vizsgált kombinációban alacsony előfordulásokkal szerepelt.

³⁶ Pontosabban a kulcsok lehetséges kombinációinak (nem az összes elméletben lehetséges kombinációnak, csak azoknak, amelyek a mintában is előfordulnak) a gyakoriságát kell megvizsgálni.

³⁷ A személyi állományhoz hozzákapsoltuk a lakásállományt, így egy rekord szinten személyeket tartalmazó, de lakásjellelmzőkkel kiegészített állományt kaptunk.

– A *foglalkozás ISCO-08*³⁸ szerinti *foglalkozási csoportja* a FOGLKOD = 03 (Fegyveres erők egyéb rangú állománya) és a FOGLKOD = 95 (Utcai és hasonló árusok, szolgáltatók) esetében több kombinációban alacsony előfordulásokat tapasztaltunk.

– A *munkáltató vagy a vállalkozás NACE Rev. 2*³⁹ szerinti *ágazata* a MUNKKOD = 03 (Halászat, halgazdálkodás), a MUNKKOD = 06 (Kőolaj-, földgáztermelés), a MUNKKOD = 07 (Fémtartalmú érc bányászata) és a MUNKKOD = 98 (Háztartás termék-előállítás, szolgáltatása saját fogyasztásra) esetében több kombinációban alacsony előfordulásokat tapasztaltunk.

Az ún. „problémás” változók tehát további beavatkozást igényeltek. A rajtuk végrehajtott módosításokat tartalmazza a 6. táblázat.

6. táblázat

További felfedés elleni védelmet igénylő változók

Problémás változó	Felfedés elleni védelem érdekében végrehajtott módosítások
Korév (KEV)	KEV > 90 értékekre felsőkódolás
Élve született gyermekek száma (EGYERUJ)	EGYERUJ > 10 értékekre felsőkódolás
Fürdőszobák száma (FURDOUJ)	FURDOUJ > 3 értékekre felsőkódolás
Vízöblítéses WC-k száma (WCUJ)	WCUJ > 3 értékekre felsőkódolás
Lakás alapterülete (LAT)	LAT < 12 és LAT > 300 értékekre alsó-, illetve felsőkódolás
Foglalkozás ISCO-08 szerinti foglalkozási csoportja (FOGLKOD)	FOGLKOD = 02 ⁴⁰ tartalmazza a 03-as kategóriát is FOGLKOD = 52 ⁴¹ tartalmazza a 95-ös kategóriát is
Munkáltató/vállalkozás NACE Rev. 2 szerinti ágazata (MUNKKOD)	MUNKKOD = 02 ⁴² tartalmazza a 03-as kategóriát is MUNKKOD = 05 ⁴³ tartalmazza a 06-os és a 07-es kategóriát is MUNKKOD = 97 ⁴⁴ tartalmazza a 98-as kategóriát is

A 6. táblázatból jól látható, hogy az azonosító jellegű változók alacsony előfordulásait alsó- és felsőkódolással, illetve kategóriák összevonásával sikerült kiküszöböl-

³⁸ ISCO (International Standard Classification of Occupations): Foglalkozások Egységes Nemzetközi Osztályozási Rendszere.

³⁹ NACE Rev. 2 (Statistical Classification of Economic Activities in the European Community): az Európai Közösség Gazdasági Tevékenységek Statisztikai Osztályozása.

⁴⁰ Fegyveres erők tiszthelyettesei.

⁴¹ Értékesítési foglalkozások.

⁴² Erdőgazdálkodás.

⁴³ Szénbányászat.

⁴⁴ Háztartási alkalmazottat foglalkoztató magánháztartás.

ni. Az így létrejött személyi- és lakásminta véleményünk szerint már nem tartalmaz reális felfedési kockázatot, figyelembe véve a felfedési forgatókönyvet.

Az információvesztésért ebben az esetben nem mértük fel, mivel a megrendelő (a Minnesotai Egyetem) ezt nem igényelte.

4. Összegzés

Nyilvános mikroadatfájlokat a világ fejlett részein már évtizedek óta publikálnak, de ahogy az 1. fejezetben bemutattuk, egységes definíció a mai napig nem létezik rá. Ebből következően a nyilvános mikroadatfájlok előállítására is számos lehetőség áll rendelkezésre. Ebben a cikkben nem vállalkoztunk, nem is vállalkozhattunk arra, hogy a teljesség igényével mutassuk be a nyilvános mikroadatfájlok előállításának módszertani kelléktárát. Csupán a módszertanilag egyszerűbb, ún. hagyományos módszereket⁴⁵ foglaltuk össze. A módszerek technikai részleteiről az érdeklődő Olvasó tájékozódhat a bőségesen rendelkezésre álló szakirodalomban. E helyett inkább a „nagy egészre” a nyilvános mikroadatok előállításának folyamatára koncentráltunk. Hangsúlyozzuk, hogy a 2. fejezetben felvázolt folyamatséma csupán egy lehetséges, de nem az egyetlen módja a nyilvános mikroadatfájlok előállításának. A cikk vezérfonalként szolgált, mely segít eligazodni a módszerek és lehetőségek tárházában.

Bízunk abban, hogy a tanulmány – elsősorban a 3. fejezet – hasznos lehet azoknak a felhasználóknak, akik már használták, vagy ezután kívánják megismerni a KSH által publikált nyilvános mikroadatfájlokat. Célunk az volt, hogy a KSH honlapján megtalálható információknál egy kicsit bővebben bemutassuk ezeknek az állományoknak az előállítási folyamatát, illetve az előállítás során felmerülő adatvédelmi dilemmákat.

A cikk azok számára is hasznos lehet, akik most ismerkednek a nyilvános mikroadatfájlokkal, és szeretnének megismerkedni az alapfogalmakkal, illetve rövid bevezetést kapni arról, hogy melyek a legalapvetőbb, leggyakrabban használt módszerek a nyilvános mikroadatok előállítása során.

Végül, de nem utolsósorban a cikk megírását az is motiválta, hogy összegezzük a KSH-ban még viszonylag gyerekcipőben járó nyilvános mikroadatokkal kapcsolatos módszertani tapasztalatainkat. Bízunk benne, hogy a közeljövőben a felhasználói igényeknek megfelelően további nyilvános mikroadatfájlok publikálására is sor kerülhet, melyek során továbbfejleszhetjük az itt bemutatott módszertani apparátust is.

⁴⁵ Nem beszéltünk például az utóbbi években egyre divatosabbá váló szintetikus adatok módszertanáról, például *Alfons et al.* [2011].

Irodalom

- ALFONS, A. – KRAFT, S. – TEMPL, M. – FILZMOSER, P. [2011]: Simulation of Close-to-Reality Population Data for Household Surveys with Application to EU-SILC. *Statistical Methods & Applications*. Vol. 20. Issue 3. pp. 383–407.
- AZ EURÓPAI UNIÓ HIVATALOS LAPJA [2009]: Az Európai Parlament és a Tanács 223/2009/EK rendelete (2009. március 11.) az európai statisztikákról és a titoktartási kötelezettség hatálya alá tartozó statisztikai adatoknak az Európai Közösségek Statisztikai Hivatala részére történő továbbításáról 1101/2008/EK, Euratom európai parlamenti és tanácsi rendelet, a közösségi statisztikákról szóló 322/97/EK tanácsi rendelet és az Európai Közösségek statisztikai programbizottságának létrehozásáról szóló 89/382/EGK, Euratom tanácsi határozat hatályon kívül helyezéséről. 2009. 3. 31. L 87/164.
- BARTUS T. [2013]: Adatcserével anonimizált mikroadatok használhatósága – Egy szimulációs vizsgálat tanulságai. *Statisztikai Szemle*. 91. évf. 5. sz. 465–497. old.
- DRECHSLER, J. – REITER, J. P. [2009]: Disclosure Risk and Data Utility for Partially Synthetic Data: An Empirical Study Using the German IAB Establishment Survey. *Journal of Official Statistics*. Vol. 25. No. 4. pp. 589–603.
- DRECHSLER, J. [2011]: *Synthetic Datasets for Statistical Disclosure Control*. Springer. New York.
- EXECUTIVE OFFICE OF THE PRESIDENT OF THE UNITED STATES [2014]: Methods for Public-Use Microdata Files. In: *Report on Statistical Disclosure Limitation Methodology*. Working Paper 22. pp. 61–72. https://www.ciser.cornell.edu/NYCRDC/helpful_links/WP-22-OMB-totalreport.pdf
- HUNDEPOOL, A. – DOMINGO-FERRER, J. – FRANCONI, L. – GIESSING, S. – NORDHOLT, E. S. – SPICER, K. – DE WOLF, P.-P. [2012]: *Statistical Disclosure Control*. John Wiley & Sons, Ltd. London.
- HUNDEPOOL, A. – DE WOLF, P.-P. – BAKKER, J. – REEDIJK, A. – FRANCONI, L. – POLETTINI, S. – CAPOBIANCHI, A. – DOMINGO-FERRER, J. [2014]: *μ -Argus 5.1 User's Manual*. Statistics Netherland. Amsterdam.
- KSH (KÖZPONTI STATISZTIKAI HIVATAL) [2013]: *A Központi Statisztikai Hivatal elnökének 27/2013. KSH utasítása a Központi Statisztikai Hivatal adatvédelmi szabályzatáról*. Budapest. https://www.ksh.hu/docs/szolgáltatások/adatigenyles/ksh_adatvedelmi_szabalyzat.pdf
- LOWTHIAN, P. [2012]: *Guidelines for Anonymisation of Social Survey Microdata*. Eurostat. Luxembourg.
- OECD (ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT) [2014]: *OECD Expert Group for International Collaboration on Microdata Access – Final Report*. <http://www.oecd.org/std/microdata-access-final-report-OECD-2014.pdf>
- TEMPL, M. – MEINDL, B. – KOWARIK, A. [2014]: *Introduction to Statistical Disclosure Control (SDC)*. International Household Survey Network. Vienna. https://cran.r-project.org/web/packages/sdcMicro/vignettes/sdc_guidelines.pdf
- WILLENBORG, L. – DE WAAL, T. [2001]: *Elements of Statistical Disclosure Control*. Springer-Verlag. New York.

Summary

In the first part of this study, the main types and definitions of public use files are summarized. Beyond that, key issues and the most important steps of generating such files are outlined briefly at an introductory level. The second part of the paper reflects a more practical point of view, since the elaboration of the three types of public use files (research, teaching and test files) is demonstrated. These files are derived from the 2011 Population and Housing Census of Hungary conducted by the Hungarian Central Statistical Office. If the public use files described in the study become popular among users, the Hungarian Central Statistical Office may publish more such files in the future.

Az EU-SILC- és az EU-LFS-adatok nyilvános mikroadatfájljai*

de Wolf, Peter-Paul

a Holland Statisztikai Hivatal
statisztikusa

E-mail: pp.dewolf@cbs.nl

Részben az Eurostat által finanszírozva, hét tagállam dolgozik egy olyan harmonizált szakmai javaslaton, amely segítségével az EU-SILC¹ és az EU-LFS² adatokból nyilvános mikroadatfájlok állíthatók elő. A végleges eredmények 2015. decemberére várhatók. A cikk – melyet a UNECE³/Eurostat adatvédelemmel foglalkozó munkacsoportjának Helsinkiben, 2015. október 5–7-én tartott ülésén mutattak be – a projekt aktuális állapotát írja le és közli az (előzetes) eredményeket.

TÁRGYSZÓ:

Folyamatok fejlesztése.

EU-SILC.

EU-LFS.

* Ez a tanulmány az Európai Unió hét tagállamának projektjében résztvevők közreműködésén alapul: *Maxime Bergeat, Matthias Templ, Lydia Spies, Annu Cabrera, Kristóf Péter, Andreja Smukavec, Aleksandra Bujnowska, Peter-Paul de Wolf*. Angol nyelvű változata mellékletként elérhető a www.ksh.hu/statszemle internetes oldalon. A magyar nyelvű fordítást a *Központi Statisztikai Hivatal Módszertani főosztálya* készítette.

¹ EU-SILC (European Union statistics on income and living conditions): az Európai Unió háztartási költségvetési és életkörülmény-felvétele.

² LFS (labour force survey): munkaerő-felmérés.

³ UNECE (United Nations Economic Commission for Europe): az Egyesült Nemzetek Szervezetének Európai Gazdasági Bizottsága.

Az Eurostat számos mikroadat-állományhoz enged hozzáférést külső kutatók számára. Az állományok elérhetők kutatószobában használt közvetlen azonosítótól mentes mikroadat-állományként vagy anonimizált mikroadatként. Mind a kutatószobai állományokhoz mind az anonimizált mikroadatokhoz csak akkreditált kutatók férhetnek hozzá: az adatok bizalmasak, ezért korlátozott az elérhetőségük. A kutatókra az akkreditáción kívül még jogi előírások is vonatkoznak (ezeknek a megszegése büntetést von maga után), ennek következtében a mikroadatfájlok nem teljesen anonimizáltak. Az egyedi adatszolgáltatók érzékeny adatai nyilvánosságra kerülésének valószínűségét statisztikai felfedés elleni védelmi módszerekkel szükséges csökkenteni, elsősorban a globális átkódolás és a lokális elnyomás technikájának alkalmazásával.

Az anonimizált mikroadat-állományokhoz úgy lehet hozzáférni, hogy első lépésként azt a szervezetet, ahová a kutató tartozik, el kell fogadtatni hiteles kutatási intézményként az Eurostatnál, ahová minden egyes új projekthez kutatási tervet kell beadni. Mivel az akkreditációs folyamat elég sok időt vehet igénybe (néha akár 10 hetet is), ezért a kutató és a kutatóintézet számára is előnyös, ha az adatfájl tartalmáról előzetes ismeretekkel rendelkezik. Ez alapján körültekintőbb döntést lehet hozni arról, hogy valóban érdemes-e ennyi időt fektetni az akkreditációs eljárásba.

Tehát hasznos lenne, ha rendelkezésre állna egy mikroadatfájl, ami a kutató számára már az akkreditációs folyamatot megelőzően elküldhető és minden kontrol nélkül használható lenne. Ezt az állományt nyilvános mikroadatfájlnak kell tekinteni és annak megfelelően kezelni (például tükrözze az anonimizált mikroadat-állomány szerkezetét). Így a kutató tisztelni tudja az előzetesen megírt programkódjait, és általános képet kaphat az állomány tartalmáról.

Az Eurostat 2015 januárjában egy olyan projektet⁴ indított, melynek célja nyilvános mikroadatfájlok készítése azokhoz az állományokhoz, melyek már elérhetők anonimizált mikroadatként. Ez a lehetőség nagymértékben megkönnyíti az akkreditációra váró kutatók munkáját. A nyilvános mikroadatfájlok másodlagos felhasználási céljaként megfogalmazódott, hogy ezeket a fájlokat hozzáférhetővé kellene tenni statisztikai képzések számára is. A projekt végére a kész nyilvános mikroadatfájlokat olyan dokumentációval együtt kell elérhetővé tenni, amely tartalmazza létrehozásuk harmonizált módszertanát.

Felmérés készült arról, hogy az egyes mikroadat-állományokhoz⁵ hány esetben érkezett hozzáférési igény az Eurostathoz. Az EU-SILC- és az EU-LFS-adatok bizo-

⁴ SGA 11112.2014.067-2014.765 under FPA 11112.2014.005-2014.533

⁵ Eurostat-prezentáció a második Európai Adathozzáférési Fórumon, 2015. március 24–25, Luxembourg.

nyultak a legnépszerűbbnek, ezért dolgoztunk ezekkel a projekt során. (Részletesebb információ érhető el *Bujnowska* [2015] írásában az Eurostat mikroadat-állományainak hozzáféréséről.)

A nyilvános mikroadatfájlok előállítása kétféleképpen valósulhat meg: „hagyományos” vagy (teljesen) szintetikus adatokat eredményező eljárással. Előbbi esetén csak olyan statisztikai felfedés elleni védelmi módszereket alkalmazunk, mint a globális átkódolás, lokális elnyomás és a PRAM⁶ egyszerű formája. A szintetikus megközelítés pedig egy adatgeneráló eljárás kidolgozását jelenti.

Az EU-SILC-állomány számos jövedelemváltozót tartalmaz. Néhány ország nem engedélyezi, hogy érzékeny változók is megjelenjenek nyilvános mikroadatfájlokban. Amennyiben kivesszük az összes jövedelemváltozót az EU-SILC-állományból, akkor olyan állományt kapunk eredményül, amely nem felel meg a leírt céloknak. Ezért döntöttünk úgy, hogy az EU-SILC nyilvános mikroadatfájl előállításához a teljesen szintetikus formát alkalmazzuk. Az EU-LFS-állomány esetében a hagyományos megközelítés mellett döntöttünk.

A tanulmányban leírjuk a nyilvános mikroadat-állományok előállítási folyamatát. Bemutatjuk az EU-SILC szintetikus és az EU-LFS-adatok hagyományos megközelítésszerű előállítási módját. A cikk írásának időpontjában az említett módszereket még csak néhány állományon alkalmaztuk. Ismertetjük a különböző megközelítésekhez tartozó felfedési kockázatokat. Néhány hasznossági (információvesztési) mérőszámot is definiáltunk annak illusztrálására, hogy vajon a kutatók számára mennyire jól használhatók a nyilvános mikroadatfájlok előzetes adatként, miközben az akkreditációs eljárás eredményére várnak. Végül a projekt jelenlegi helyzetéről számolunk be.

1. Az EU-SILC-állomány esetén használt megközelítés

Az EU-SILC az Eurostat által koordinált, az EU jelenlegi és leendő tagállamai, valamint néhány EFTA-ország adatait tartalmazó keresztmetszeti és longitudinális mintás felvétel. A keresztmetszeti adatok rögzített, állandó időszakokra vonatkoznak, olyan változókkal, mint a jövedelem, a szegénység, a szegregáció és az életkörülmények. A longitudinális adatok az egyéni szintű változásokat mutatják, általában négyévnvi periódust figyelnek meg.

Tudomásunk szerint, nincs olyan kész szintetikus adatgenerálási eljárás, amellyel teljesen szintetikus longitudinális állományt lehet előállítani, és ami konzisztens a

⁶ PRAM (post randomization method): utólagos randomizációs módszer. Lásd *Hundepool et al.* [2012] és *Gouweleeuw et al.* [1998] leírását a PRAM-ról.

kapcsolódó keresztmetszeti adatokkal, valamint hasonló az anonimizált mikroadat-állományhoz. Ezért úgy döntöttünk, tekintve a projekt korlátozott időkeretét, hogy az EU-SILC-állomány esetében csak a keresztmetszeti adatok nyilvános mikroadat-fájljának előállítására koncentrálunk.

1.1. A szintetikus megközelítés

A módszertanról részletesebb leírás *Alfons et al.* [2011] munkájában olvasható, alkalmazásáról – kifejezetten az EU-SILC-állományon – a projekt egy később kiadandó munkájában ad számot (megjelenése 2015. decemberre várható).

Az eljárás fő célja, hogy olyan szintetikus állományt állítsunk elő, amely „hasonlít” az Eurostat EU-SILC anonimizált mikroadat-állományához. Általánosságban elmondható, hogy egy teljesen szintetikus adatbázis előállítása (egyik változó sem „valós”) biztonságos nyilvános mikroadatot eredményez, a felfedési kockázatot a későbbiekben tárgyaljuk.

A szintetikus adatok előállításához használt modellek az eredeti nyers adatokon alapszanak, mivel olyan nyilvános mikroadatfájlt kell készíteni, amely szerkezetében hasonlít az anonimizált mikroadathoz, továbbá olyan információkat kell tartalmaznia, amelyek a lehető legközelebb esnek az eredeti mintához.

Ez azt jelenti, hogy az Eurostatnak küldött adatokat használjuk, azokat, amelyekből az Eurostat anonimizált mikroadatot állít elő. Ennek eredményeként olyan szintetikus állomány jön létre, amelynek struktúrája hasonló a nyers adatokéhoz. Így viszont alkalmaznunk kell a szintetikus adatokon azt a transzformációt, amelyet az Eurostat is elvégzett, hogy az anonimizált mikroadat-állomány struktúráját kapjuk.

Az EU-SILC alapvetően háztartásfelvétel. Ennek következményeként bizonyos háztartásstruktúra jelen van az állományban. A szintetikus megközelítést használva megtartjuk e struktúrát a nyilvános mikroadatban is.

Első lépésként létrehozunk egy szintetikus alapsokaságot, amelyhez nyers adatokon alapuló modelleket használunk. Ezután ebből a szintetikus alapsokaságból a nyers adatokéval megegyező méretű mintát veszünk. Végül ezt transzformáljuk, hogy tükrözze az anonimizált mikroadat struktúráját. Ahhoz, hogy a teljes alapsokaságot szimulálni tudjuk, a nyers adatok keresztmetszeti súlyait használjuk.

A szintetikusadat-szimuláció négy lépésből áll, ezeket a lépéseket minden (regionális) rétegre függetlenül alkalmazzuk:

1. A háztartásstruktúra felállítása.
2. Kategoriális változók szimulációja.
3. (Félig) folytonos változók szimulációja.
4. (Félig) folytonos változók különböző komponensekre osztása.

Az 1. lépés, a háztartásstruktúra felállítása, a következőt jelenti. Először a Horvitz–Thompson-becsléssel meghatározzuk az alapsokaságban minden háztartásmérethez a háztartások számát. Majd a szintetikus alapsokaságot állítjuk elő, pontosan ennyi háztartást létrehozva. A háztartásstruktúrát (kor és nem szerinti megoszlás a háztartáson belül) a nyers állomány azonos méretű háztartásainak struktúrájából választjuk ki újra-mintavételezéssel, a szintetikus alapsokaság minden l méretű háztartásához. Ezt az illogikus háztartásstruktúrák elkerülése érdekében tesszük. Egyrészt ez azt is jelenti, hogy amennyiben egy bizonyos háztartásméret egy adott rétegben egyedi a mintában, a struktúrája (nem és kor eloszlása) mindig átadódik az adott rétegben az adott méretű minden egyes szimulált háztartásához. Másrészt viszont lehetséges, hogy ebben a rétegben és háztartásméretben több háztartást is szimulálunk.

A 2. lépésben a kategorikus változókat szimuláljuk a háztartások mindegyikére a szintetikus alapsokaságban. Ez többlépcsős folyamat (minden változót az előzőekben szimulált változóktól függően szimulálunk). A nyers adatokat véve multinomiális logisztikus regressziós modellt illesztünk a kategorikus változókra az előzőekben szimulált változókkal, mint prediktorokat használva. A multinomiális eloszlásból ezután kapunk egy értéket a becslt (feltételes) valószínűségekkel. A változók, amelyeket e módon szimulálunk (ebben a sorrendben): az önbevalláson alapuló gazdasági státus, az állampolgárság, a családi állapot, az iskolázottság, a foglalkozás (egy számjegy, a második számjegyet véletlenszerűen választjuk az előzőre, mint feltételre nézve), a NACE-kód (egy számjegyen).

A 3. lépésben a (fél) folytonos változókat két lépésben szimuláljuk. Először a változót leképezzük a változó kategóriarendszerébe (például jövedelemosztályok), majd ugyanazt a megközelítést alkalmazzuk, mint az igazi kategoriális változók esetén. Végül egy véletlen értéket választunk az adott jövedelemkategórián belül, hogy folytonos változót kapjunk.

A 4. lépésben a jövedelemváltozó esetén a teljes jövedelmet felosztjuk különböző komponensekre. A különböző komponensek eloszlásának meghatározása donor imputálással történik, ahol minden szimulált rekordhoz az adott rétegből választunk egy rekordot, és a szimulált rekord a donorrekord arányait kapja. Ez függetlenül történik a háztartásra és a személyekre vonatkozó jövedelemkomponenseknél. Vegyük észre, hogy egy donorrekord jövedelemváltozójának csak az arányszámait használjuk.

A nyilvános mikroadatfájl végül a szintetikus alapsokaságból választjuk. Ezt rétegzett véletlen, visszatevéses mintavétellel végezzük, ahol a rétegváltozó a régió és a háztartás mintavételi egység. Régióként ugyanannyi háztartást választunk a szintetikus alapsokaságból visszatevéssel, mint amennyi az anonimizált mikroadatban van (de lehetőleg különböző háztartásméretekkkel).

1.2. Néhány gyakorlati probléma

Az előbbiekben leírt szintetikusadat-előállítási folyamat gyakorlati alkalmazása során néhány probléma merült fel. Kiderült, hogy van olyan változó, melynek eloszlása meglehetősen ritka bizonyos rétegekben, mely befolyásolja a modell becsléseit. Emiatt esetenként úgy döntöttünk, hogy egy lépésben, egyszerre végzünk becslést az egész országra.

Egy másik probléma a populáció méretéhez kapcsolódik. Az az általános ötlet, mely szerint minden egyes ország esetén generáljuk a teljes populációt, nem lehetséges (figyelembe véve az ország teljes népességét és a rendelkezésre álló számítógépkapacitást). Elegendő a tényleges népességszámnál kisebb méretű populációt generálni, mely lényegesen nagyobb, mint a SILC mintanagysága. Ezt természetesen figyelembe kell venni a súlyképzésnél.

Tekintettel arra, hogy az EU-SILC számos változót tartalmaz, a bemutatott szimulációs eljárás gyakorlati alkalmazása közben kiderült, hogy lehetetlen az összes változót az előbbiekben leírt módon szimulálni: a már generált, nagyszámú változótól való függés jelentősen megnöveli a számítási időt. Ezért két csoportra osztottuk a változókat: az első csoport esetén a generálásánál figyelembe vettük az összes többi, már szimulált változót, míg a másik csoportnál a függőség csak egyetlen változóra vonatkozott (öt osztályra bontott bevétel). A második csoport esetén a szimulációhoz használt eloszlást ezen változó (az öt osztályra bontott bevétel) anonimizált mikroadatként súlyozott eloszlásából becsültük. Ehhez először a bevételváltozót szimuláltuk.

E megközelítés végrehajtásához R szkripteket fejlesztettünk ki, ezek magukban foglalják az adatok előzetes feldolgozását, a szintetikus populáció létrehozását, a minta kiválasztását és olyan alakra hozását, mely összhangban van az anonimizált mikroadat szerkezetével. A szintetikus populáció megalkotásához a simPop nevű R csomagot használtuk fel.

2. Az EU-LFS-állomány esetén alkalmazott megközelítés

Az EU-LFS az EU által koordinált, jelenlegi és leendő tagállamai, valamint néhány EFTA⁷-ország adatait tartalmazó, keresztmetszeti és longitudinális háztartási mintavételen alapuló felvétel. Az adatok a munkaerőpiac aktív résztvevőiről, és azon kívüli személyekről tartalmaz megfigyeléseket. Az adatok negyedéves, és éves bon-

⁷ EFTA (European Free Trade Association): Európai Szabadkereskedelmi Társulás.

tásban is rendelkezésre állnak. Döntésünk alapján a negyedéves adatokkal kezdtük, majd az éves adatokat tartalmazó nyilvános mikroadatot ezekből a negyedéves állományokból állítottuk össze. Azonban néhány, az éves adatbázisban levő változó nem található meg a negyedéves adatok között, így ezen változók esetén kicsit másképpen jártunk el.

2.1. A hagyományos megközelítés

Alapvetően három módszert alkalmaztunk:

1. Változókat távolítottunk el (mely értékeit globálisan „hiányzóra” állítottuk).
2. Globálisan átkódoltunk.
3. a) Lokálisan elnyomtunk a k anonimitás alapján az azonosító változók egy speciális részhalmazán, majd a fennmaradó részén pedig a PRAM-technikát alkalmaztunk.
b) Lokálisan elnyomtunk az azonosító változók összes m dimenziós kombinációja alapján.

Az első módszert (változók eltávolítása) természetesen úgy alkalmaztuk, hogy az előállt nyilvános mikroadat szerkezete összhangban maradt az anonimizált mikroadatével, azaz egy változó eltávolítása alatt azt értjük, hogy minden egyes értéket „hiányzóra” állítjuk.

Első lépésként a háztartások előállítását lehetővé tevő változókat töröltük. Ez nemcsak kötelező jellegű néhány tagállamban, hanem ahogy kiderült, a kutatók fő érdeklődési köre nem a háztartásokra, hanem a személyekre vonatkozó információkat érinti. A regionális változókat országos szintűre kódoltuk át, így lényegében azokat is töröltük. Néhány egyéb változót szintén globálisan „hiányzóra” (lényegében eltávolítottuk) kódoltunk, mivel azok összetett módon kapcsolatban állnak olyan változókkal, melyek átkódolására (például a NACE egy korábbi változatára) később kerül sor.

Miután eltávolítottunk néhány változót, 12 azonosító változót találtunk: urbanizáció foka, nem, életkor, állampolgárság, foglalkozási csoport (ISCO-08)⁸, lakóhelyen történő tartózkodás években kifejezve, legmagasabb iskolai végzettség, születési ország, gazdasági tevékenység (Nace Rev. 2)⁹, foglalkoztatási forma, munkavégzés

⁸ ISCO (International Standard Classification of Occupations): Foglalkozások Nemzetközi Osztályozási Rendszere.

⁹ NACE Rev. 2 (Statistical Classification of Economic Activities in the European Community): az Európai Közösség Gazdasági Tevékenységek Statisztikai Osztályozása.

helye ország szerint és az ILO¹⁰ szerinti munkaerő-piaci státus. További információkat ezen változókról az Eurostat [2014] tud szolgáltatni az anonimizált mikro-adatoknál.

Globális átkódolásra számos változónál szükség volt, már az anonimizált mikrodátum esetén is, itt azonban még szélesebb kategóriákat alkalmaztunk, és a következő csoportosítást alakítottuk ki:

- életkort: 0–14, 15–24, 25–39, 40–54, 55–74 éves és 75 év feletti;
- állampolgárság: saját ország, EU28 és nincs/egyéb válasz;
- születési ország: saját ország, EU28 és nincs/egyéb válasz;
- a foglalkozási csoport: egyjegyű ISCO kód;
- adott tagállamban leélt évek száma: 0, 1–9 év, 10 év felett;
- iskolázottság szintje: alap-, közép-, felső fok;
- foglalkoztatási forma: az alkalmazottat és a segítő családtagot egy csoportba soroltuk;
 - a munkavégzés helye: saját ország, EU28 és nincs/egyéb válasz;
 - az urbanizáció foka: a sűrűn lakott és a közepesen lakott körzetek egy csoportba kerültek;
- NACE (Rev. 2): A, B–E, F, G, H–S, U és T kategória.

Az eredeti ötlet az volt, hogy a globális átkódolás elvégzése után a felfedési kockázat mértékét ellenőrizzük, és a továbbra is kockázatos eseteket lokálisan elnyomjuk. A kockázat mérésére leggyakrabban két módszert alkalmaznak: a k anonimitást a kulcsváltozók egy limitált részhalmazán, valamint vizsgálják a kulcsváltozók összes m dimenziós kombinációját (nevezzük ezt a későbbiekben teljes m dimenziós megközelítésnek). Mi a munkánk során $k = 5$ paraméter mellett használtuk a k anonimitást a következő kulcsváltozókra: urbanizáció foka, nem, életkor, állampolgárság, foglalkozási csoport, lakóhelyen történő tartózkodás éveiben kifejezve, valamint a legmagasabb iskolai végzettség. A többi azonosító változó esetén a PRAM-technikát választottuk 80 százalékos változatlanul maradási valószínűséggel, a fennmaradó valószínűségeket pedig egyenletes osztottuk el az adott változó kategóriái között. A teljes m dimenziós megközelítés vizsgálata esetén paraméterértéknek $m = 4$ -et, küszöbértékként pedig 10-et alkalmaztunk minden egyes dimenzió esetén.

2.2. Néhány gyakorlatban felmerülő probléma

Néhány változó más változókról adhat információkat, ezért ezek különleges bánásmódot igényelnek. Például:

¹⁰ ILO (International Labour Organization): Nemzetközi Munkaügyi Szervezet.

– Állampolgárság: az állampolgárság változó elnyomása után a lakóhelyen történő tartózkodás éveiben kifejezve változó továbbra is adhat információt arról, hogy az illető külföldi-e vagy sem.

– Foglalkozási forma: ha a foglalkozási forma megegyezik az alkalmazottal, akkor a munkaerő-piaci státus a referenciahét folyamán két értéket vehet csak fel: „dolgozott a referenciahét folyamán”, vagy „nem dolgozott távollét miatt a referenciahét során”.

– Legmagasabb iskolai végzettség: ha a referenciahét folyamán a foglalkozási formából, valamint a munkaerő-piaci státusból is arra lehet következtetni, hogy az illető 15 év alatti, a legmagasabb iskolai végzettség csak egyetlen értéket vehet fel, így az utóbbi elnyomása maga után vonja a másik két változó elnyomását is.

A tradicionális megközelítés végrehajtásához a lokális elnyomás kivételével az sdcMicro, valamint a μ -ARGUS is használható. Ahogy említettük, két kockázat mértéket javasoltunk: a k anonimitást és az m dimenziós kombinációk vizsgálatát. A választott kockázat mértéktől függően több lehetőség is rendelkezésre áll a lokális elnyomás optimális végrehajtására. A k anonimitás esetén az sdcMicro használatát javasoljuk, míg az m dimenziós kombinációknál a μ -ARGUS-t.

3. A felfedési kockázat tárgyalása

Mivel a felfedési kockázatot másképpen kell értelmezni a szintetikus és a hagyományos megközelítés esetén, ezért tárgyalása is külön történik e két módszer esetén.

3.1. EU-SILC

A szintetikus megközelítés esetén az adatállomány teljesen szintetikus, vagyis mi állítottuk elő, szimuláltuk az adatokat az eredeti állomány változóinak becsült eloszlása alapján. *Templ* és *Alfons* [2010] által készített tanulmányban található egy általános megközelítés a felfedési kockázatra teljesen szintetikus állományoknál, valamint ezen megközelítés gyakorlati alkalmazására egy példa az AMELI¹¹-projekt keretében szimulált EU-SILC állományra. A tanulmány öt felfedési forgatókönyvet említ. Az általános következtetés az, hogy még a legtöbb előzetes ismerettel rendel-

¹¹ AMELI (advanced methodology for European Laeken indicators): az európai laekeni indikátorok részletes módszertana.

kező (a szintetikusadat előállításának folyamatát ismerő) támadó esetén is nagyon alacsony a felfedési kockázat. Továbbá, még ha a támadó képes is sikeresen azonosítani egy egyedet, annak a valószínűsége, hogy az adott egyedhez tartozó adatok közel vannak a valósakhoz, rendkívül csekély.

A mi esetünkben az egyedi szerkezetű (egyedi szerkezethez közel eső) háztartások azonosíthatók. Például egy nagy háztartás, mely többször is szerepel a nyilvános mikroadatban ugyanazzal a szerkezettel (kor és nem eloszlás), valószínűleg egyedi az eredeti mintában. Azonban a hozzá tartozó bevételi adatok eltérnek a valósaktól.

Az ilyen egyedi háztartások felfedési kockázatának csökkentése érdekében megfontolandó ezen háztartások eltávolítása a nyilvános mikroadatból. Habár ez a nyilvános mikroadatfájlból számolható becslések torzításához vezetne, azonban figyelembe véve a fájlok célját, ez talán nem tekinthető akkora problémának.

3.2. EU-LFS

A hagyományos megközelítés esetén az azonosító változókat részletesen szemügyre kell venni. Az általános vélekedés alapján a k anonimitás jobban megfelelne, azonban az azonosító változók nagy száma miatt azokat kulcsváltozóként használva a k anonimitás, mint kockázati mérték, igencsak problémás lenne. Ezért ezen azonosító változók egy részhalmazát kellett kijelölni, és azokra alkalmazni a szabályt.

A másik lehetőség az azonosító változók m dimenziós kombinációinak vizsgálata lenne. Bizonyos értelemben ezekre úgy lehet tekinteni, mint a k anonimitás mértéknek az m dimenziós marginálisaira, ahol minden egyes azonosító változót kulcsváltozóknak vesszünk. Magasabb küszöbérték használata ezen m dimenziós kombinációkra nagyobb mozgásteret enged a magasabb dimenziós kombinációknak.

4. A nyilvános mikroadatfájlok használhatósága

A nyilvános mikroadat fájlok felhasználhatóságának mérésére néhány, ún. hasznossági mértéket definiáltunk, melyek az anonimizált mikroadat-állomány és a nyilvános mikroadatfájl közötti relatív eltérést számszerűsítik. Lényegében ezen mértékek

$$\frac{\left(\begin{array}{c} \text{Egy nyilvános mikroadatfájlbeli} \\ \text{indikátor értéke} \end{array} \right) - \left(\begin{array}{c} \text{Egy anonimizált mikroadatfájlbeli} \\ \text{indikátor értéke} \end{array} \right)}{\left(\begin{array}{c} \text{Egy anonimizált mikroadatfájlbeli} \\ \text{indikátor értéke} \end{array} \right)}$$

alakban írhatók fel, ahol indikátor alatt például az egyedek nem, életkor, iskolai végzettség stb. szerinti eloszlását értjük, tehát az adatok alapjellemezőit tükröző mértékeket.

Az ebben a formulában levő indikátor lehet adatspecifikus is, például az EU-SILC esetén használhatunk olyan „fő indikátort”, mint az elszegényedéssel veszélyeztetettek aránya, az EU-LFS-nél pedig a foglalkoztatottsági ráta.

Továbbá a Drechsler [2009] által javasolt konfidenciaintervallum-átfedés mértéken alapulva néhány modellalapú hasznossági mérték is szóba került. Az EU-SILC kapcsán például a következő logisztikus regressziós modell használható:

$$\log(\text{egyenértékesített rendelkezésre álló jövedelem}) \sim \text{életkor} + \text{nem} + \text{végzettség} + \text{állampolgárság} + \text{háztartás mérete}.$$

Az LFS esetén pedig a „foglalkoztatott-e” változóra vonatkozó logisztikus regresszió az életkor, iskolai végzettség, állampolgárság és háztartás mérete magyarázóváltozókkal becsülhető.

5. A projekt jelenlegi helyzete

Jelen cikk megírásakor a projektben résztvevő csapat a nyilvános mikroadatfájlok előállítására javasolt módszereket értékelte ki. A szóba jöhető eljárásokat néhány tagországi adatállományon teszteltük. Az idő nagy részét a technikai megvalósításhoz szükséges eszközök fejlesztésére, illetve a lehetséges kockázati és hasznossági mértékek megvitatására szántuk. A projekt végére (2015 december) minden résztvevőnek lehetőség szerint elő kell állítania prototípusfájlokat mind az EU-SILC-ből, mind az EU-LFS-ből. Ahhoz, hogy ezen prototípusokból „valós” nyilvános mikroadatfájlok váljanak, szükséges az Eurostat adatvédelemmel foglalkozó munkacsoportjának, valamint a témában érintett egyéb munkacsoportok jóváhagyása. Abban az esetben, ha az előbb említett munkacsoportok pozitívan döntenek a prototípusok sorsáról, még mindig nélkülözhetetlen a tagállamok jóváhagyása is ahhoz, hogy nyilvános mikroadatfájlokat adjanak ki a saját EU-SILC és/vagy EU-LFS állományaikról.

A projekt a nyilvános mikroadatokat mellett azok előállításához szükséges módszertani dokumentációt, valamint a technikai megvalósításhoz szükséges R szkripteket és egyéb szoftvermegoldásokat is szolgáltat.

Irodalom

ALFONS, A. – KRAFT, S. – TEMPL, M. – FILZMOSE, P. [2011]: Simulation of Close-to-Reality Population Data for Household Surveys with Application to EU-SILC. *Statistical Methods & Applications*. Vol. 20. Issue 3. pp. 383–407.

- BUJNOWSKA, A. [2015]: *Access to EU Microdata for Research Purposes*. Presented at the UNECE/Eurostat Work Session on statistical data confidentiality. United Nations Economic Commission for Europe. Helsinki.
- DRECHSLER, J. – REITER, J. P. [2009]: Disclosure Risk and Data Utility for Partially Synthetic Data: An Empirical Study Using the German IAB Establishment Survey. *Journal of Official Statistics*. Vol. 25. No. 4. pp. 589–603.
- EUROSTAT [2014]: *EU Labour Force Survey Database User Guide*. Version December 2014. <http://ec.europa.eu/eurostat/documents/1978984/6037342/EULFS-Database-UserGuide.pdf>
- GOUWELLEEUW, J. M. – KOOIMAN, P. – WILLENBORG, L. C. R. J. – DE WOLF, P.-P. [1998]: Post Randomisation for Statistical Disclosure Control: Theory and Implementation. *Journal of Official Statistics*. Vol. 14. No. 4. pp. 463–478.
- HUNDEPOOL, A. – DOMINGO-FERRER, J. – FRANCONI, L. – GIESSING, S. – SCHULTE NORDHOLT, E. – SPICER, K. – DE WOLF, P.-P. [2012]: *Statistical Disclosure Control*. Wiley. Hoboken.
- TEMPL, M. – ALFONS, A. [2010]: Disclosure Risk of Synthetic Population Data with Application in the Case of EU-SILC. *Computer Science*. Vol. 6344. pp. 174–186.

Summary

Partly financed by Eurostat, seven member states (represented by *Maxime Bergeat, Matthias Templ, Lydia Spies, Annu Cabrera, Péter Kristóf, Andreja Smukavec, Aleksandra Bujnowska, Peter-Paul de Wolf*) have been working on a proposal for a harmonized approach to produce public use files of the EU-SILC and EU-LFS data. The final results are due in December 2015. The current paper that was presented at the UNECE/Eurostat Work Session on Statistical Data Confidentiality in Helsinki, on 5–7 October 2015, describes the state of affairs and discusses the (preliminary) results.

For the original study, please see the Annex online at <https://www.ksh.hu/statszemle>.

A célzott adatcsere módszere a térstatisztikában

Nagy Beáta,
a KSH osztályvezetője
E-mail: Beata.Nagy@ksh.hu

A térstatisztikai adatok felfedés elleni védelmének kidolgozása és alkalmazása számos országban jelentős kihívással jár. Megfelelő felfedés elleni módszerek szükségesek a térstatisztikán alapuló hivatalos statisztikai adatok védelmére azért, hogy ne történhessen azonosítás, illetve felfedés a kiadott adatokból.

A tanulmány a térstatisztikán alapuló hivatalos statisztikai adatok egy – nem cellaelnyomást alkalmazó – felfedés elleni védelmi módszerét mutatja be, mely mind lakossági, mind gazdaságstatisztikai adatok esetén megfelelő.

Az eljárást a Központi Statisztikai Hivatal az 1 km²-es rácshálón megjelenített 2011. évi népszámlálási adatokon alkalmazta először. A szerző összefoglalja, hogy miért van szükség felfedés elleni védelemre, illetve a védendő rekordokat hogyan választják ki. Ismerteti mi is az adatcsere, valamint azt is, hogy a hivatal ezt hogyan alkalmazta.

TÁRGYSZÓ:
Térstatisztika.
Adatcsere.
Adatvédelem.

A világ egyre inkább számokban mérhető környezetté válik. Egyre több szakma – a szokványostól a nem mindennaposig – adatoktól, számszerű következtetésektől függ.

Az adatok nem egyszerűen számok, azok információt hordoznak magukban egy adott jelenségről, jelentésüket mindig az adott környezet határozza meg.

A statisztika az adatok gyűjtésének, feldolgozásának, elemzésének, értelmezésének és bemutatásának a tudománya; a természettudományi, társadalmi és gazdasági jelenségeket térben és időben is megfigyeli, így a megfigyelt egység térbeli elhelyezkedése mindig fontos szerepet játszik.

A statisztika keretein belül megkülönböztetjük a hivatalos statisztikát, melyet Magyarországon a Hivatalos Statisztikai Szolgálat tagjai készítenek. A hivatalos statisztikák földrajzi bontását jellemző módon a hierarchikus közigazgatási egységeknek megfelelően képezzük. Bár ez az eljárás tökéletesen alkalmas arra, hogy jelentések, kimutatások készüljenek adott közigazgatási határok szerinti bontásban, azonban nem teszi lehetővé azt, hogy számos olyan társadalmi-gazdasági, környezeti jelenség, melyek földrajzi és nem közigazgatási egységhez kötődnek (mint például árvizek, légszennyezés stb.) kellő mélységben elemezhetők legyenek. Ennek vizsgálatára térstatisztikai módszerek és eszközök a legalkalmasabbak.

A fogalmak használatában a KSH¹ igyekezett a nemzetközi gyakorlatot követni, ahol azonban fogalmi ütközés keletkezett. A kezdetektől használt geostatisztika (geostatistics) szó alatt egyrészt a társadalmi és gazdasági adatok térben való elhelyezését, statisztikai kutatását érthetjük, másrészt ezt a kifejezést használják az alkalmazott földtudományokban valószínűség-számítási és matematikai-statisztikai kutatásokra is (például lelőhely méretének becslésére mintafúrások alapján). Ezért célszerűbb a térstatisztika (spatial statistics) szóhasználat, melyet az Eurostat GISCO² munkacsoportja is szorgalmaz (Kádár *et al.* [2015]).

Ami igazán különböző a térstatisztika és a statisztika között az az, hogy akármi-lyen komplex és rendellenes egy jelenség, a statisztika ezen területe mindig valamilyen térbeli kapcsolatrendszer magába foglaló struktúrát keres mögötte. Ez visszavezethető arra az intuíción alapuló gondolatra, hogy a térben egymáshoz közel levő pontok rendszerint hasonló tulajdonságokkal rendelkeznek.

A térstatisztikai jelenségek elemzésére egy egyenlő oldalú rácshálós (grid) szerkezet sokkal alkalmasabb, mivel a rácscellák

¹ KSH: Központi Statisztikai Hivatal.

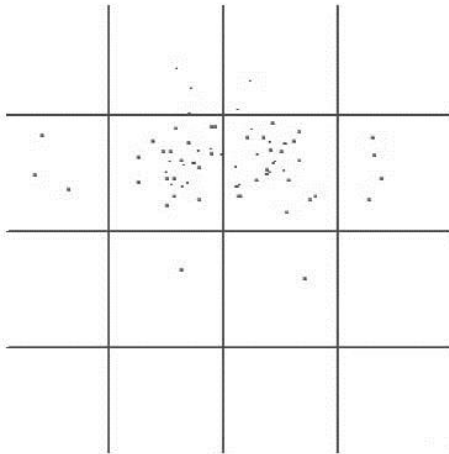
² GISCO (Geographical Information System at the Commission): a Bizottság földrajzi információs rendszere.

- mérete azonos, egyszerűvé téve így az összehasonlítást;
- időben állandók, nem változnak, mint a közigazgatási területi egységek;
- könnyen integrálhatók más tudományos témájú (például meteorológiai) adatokkal;
- méretét tekintve a rácshálós rendszerek hierarchikusan is felépíthetők, így alkalmazkodnak a megfigyelni kívánt területhez;
- összevonhatók olyan területté (például hegyvidéki régióvá, vízgyűjtő területté stb.), ami egy adott elemzés speciális igényeinek leginkább megfelel.

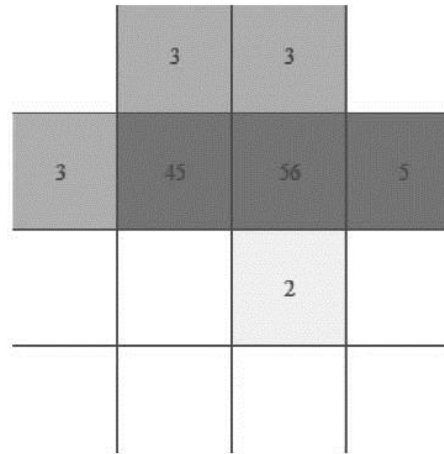
Az egységes rácshálón megjelenített statisztikai adatok tehát földrajzilag egy Descartes-koordinátákkal ellátott rácsháló – többnyire négyzet alapú – celláihoz vannak rendelve.

A rácscellák létrehozásához nélkülözhetetlen egy térbeli pontosságú geokódokkal³ rendelkező ponthalmaz. (Lásd az 1. ábrát.) Az esetek többségében ezek vállalati és címregiszterek földrajzi koordinátákkal, amelyekhez az adott statisztikai információ köthető. Ezen elemi szintű adatok (amelyek a térkép egy pontjára mutatnak) magasabb szintre, például négyzet alakú cellákra aggregálhatók az elemzés céljától függően. (Lásd a 2. ábrát.)

1. ábra. Rácshálóba szervezett pontok



2. ábra. Adatok aggregálása rácscellákba



³ Földrajzi azonosító. Az egyes objektumokhoz vagy objektumcsoportokhoz rendelt kód, amely egyértelműen azonosítja az objektum jellegét és földrajzi helyzetét. Felhasználásával többek között lehetővé válik a különböző adatbázisok összekapcsolása. Magyarországon rendelet szabályozza a geokód előállítási formáját.

2010 elején egy GEOSTAT nevezetű ESSnet⁴ projektet indított az Eurostat azért, hogy népszerűsítse a rácshálóalapú statisztikát, továbbá hogy elősegítse az EU-n belül a statisztikai és térstatisztikai adatok közös informatikai infrastruktúrájának kialakítását. A cél az egységes irányelvek kifejlesztése volt a térstatisztikai adatgyűjtéshez és az adatok feldolgozásához az ESR⁵-en belül. (Petri [2014])

Az előbb említett kezdeményezésnek része a „Népszámlálási adatok megjelenítése egy európai rácsháló adatbázisban” elnevezésű GEOSTAT 1 projekt, melynek célja a 2011. évi népszámlálás háztartás- és személyi állomány számos jellemzőjének megjelenítése volt 1 km²-es rácshálón. (European Forum for Geostatistics [2011]).

A célzott adatcsere magyarországi implementálása a GEOSTAT 1C projekt keretében valósult meg.

1. Miért van szükség felfedés ellen védelemre?

A hivatalos statisztikai adatok számos adathozzáférési csatornán⁶ érhetőek el. A legtöbb felhasználó a kiadványokat vagy a honlapon található táblázatos adatokat használja. De speciális igények esetén az adathozzáférés történhet például egyedi kérésre összeállított táblázatos aggregált adatok, anonimizált mikroadatok, biztonságos környezetben történő hozzáférés vagy nyilvános mikroadatfájlok⁷ formájában. (Részletesebb információk az adat-hozzáférésről jelen szám több cikkében elérhetők: Mátyás-Bodovics [2015], Vereczkei [2015].) Az adatok magukban hordozzák az azonosítás⁸ és a felfedés kockázatát,⁹ ezért a statisztikai adatokat azok kiadása előtt minden esetben felfedés elleni védelemmel kell ellátni.

Az adatvédelem fő feladata, hogy különböző matematikai-statisztikai módszerek és eljárások alkalmazásával minimalizálja a statisztikai egységek azonosításának, illetve felfedésének a kockázatát úgy, hogy közben a lehető legkisebb információvesztést okozza.

⁴ ESS (European Statistical System): Európai Statisztikai Rendszer (ESR). <http://www.essnet-portal.eu/project-information/geostat/>

⁵ https://www.ksh.hu/nemzetkozi_adatok_europai_statisztikai_rendszer

⁶ Adat-hozzáférési csatorna: az adathozzáférés módjai, melyeken keresztül az adatok a felhasználókhöz jutnak. Különböző adat-hozzáférési csatornák esetén különböző feltételek mellett lehet az adatokhoz jutni.

⁷ Nyilvános mikroadatfájl (public use file): olyan nyilvánosan hozzáférhető mikroadat-állomány, melyben a statisztikai egységek azonosítási, felfedési kockázata minimális.

⁸ Azonosítási kockázat: annak valószínűsége, hogy a támadó legalább egy statisztikai egységet azonosít a közzétett adatok alapján.

⁹ Felfedési kockázat: az azonosítási kockázat, valamint a statisztikai egységre vonatkozó új információ nyilvánosságra kerülésének valószínűsége.

A területi változó mindig kiemelt figyelmet igényel a felfedés elleni védelem végrehajtása során, hiszen minél kisebb a vizsgált területi egység és minél kisebb esetszámok fordulnak elő, az adatvédelmileg annál nagyobb kockázatot jelent. Alacsony gyakoriságú cellaértékeknel ugyanis megnő a kockázata az egyedek beazonosíthatóságának, így jelentősen megnövekszik a felfedés lehetősége is. Ez a veszély hatványozottan előfordulhat a térstatisztika területén, ahol az adatok speciális megjelenítése tovább nehezíti az adatvédelmet.

Feltételezhetjük, hogy az emberek ismerik azt a környéket, ahol élnek vagy azokat (főbb demográfiai jellemzőiket, például foglalkozásukat, nemüket, életkorukat), akik közel laknak hozzájuk, ezért egy esetleges támadó¹⁰ gyakran az adott földrajzi területen belül próbál egyedeket beazonosítani, illetve róluk információt felfedni. Ez a felfedési forgatókönyv végigkísérte a felfedés elleni védelem kialakítását, mivel Magyarországon számos ritkán lakott terület van, mindemellett a hazai jogszabályok betartására is kiemelt figyelmet kellett fordítani.

Az irányadó jogszabályok a következők voltak:

– *1993. évi XLVI. törvény a statisztikáról*,¹¹ melyben a 17. § (2) kimondja, hogy „nem lehet nyilvánosságra hozni természetes és jogi személy, valamint a jogi személyiséggel nem rendelkező adatszolgáltatóval kapcsolatba hozható adatot, azaz egyedi adatot”.

– *170/1993. (XII.3.) Kormányrendelet a statisztikáról szóló 1993. évi XLVI törvény végrehajtásáról*,¹² melyben a 19. § kimondja, hogy „összesítve sem lehet nyilvánosságra hozni olyan adatot, amelynél az adatszolgáltatók száma háromnál kevesebb”.

2. Felfedés elleni védelmi módszerek a térstatisztikában

Az adatok típusát tekintve megkülönböztetünk mikro- és a táblázatos adatokat.¹³ A mikroadat-állomány rekordok sorozatából áll, és a megfigyelési egységek adatait tartalmazza. Személyre vonatkozó mikroadat például az állomány egy sora, mely egy személy adatait reprezentálja. A megfigyelési egység bármilyen statisztikai egység

¹⁰ Támadó: olyan felhasználó, aki megkísérel egy statisztikai egységet a mikroadat valamely rekordjához hozzárendelni, vagy aggregált adatokból bizonyos statisztikai egységek tulajdonságaira következtetni.

¹¹ http://www.ksh.hu/docs/bemutakozas/hun/1993_evi_XLVI_torveny_a_statisztikarol.pdf

¹² http://www.ksh.hu/docs/bemutakozas/hun/Statisztikai_torveny_vegrehajtasa.pdf

¹³ Táblázatos adat: olyan táblázatos formába rendezett adatállomány, amely aggregált adatokat tartalmaz.

lehet (például vállalkozás, háztartás, személy stb.) Táblázatos adatok előállhatnak a mikroadatokban található megfigyelési egységek aggregálásával, illetve más táblázatok felhasználásával.

Esetünkben a 2011. évi népszámlálási mikroadat-állomány a következő információkat tartalmazta a személyekről: nem, korcsoport,¹⁴ a személyhez tartozó rácscella kódja, településkód, technikai azonosító. Ebből a mikroadatból a következő népszámlálási témakörökre rácsáló-adatbázist (táblázatos adat) kellett előállítani: nem, korcsoport, nem és korcsoport.

A térstatistikai adatok védelmére különböző felfedés elleni módszer létezik.

– *Védendő cellák egységes szimbólummal történő elnyomása, helyettesítése.* Lehetséges szimbólum a védendő cellák helyettesítésére a „99 999” vagy „0”. Az utóbbi Norvégiában gyakran alkalmazott eljárás, annak ellenére, hogy a 0-val való helyettesítéssel számos cella „lakatlan” válik, valamint így a populáció összesenre vonatkozó értékek sem egyeznek meg a felfedés elleni védelemmel nem ellátott mikroadatból származó összesenekkel.

– *Védendő cellák 3-mal való helyettesítése.* Ebben az esetben a cellák azon jellemzője, hogy lakott-e vagy lakatlan, megmarad, de a populációra vonatkozó összesenek itt is eltérők lesznek a felfedés elleni védelemmel nem ellátott mikroadatétól.

– *A cellák méretének növelése.* Például a cellák 1×1 km-ről 2×2 km-re növelésével csökkenteni lehet a védendő cellák számát.

– *Cellák egyesítése.* Németországban használt eljárás, mely során a cellák különböző méretűek és alakzatúak lehetnek a felfedés elleni védelem elérése érdekében.

A GEOSTAT 1C projekt kritériumai, illetve az adatvédelemre vonatkozó előírásai¹⁵ megnehezítették a térstatistikában gyakran használt módszerek alkalmazását, így olyan védelmi eljárásokat is fontolóra kellett venni, melyek nem térstatistika

¹⁴ Korcsoport: 0–14 éves; 15–64 éves; 65 év feletti.

¹⁵ A GEOSTAT 1C projekt adatvédelemre vonatkozó előírásai: „1. A nemzeti rácsáló-adatbázisok teljes népessége vagy népességbontása meg kellett, hogy egyezzen a népszámláláskor közzétett azonos témájú népességgel. 2. Ha ez a feltétel az alkalmazott adatvédelmi eljárások miatt nem volt teljesíthető, akkor a hiányzó vagy hozzáadott személyek számát, a módszert és az okokat dokumentálni kellett a metaadatokban. 3. Az adatvédelem miatt más cellába helyezett személyeket, és általánosságban az adatvédelem alá tartozó személyek számát is dokumentálni kellett. 4. A lakott cellákat meg kellett tartani, és nem szabadott lakatlannak jelölni. Ez vonatkozott a teljes népességre, és a népességbontásokra is. 5. A nem lakott cellákat meg kellett tartani, és nem szabadott lakottnak jelölni. Ez vonatkozott a teljes népességre, és a népesség bontásokra is. 6. Integer típusú adatként kellett jelenteni a népességadatokat. Ez azt jelentette, hogy az adatvédelmi intézkedések nem használhattak string vagy alfanumerikus típusú adatokat. 7. A cellarácsoknak meg kellett maradniuk 1 km^2 -esnek és négyzet alakúnak. Ez azt jelentette, hogy az egyes lakossági küszöbértékek elérése céljából történő cellaegyesítések nem voltak megengedettek” (Kádár et al. [2015] 384. old.)

specifikusak, de használatukkal biztosítható az adatok felfedés elleni védelme és emellett teljesíthetők a projekt feltételei is. Mindezeket szem előtt tartva a legmegfelelőbbnek az adatsere módszere bizonyult, melyet a következő fejezetben részletezünk.

3. Az adatsere módszeréről általánosságban

Az adatsere mikroadatok védelmére alkalmazott eljárás, melynek lényege, hogy megcseréljük a kiválasztott változó vagy változók egyes rekordjaihoz tartozó értékeket. Alapesetben minden mikroadatrekord két részvektorra, v_i -re és z_i -re osztható, ahol z kulcsváltozó vagy a kulcsváltozók vektora. Definiáljunk (i, j) párokat egy P halmazban (ez a halmaz jelöli azon párok halmazát, melyek értékeit az adatsere során kicseréljük egymásra), ahol i és j két különböző eleme az $\{1, \dots, n\}$ mikroadatrekordok halmazának. Az adatsere folyamán z_i és z_j értékei cserélődnek minden (i, j) pár esetén a P halmazban. (Willenborg [2001])

Az adatsere formálisan a következőképpen írható le:

Legyen $\delta_{z_i, j} = 1$, ha a z változó értékét az i és j -edik rekordok között kicseréljük; különben $\delta_{z_i, j} = 0$. Ekkor (Bartus [2013]):

$$z_i^a = (1 - \delta_{z_i, j})z_i + \delta_{z_i, j}z_j,$$

$$z_j^a = (1 - \delta_{z_i, j})z_j + \delta_{z_i, j}z_i.$$

A csere, mely magába foglalja a megfelelő pár megtalálását, egy transzformációs technika, ami garantálja a sor- és oszlopösszesenek változatlanóságát.

Példa: A következő mikroadatban található három változó (U, V, T) és hét rekord. Azt feltételezzük, hogy az U változót felfedés elleni védelemmel kell ellátni, mert az ötödik sorban lévő értéke egyedi, vagyis az eredeti mikroadat nem adható ki

Az 1. b) táblázat tartalmazza az adatsere utáni mikroadatot. Az U' változó az U változóból annak 4. és 5. rekordjához tartozó értékének megcserélésével áll elő. A támadó az adatsere utáni mikroadatból már csak annyit tudhat, hogy az U' változón adatszerét végeztünk el, de azt nem tudhatja, hogy mely rekord(ok) kerültek védelem alá.

1. táblázat

Általános példa az adatsere módszerére

a) Eredeti mikroadat				b) Adatsere utáni mikroadat			
Rekord	U	V	T	Rekord	U'	V'	T'
1.	0	1	0	1.	0	1	0
2.	0	1	0	2.	0	1	0
3.	0	0	1	3.	0	0	1
4.	0	0	1	4.	1	0	1
5.	1	1	1	5.	0	1	1
6.	1	0	0	6.	1	0	0
7.	1	0	0	7.	1	0	0

Az adatsere elvégzésével a mikroadatból nyert két-, illetve háromdimenziós gyakorisági táblázatban az összesenek megegyeznek az eredeti mikroadatból nyerhető kétdimenziós gyakorisági táblázat összesenjeivel, annyi különbséggel, hogy a két-, illetve háromdimenziós táblázat szerinti megoszlások eltérnek az eredeti mikroadatból nyerhető megoszlásoktól. Ezzel az eltéréssel biztosítható az, hogyha a támadó azonosítást próbál végezni, akkor nem lehet biztos abban, hogy az adott rekordhoz valóban az adott értékek párosulnak.

3. ábra. Az 1. táblázat mikroadataihoz tartozó kétdimenziós gyakorisági táblázatok

		V					V'		
U	0	1	Összesen		U'	0	1	Összesen	
0	2	2	4		0	1	3	4	
1	2	1	3		1	3	0	3	
Összesen	4	3	7		Összesen	4	3	7	

		T					T'		
U	0	1	Összesen		U'	0	1	Összesen	
0	2	2	4		0	2	2	4	
1	2	1	3		1	2	1	3	
Összesen	4	3	7		Összesen	4	3	7	

4. ábra. Az 1. táblázat mikroadataihoz tartozó háromdimenziós gyakorisági táblázatok

		T				Összesen
		0		1		
		V		V		
		0	1	0	1	
U	0	1	0	1	Összesen	
0	0	2	2	0	4	
1	2	0	0	1	3	
Összesen	2	2	2	1	7	

		T'				Összesen
		0		1		
		V'		V'		
		0	1	0	1	
U'	0	1	0	1	Összesen	
0	0	2	1	1	4	
1	2	0	1	0	3	
Összesen	2	2	2	1	7	

4. A felfedés elleni védelemmel ellátandó rekordok meghatározása

A felfedés elleni védelmi módszer alkalmazásához meg kell határozni a felfedés elleni védelemmel ellátandó rekordokat a mikroadat-állományban. Ezek azok a rekordok, melyek valamely jellemzői (például nem, korcsoport) vagy jellemzőkombinációi (például nem-korcsoport) alapján három alatti gyakoriságban fordulnak elő az általunk vizsgált mikroadat-állományban, így esetükben a ritka előfordulások miatt megalapozott az azonosítás, valamint a felfedés veszélye.

A felfedés elleni védelemmel ellátandó rekordok meghatározásához először ki kell jelölni a mikroadat bizonyos jellemzőit (például kulcsváltozóit¹⁶), melyekből kulcsokat¹⁷ szükséges képezni.

Legyen A , B , C a vizsgált mikroadatban kijelölt három változó. A kijelölt változókból azok értékeinek lehetséges kombinációit képezzük. A 2. táblázat második sora mutatja, hogy az A , B , C változókból milyen típusú (egyszeres, kétszeres, háromszoros) kombinációk állíthatók elő.

A kijelölt változók, illetve a kombinációik értékeinek mikroadat-állományban előforduló gyakoriságai alapján ún. „flag”¹⁸ értékek képzésére kerül sor. Ezt úgy végezzük el, hogy az állomány minden egyes rekordjához hozzárendeljük az alapjellemezőik (azaz a három kijelölt változó), illetve a kombinációik szerinti gyakoriságokat. Ezekből

¹⁶ Olyan változókat választunk kulcs változónak, amelyek a mikroadat-állományból általában ismertek egy vagy több statisztikai egységre, ezáltal lehetségessé válhat a statisztikai egységek azonosítása a választott változók alapján.

¹⁷ Kulcs: mikroadat-állomány néhány változójának vagy néhány változó kategóriájának kombinációja.

¹⁸ A flag egy kétértékű változó, amely valamilyen előre definiált feltétel(ek) teljesülését vagy nem teljesülését jelöli.

kiindulva megállapítható, hogy mely rekordok fordulnak elő az adott jellemző vagy kombináció szerint 3 alatti gyakoriságban a mikroadatban. Tehát ha tudjuk a jellemzők és a kombinációk szerint a rekordokhoz tartozó gyakoriságokat, akkor könnyedén képezhetők a flag értékek. A flag a következő értékeket veheti fel:

- 1, ha felfedés elleni védelemmel kell ellátni a mikroadat-állományban az adott rekordot (azaz ha a jellemző vagy jellemző kombináció adott rekordra felvett értéke 3 alatti előfordulással rendelkezik a teljes mikroadat-állományban);
- 0, egyébként.

2. táblázat

Példa a rekordonkénti flag táblázatra

Rekord/ID-azonosító	A	B	C	$A \times B$	$A \times C$	$B \times C$	$A \times B \times C$	Flag összes
1	1	1	0	1	1	1	1	6
2	1	0	0	1	1	0	1	4
3	0	0	1	0	1	1	1	4
4	0	0	0	1	0	0	1	2
5	0	0	0	0	0	0	1	1

Megjegyzés. ID: adott személy technikai azonosítója.

A rekordonkénti flag értékek, majd ezen értékek összesen adatának képzésével határozható meg, hogy mely rekordokat és milyen sorrendben kell felfedés elleni védelemmel ellátni. A flag-ek összegét rekordonként szükséges képezni, majd csökkenő sorrendbe rakni. Így az első sorokban szerepelnek azok a rekordok, amelyek a legkönnyebben azonosíthatók, azaz amelyek a legkockázatosabbak (Lásd a 2. és a 3. táblázatot.)

A flag felvehető értékeiből következik, hogy egy rekordot kockázatosnak nevezünk, ha a flag értékek összege nagyobb, mint 0, azaz a rekordon felfedés elleni védelmet szükséges alkalmazni. Következésképpen a végső táblázatos adatban egy rácscella kockázatosnak tekinthető, ha legalább egy felfedés elleni védelemmel ellátandó rekordja van.

Az általunk vizsgált mikroadat-állományban három jellemzőt (a nemet, a korcsoportot, a rácsellát) választottunk ki, mivel ezen jellemzők alapján állnak elő a végső „táblázatos adatok”.

A kulcsok – a kiválasztott jellemzők, illetve kombinációik (esetünkben valamennyi változó) – a következők:

- *alapp jellemzők*: 1. nem, 2. korcsoport, 3. rácscella;
- *kettes kombináció*: 4. nem \times rácscella, 5. rácscella \times korcsoport, 6. nem \times korcsoport;
- *hármass kombináció*: 7. nem \times korcsoport \times rácscella.

3. táblázat

Rekordonkénti flag táblázat előállítás, melynek utolsó oszlopa azt mutatja, hogy mennyire kockázatos egy adott rekord

Mikroadat					Gyakoriság							Flag táblázat							
ID	Településkód	Nem	Rácscella	Korcsoport	1.	2.	3.	4.	5.	6.	7.	1.	2.	3.	4.	5.	6.	7.	Flag összes
1	11	1	53	64	247 592	352 642	4	4	2	175 285	2	0	0	0	0	1	0	1	2
2	11	1	53	65	247 592	91 582	4	4	2	33 317	2	0	0	0	0	1	0	1	2
3	11	1	53	65	247 592	91 582	4	4	2	33 317	2	0	0	0	0	1	0	1	2
4	11	1	53	64	247 592	352 642	4	4	2	175 285	2	0	0	0	0	1	0	1	2
5	11	1	57	65	247 592	91 582	11	7	2	33 317	1	0	0	0	0	1	0	1	2
6	11	2	57	65	272 739	91 582	11	4	2	58 265	1	0	0	0	0	1	0	1	2
7	11	1	33	14	247 592	76 107	46	23	5	38 990	2	0	0	0	0	0	0	1	1
8	11	1	33	14	247 592	76 107	46	23	5	38 990	2	0	0	0	0	0	0	1	1
.	11	1	57	64	247 592	352 642	11	7	9	175 285	6	0	0	0	0	0	0	0	0
.
.	44	2	30	64	272 739	352 642	3	2	2	177 357	2	0	0	0	1	1	0	1	3
.	44	1	28	64	247 592	352 642	4	2	3	175 285	2	0	0	0	1	0	0	1	2
<i>i</i>	44	1	56	14	247 592	76 107	180	75	15	38 990	6	0	0	0	0	0	0	0	0
.	44	2	56	65	272 739	91 582	180	105	45	58 265	31	0	0	0	0	0	0	0	0
.
.	88	1	9	64	247 592	352 642	1	1	1	175 285	1	0	0	1	1	1	0	1	4
.	88	2	13	65	272 739	91 582	5	3	2	58 265	2	0	0	0	0	1	0	1	2
.	88	2	22	14	272 739	76 107	30	15	4	37 117	2	0	0	0	0	0	0	1	1
.	88	1	14	64	247 592	352 642	11	5	6	175 285	3	0	0	0	0	0	0	0	0
<i>n</i>	88	2	14	64	272 739	352 642	11	6	6	177 357	3	0	0	0	0	0	0	0	0

5. Célzott adatcsere

Az adatcsere lehet random vagy célzott. Előbbi esetben értelemszerűen random módon történik a csere, míg célzott adatcserekor (ahogyan a neve is utal rá) a felfedés elleni védelemmel ellátandó rekordokat párosítjuk, azt nem bízunk a teljes véletlenre.

Esetünkben minden egyes felfedés elleni védelemmel ellátandó rekord korcsoport szerinti cseréjére (swapping) sor kerül. Ez azt jelenti, hogy minden egyes felfedés elleni védelemmel ellátandó rekordhoz (A -hoz) ún. „partnert” (cserepartner B -t) szükséges találni.

A „cserepartnernek” a következő feltételeknek kell megfelelnie rekord szinten:

- a vizsgált korcsoporttól különböző korcsoportú legyen;
- a vizsgált cellától eltérő cellában helyezkedjen el, azaz amely cellába tartozik egy adott rekord, attól különböző cellában kell keresni annak „párját”;
- a vizsgált településen belül tartózkodjon, melynek fő oka, hogy lecsökkenjen a párkeresés során a cellák közötti távolság, azaz a vizsgált rekord párja csak egy bizonyos területről legyen választható.

A cserepartner keresését mindig a felfedés elleni védelemmel ellátandó rekordok között kell kezdeni azért, hogy minél kevesebb olyan rekordot vonjunk be az adatcserebe, amelyet nem kellene felfedés elleni védelemmel ellátni. Ha nincs a kritériumoknak megfelelő rekord, akkor a felfedés elleni védelemmel nem jelölt rekordok (azaz ahol a „flag érték összege” 0-val egyenlő) közül választható ki a partner, de még mindig településen belül maradva.

A partner kiválasztása akkor történik meg, ha nemcsak a rekord feltételeknek felel meg, hanem a rácscellákra vonatkozó feltételek is teljesülnek rá, amelyek a következők:

- a felfedés elleni védelemmel ellátandó rekord cellájától eltérő cellában levő rekord legyen a partner, de olyan cella, amelynek legalább egy rekordját felfedés elleni védelemmel kellett ellátni;
- ezzel egy időben az is teljesüljön, hogy a fennmaradt legkisebb gyakoriságú (adott cellára vonatkozó populációs gyakoriság) „adatvédelmileg kezelt” cellából kerüljön ki a partner (tehát olyan cellából, melyben már történt csere).

Ha nincs olyan rekord partner (cserepartner B), amelyre a rekord és cellára vonatkozó feltételek teljesülnének, akkor a legalacsonyabb gyakoriságú nem védendő cellából választható ki a megfelelő cserepartner. Ezáltal biztosítva azt, hogy minél

kevesebb olyan cellát érintünk az adatcsere során, amelyet nem kellene felfedés elleni védelemmel ellátni.

4. táblázat

Célt az adatcsere előtti mikroadatban, illetve a csere utáni mikroadat

Célt az adatcsere előtti mikroadat					A módszerhez kapcsolódó releváns információk: rácscellákról és rekordokról			Cserepárok	Mikroadat célt az adatcsere után					
ID	Településkód	Nem	Korcsoport	Rácscella	Flag összes	Rácscellaszakosság	Rácscella flag		ID	Településkód	Nem	Korcsoport	Rácscella	Adv
1	12	2	65	15	3	3	1	1	1	12	2	64	15	sw
3	12	2	64	13	3	4	1	1	3	12	2	65	13	sw
7	12	1	64	20	2	3	1	10	7	12	2	65	20	sw
8	12	1	64	16	2	6	1	12	8	12	1	14	16	sw
9	12	1	64	16	2	6	1	13	9	12	1	14	16	sw
10	12	2	65	48	2	11	1	10	10	12	1	64	48	sw
11	12	2	64	48	1	11	1	15	11	12	2	14	48	sw
12	12	2	64	48	1	11	1	16	12	12	1	65	48	sw
13	12	1	14	21	1	61	1	12	13	12	1	64	21	sw
14	12	1	14	21	1	61	1	13	14	12	1	64	21	sw
15	12	2	64	16	0	6	1		15	12	2	64	16	
17	12	2	14	21	0	61	1	15	17	12	2	64	21	sw
18	12	1	65	21	0	61	1	16	18	12	2	64	21	sw

$i+1$	113	1	65	62	1	11	1	21359	$i+1$	113	1	64	62	sw
	113	2	65	62	1	11	1	21360		113	2	64	62	sw
	113	2	65	62	1	11	1	21361		113	1	64	62	sw
	113	2	64	62	0	11	1			113	2	64	62	
	113	1	64	62	0	11	1			113	1	64	62	
	113	1	64	61	0	222	0	21359		113	1	65	61	sw
	113	2	64	61	0	222	0	21360		113	2	65	61	sw
	113	1	64	61	0	222	0	21361		113	2	65	61	sw
n	113	2	65	61	0	222	0		n	113	2	65	61	

Megjegyzés. Adv – felfedés elleni védelemmel ellátott rekordok függetlenül attól, hogy kockázatos volt-e a rekord vagy sem; sw – a cserét jelöli.

Minden egyes felfedés elleni védelemmel ellátandó rekordot (A-t) és párjukat (cserepartner B-t) megkerestük a bemutatottak szerinti célt az adatcsere során.

Minden olyan cella adatvédelmileg kezelésként kerül megjelölésre a táblázatos adatban, amelyet valóban felfedés elleni védelemmel kell ellátni (azaz tartalmaz legalább egy olyan rekordot, amely kockázatos), továbbá azon cellákat is, melyek bár nem tartalmaznak kockázatos rekordot, de partnerkeresés során kénytelenek voltunk belőlük választani cserepartneret egy kockázatos rekordhoz.

A párok megtalálása után a „korcsoport” egyértelműen meg fog változni minden cserébe bevont rekord esetén, mivel feltételként szerepelt a párkeresésben. Fennmaradt azonban még egy változó, a „nem”, amit szintén csere alá vontunk, hogy ezen változó alapján se legyen lehetőség azonosításra. Ez azt jelenti, hogy amikor a korábbiakban leírt feltételek alapján megtaláljuk A -hoz a B párt, akkor nemcsak a „korcsoport” cserélődik ki a párok között, hanem a „nem” is. Mivel a „nem” változóhoz két felvehető érték tartozik (férfi/nő), így a „nem” csak bizonyos esetekben módosul, mégpedig akkor, amikor az A neme különböző a cserepartner B nemétől. (Lásd a 4. táblázatot.) A párok (azaz a felfedés elleni védelembe bevont rekordok) „nem” ismérvértéke az esetek 49,89 százalékában változott meg.

Technikai megvalósítás. A felfedés elleni védelmi módszer során két szoftvert alkalmazunk. A kockázatos rekordok megtalálása STATA szoftverrel, míg a célzott adatcsere Excel makrók segítségével történik.

6. A módszer előnyei

– A módszer biztosítja azt, hogy a teljes népességre és különböző népességbontásokra vonatkozó adatok megegyezzenek a népszámlálásból az ország összesenre vonatkozó népességszámokkal, továbbá a cellákra vonatkozó gyakoriságok is változatlanok maradnak. (Lásd az 5. táblázatot.)

– Minden kockázatos rekord cseréjére sor kerül.

– Nem történik változás a cellák lakottságára, illetve lakatlanságára vonatkozólag (tehát a lakott cella lakott, a lakatlan pedig lakatlan marad a csere után).

– Az adatcsere egyidejűleg egy vagy több változón is alkalmazható, a felhasználói igények figyelembevételével történhet a változók kiválasztása.

– Az adatcsere megnehezíti a támadó dolgát, hiszen nem lehet biztos abban, hogy a védett állomány a tényleges értékeket mutatja.

5. táblázat

Rácscellánként aggregált fiktív adatok korcsoportos bontásban, célzott adatscere előtt és után

Cella	Célzott adatscere								
	előtt				Adatvédelmileg kezelt	után			
	0–14	15–64	65+	Összesen		0–14	15–64	65+	Összesen
	éves				éves				
444	3	15	4	22	+	3	16	3	22
555	3	17	4	24	+	1	17	6	24
888	3	17	7	27	+		20	7	27
889		4	4	8	+	2	2	4	8
999	1	5		6	+	4	1	1	6
111	1	2		3	+	1	1	1	3
112	47	254	109	410		47	254	109	410
113	61	327	121	509		61	327	121	509
115		10	3	13	+	2	11		13
120	123	308	124	555		123	308	124	555
130	69	307	98	474		69	307	98	474
190		1	2	3	+	1	2		3
⋮									⋮
Összesen	8 333	37 778	9 444	55 555		8 333	37 778	9 444	55 555

Megjegyzés. A számok az adott kategóriába tartozó személyek számát reprezentálják, például a 444-es cellában, a 0–14 éves korcsoportban 3 válaszadó szerepel. Az „+” jelöli azt, hogy az adott cella felfedés elleni védelemmel ellátott-e. Minden egyes „+”-szal jelölt cella esetében a cellán belüli megoszlások megváltoznak, míg a sor és oszlop összesenek változatlanok maradnak.

7. A felfedés elleni védelem adatok minőségére gyakorolt hatása

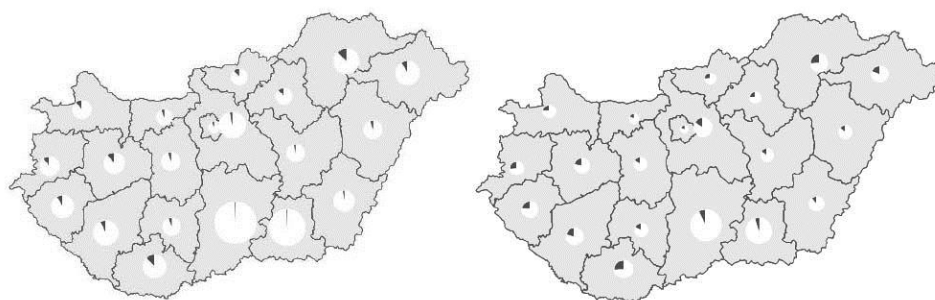
Minden felfedés elleni védelmi módszernek negatív hatása van az adatok használhatóságára vonatkozóan. A lehető legkisebb információvesztés elérése érdekében szükséges megtalálni azt az optimális megoldást, mellyel a lehető legkevesebb rekord kerül bevonásra az adatscere folyamatába.

Az 5. ábra szemlélteti az adatvédelmileg kezelt rekordok és cellák arányát meggyéknként.

5. ábra. Megyéenkénti adatvédelem

a) Adatvédelmileg kezelt rekordok megyéenként*

b) Adatvédelmileg kezelt cellák megyéenként**



* A kör mérete az adatvédelmileg kezelt rekordok számát szimbolizálja megyéenként. Fekete a körben: azon rekordok aránya, melyek nem voltak kockázatosak, de módosításukra sor került az adatvédelel. Fehér a körben: azon rekordok aránya, melyek kockázatosak voltak.

** A kör mérete az adatvédelmileg kezelt cellák számát szimbolizálja megyéenként. Fekete a körben: azon cellák aránya, melyek adatvédelmileg nem voltak kockázatosak, de a megfelelő felfedés elleni védelem ellátásához szükség volt a bevonásukra, így adatvédelmileg kezeltként jelöltük őket. Fehér a körben: azon cellák aránya, melyek kockázatosak voltak.

A magyarországi népesség 0,38 százaléka bizonyult kockázatosnak a védendő rekordok meghatározása során. A kiválasztott felfedés elleni védelmi módszer alkalmazása után a teljes népesség 0,45 százaléka került csere alá, azaz a teljes népesség 0,07 százalékát vontuk be a cserébe annak ellenére, hogy nem kellett volna felfedés elleni védelemmel ellátni.

A lakott cellák 47,64 százalék bizonyult kockázatosnak. Az adatcsere elvégzése után ez a százalék 50,73 százalékra növekedett, azaz a cellák 3,09 százaléka esetén került sor adatcsereére, annak ellenére, hogy nem kellett volna felfedés elleni védelemmel ellátni.

8. Összegzés

A statisztika terén is új lehetőséget és kihívást jelent a térstatisztika. A módszerek segítségével az eddigieknél sokkal komplexebb adatok nyerhetők ki, illetve a jelenségek rugalmasabban elemezhetők, melyekhez térképes megjelenítés is társul. Egy egységes rácsháló-struktúra olyan módon ad lehetőséget az adatok összehasonlítására, mint eddig még soha.

Ezekhez az új megjelenítési formákhoz az adatvédelem terén is alkalmazkodni kell. Számos okból kifolyólag a célzott adatcsere módszere a térstatisztikában is

alkalmasnak bizonyult. Először is a módszer alkalmazása mikroadatokon biztosítja azt, hogy a belőle képzett aggregátumok is felfedés elleni védelemmel rendelkeznek, így nem szükséges további felfedés elleni védelmi technikát alkalmazni a képzett aggregátumokon. Továbbá változatlanok maradnak a teljes népességre és a különböző népességbontásokra vonatkozó adatok és a cellákhoz tartozó gyakoriságok is. Csupán a cellán belüli megoszlások térnek el az eredetitől az adatcserébe bevont celláknál.

A módszerrel a rekordoknak csak nagyon alacsony százalékát (1%) szükséges felfedés ellen védeni, ellentétben olyan eljárásokkal, ahol a rekordok 5–15 százaléka is védelem alá kerül. Ebből fakadóan az információvesztés is igen alacsony.

A célzott adatcsere egy lehetséges felfedés elleni védelemként tartandó számon, ha a cellaelnyomás nem lehetséges a túl nagy információvesztés miatt, vagy az adott igények nem kielégíthetők szimpla cellaelnyomással. Mindemellett hasznos tapasztalatként szolgál a KSH jövőbeli térstatisztikai fejlesztéseire.

Irodalom

- BARTUS T. [2013]: Adatcserével anonimizált mikroadatok használhatósága – Egy szimulációs vizsgálat tanulságai. *Statisztikai Szemle*. 91. évf. 5. sz. 466–497. old.
- BROWN, D. [2003]: *Different Approaches to Disclosure Control Problems Associated with Geography*. Working Paper No.14. Eurostat. Brussels.
- EUROSTAT [2011]: *Population Grids*. http://ec.europa.eu/eurostat/statistics-explained/index.php/Population_grids
- EUROPEAN FORUM FOR GEOSTATISTICS [2011]: *Geostat 1A – Representing Census Data in a European Population Grid*. Final report. http://ec.europa.eu/eurostat/documents/4311134/4350174/ESSnet-project-GEOSTAT1A-final-report_0.pdf
- HUNDEPOOL, A. – DOMINGO-FERRER, J. – FRANCONI, L. – GIESSING, S. – LENZ, R. – NAYLOR, J. – SCHULTE NORDHOLT, E. – SERI, G. – DE WOLF, P-P. [2010]: *Handbook on Statistical Disclosure Control*. Version 1.2. http://neon.vb.cbs.nl/casc/.%5CSDC_Handbook.pdf
- KÁDÁR I. – BARTHA K. – NAGY B. – FÁBIÁN ZS. [2015]: Térstatisztika a Központi Statisztikai Hivatalban. *Területi Statisztika*. 55. évf. 4. sz. 380–391. old.
- LONGHUST, J. – TROMANS, N. – YOUNG, C. [2011]: *Statistical Disclosure Control for the 2011 UK Census*. Office for National Statistics. London.
- MATHERON, G. [1963]: Principles of Geostatistics. *Economic Geology*. Vol. 58. pp. 1246–1266.
- MÁTYÁS-BODOVICS E. [2015]: A Központi Statisztikai Hivatal által kezelt statisztikai adatokhoz történő hozzáférés a gyakorlatban. *Statisztikai Szemle*. 93. évf. 11–12. sz. 1170–1176. old.
- PETRI, E. [2014]: *Integration of Statistical and Geospatial Information – An Overview of European and Global Initiatives*. “European Forum for Geography and Statistics” Conference. 22–24 October. Krakow.
- STATISTICS AUSTRIA [2011]: *Target Record Swapping. Version 2.0*. Wien. www.statistik.at/web_de/static/record_swapping_063...

- TAMMILEHTO-LOUDE, M. [2011]: *Opportunities and Challenges of Grid-based Statistics*. 58th World Statistical Congress. 21–26 August. Dublin. <http://2011.isiproceedings.org/papers/650400.pdf>
- VERECZKEI Z. [2015]: Az adat-hozzáférési igények és az adatvédelmi kötelezettségek egyensúlyáról. *Statisztikai Szemle*. 93. évf. 11–12. sz. 1089–1111. old.
- WILLENBORG, L. – DE WAAL, T. [2001]: *Elements of Statistical Disclosure Control*. Springer. New York.

Summary

Developing and applying statistical disclosure control (SDC) techniques on grid-based statistics are considered as major challenges in many countries. Appropriate SDC methods are needed also for the protection of grid-based official statistical information against identification or disclosure from data released.

This paper describes the development and the first application of a grid-based statistical disclosure control method on a Hungarian grid-based dataset that can be applied to many official (social, business or other) statistical surveys.

This new method was first introduced on the 2011 Hungarian Population and Housing Census data presented in 1 km² grids, broken down into age groups and sex. Taking into consideration the data protection regulations of the Hungarian Central Statistical Office and the peculiarities of its data, the author presents how a suitable SDC method was developed, completed and adapted in practice.

Mátyás-Bodovics Édina,
a Központi Statisztikai Hivatal
osztályvezetője
E-mail: Edina.Matyas@ksh.hu

A Központi Statisztikai Hivatalban kezelt statisztikai adatok gyakorlati hozzáférése

A KSH¹ arra törekszik, hogy hatalmas adatvagyona lehetőleg minél szélesebb körben felhasználható legyen. Mivel a statisztikai adatok egyben közérdekű adatok, vagyis főszabály szerint nyilvánosak, így bárki számára hozzáférhető és megismerhető. Erre az információs önrendelkezési jogról és az információszabadságról szóló 2011. évi CXII. törvény is kötelezi a KSH-t, mely kötelezettségének egyrészt közzététellel, vagyis a közérdekű statisztikai adatok egy részének külön kérés vagy igény nélkül, a KSH a honlapján, illetve a nyomtatott kiadványokban történő publikálásával, másrészt pedig a statisztikai adatok megismerésére megfogalmazott felhasználói igények teljesítésével tesz eleget.

1. Az adatigények teljesítésének szervezeti kerete

A benyújtott igények teljesítését a KSH információszolgáltatási feladatokat ellátó egysége, az Információszolgálat koordinálja, amely tevékenysége során kiemelt figyelmet fordít a gyors és magas szakmai színvonalú, minőségi szolgáltatás biztosítására. A minőségre való törekvés a jól szabályozott, egységes elvek mentén felépített, standard munkafolyamatokban, illetve abban a szakmai tudásban és elkötelezettségben nyilvánul meg, amely során az információt szolgáltató ügyintéző a lehető legteljesebb körű tájékoztatást igyekszik nyújtani, az igényelt adatok legjobb minőségben és részletezettségben történő megadásával, az adatvédelmi szempontok betartása mellett.

¹ KSH: Központi Statisztikai Hivatal.

2. Az adatigények kezelése

Az igények kezelésének menete és teljesítésének feltételei attól függenek, hogy az adatkérő által megfogalmazott igény teljesíthető-e a nyilvánosan hozzáférhető adatokból, vagy kizárólag a KSH belső rendszereiben tárolt adatok alapján állítható elő.

2.1. Nyilvánosan hozzáférhető adatok

A KSH a honlapján évről évre egyre több statisztikai adatot publikál, melyeket felhasználóbarát megoldásokkal hatékonyan és gyorsan kereshető formában tesz hozzáférhetővé a felhasználók számára. Ezért az adatkérőnek – mielőtt az Információs Szolgálathoz fordulna – érdemes előzetesen tájékozódnia a publikált adatokat illetően. Azon túl ugyanis, hogy az adatok honlapról való letöltése biztosítja számára a leggyorsabb adathozzáférést, a honlapon elérhető, jól strukturált, tematikusan elhelyezett információk segítségével akár további, számára hasznos adatot is fellelhet a keresett témában. A nyilvános adatokon túl, a honlapon elérhető, szakstatisztikákra vonatkozó információk, illetve az OSAP²-tervek/kérdőívek áttekintése is segítséget nyújthat a végső adatigény kialakításában.

A honlaptartalom jól strukturált és tematikus felépítése ellenére havonta több száz olyan igény érkezik az Információs Szolgálat munkatársaihoz telefonon, személyes ügyfélfogadás keretében, illetve elektronikus úton, amelyek nyilvánosan elérhető adatra vagy információra vonatkoznak. Személyes és telefonos ügyfélszolgálat keretében az igény teljesítése során tényleges adatkiadás nem történik, az ügyintéző csak arról tájékoztatja az ügyfelet, hogy a keresett adatot vagy információt hol és milyen formában érheti el.

Amennyiben elektronikus úton érkezik a megkeresés, a választ is ebben formában küldik meg, amely minimálisan a honlap megfelelő adatot tartalmazó oldalára és a kapcsolódó információkra mutató linkek megküldését jelenti, de a válasz kiegészülhet még egyéb, a témához kapcsolódó releváns információkkal is.

2.2. Nyilvánosan nem hozzáférhető adatok

Az igények egy része azonban olyan statisztikai adatokra irányul, amelyeket a KSH korábban semmilyen formában nem publikált, vagyis azok csak a belső rendszereiben érhetők el. Az ilyen igényeket kizárólag írásos formában lehet a KSH-nak

² OSAP: Országos Statisztikai Adatgyűjtési Program.

benyújtani a honlapon elérhető „Lépjön velünk kapcsolatba!” alkalmazáson keresztül, az „Adatkérés” ügycategóriát kiválasztva. Az, hogy a kért adatok nyilvánosságra hozatala végül hogyan és milyen feltételekkel teljesül, az mindig az igényelt adatok részletezettségétől, a hozzáférési módtól és egyéb jellemzőitől függ, vagyis attól, hogy az igényt milyen mértékű adatvédelmi eljárásoknak kell alávetni. Ennek megfelelően a nyilvánosan nem hozzáférhető adatokra irányuló igények teljesítése a Hivatal által működtetett hat különböző adathozzáférési csatorna valamelyikén valósulhat meg a felhasználási cél függvényében.

2.2.1. Adathozzáférési csatornák³

a) Az általános célú adathozzáférési csatornák az általános céllal felhasználható, bárki számára elérhető, így erős felfedési megoldásokkal kezelt, végső formájában lényegében minimalizált felfedési kockázattal rendelkező adatállományok közvetítő-eszközei. Ide tartoznak az egyedi kérésre összeállított táblázatos adatok felhasználóknak történő átadása, illetve a nyilvános mikroadatfájlok elérhetővé tétele. Az, hogy az ilyen módokon hozzáfért adatokat illetően a felhasználó milyen felhasználási jogosultsággal rendelkezik, a hozzáféréssel egy időben elfogadott felhasználási feltételekben szabályozza a KSH.

b) A tudományos célú adathozzáférési csatornák az adathozzáférési csatornák azon része, mely kifejezetten a statisztikai adatok tudományos célú felhasználásra hivatott, amelyen keresztül kizárólag tudományos kutatások céljára összeállított, az egyes adathozzáférési csatornáknak megfelelő adatvédelmi megoldásokkal kezelt mikroadat-állományok hozzáférése biztosított. Ez utóbbi csoportba tartoznak azok az informatikailag biztonságos környezetben történő adathozzáférések, ahol az egyedi adatok kiemelt védelme, az adatvédelmi szabályok betartása, valamint megfelelő jogi garanciák mellett dolgozhatnak a kutatók a többi adathozzáférési csatornához mérten részletesebb, jellemzően csupán közvetlen azonosításra alkalmatlanná tett mikroadat-állományokon.

Tudományos célú adathozzáférési csatornának tekintjük a kutatószobai, a távoli hozzáférés, illetve a távoli végrehajtás adathozzáférési csatornáit, valamint az ugyancsak kutatási igények kiszolgálását hivatott anonimizált mikroadat-kiadást biztosító csatornát. Utóbbi esetben az intézményi háttérrel rendelkező kutatók olyan anonimizált mikroadatokhoz férhetnek hozzá, melyek a biztonságos környezetben elérhető állományokhoz viszonyítva erősebb felfedés elleni védelemmel vannak ellátva. Bár az erősebb felfedés elleni védelem miatt nem szükséges a biztonságos környezetben történő hozzáférés, a KSH ebben az esetben is kiemelt figyelmet fordít az adatok teljes védelmére a megfelelő jogi garanciák biztosításával, illetve a kutatói

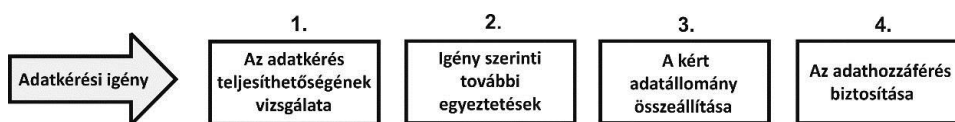
³ Lásd: <http://www.ksh.hu/adatigenyles>

akkreditációval. A különböző adathozzáférési csatornák tehát különböző szintű jogi, módszertani és informatikai garanciák mellett érhetők el. A beérkező adatigények kezelésének folyamatát az is meghatározza, hogy az adott igény mely adathozzáférési csatornán teljesíthető, illetve – amennyiben több adathozzáférési csatorna is rendelkezésre áll a teljesítéshez – az adatkérő melyik adathozzáférési csatornát preferálja.

2.2.2. KSH belső rendszereiben elérhető adatokból teljesíthető igények ügymenete

Az ábra annak a folyamatnak a főbb állomásait mutatja, amely során a KSH belső rendszereiből történő leválogatás/adat-összeállítás eredményeként teljesül az adatigény.

Az adatkérési igény benyújtása és fogadása



Az adatkérő a statisztikai adat megismerési igényét kizárólag a megfelelő adathozzáférési csatornára vonatkozó igénybejelentő adatlapon nyújthatja be, ennek hiányában ugyanis a KSH nem fogadja be a megkeresést. A különböző adathozzáférési csatornák elérésének feltételeiről és a benyújtandó igénylőlapokról részletes tájékoztató érhető el a KSH honlapján (<http://www.ksh.hu/adatigenyles>), melyet az adatkérési igény benyújtását megelőzően az adatkérőnek célszerű alaposan áttanulmányoznia.

Amennyiben az adatkérő számára nem teljesen egyértelmű, hogy igényét mely csatornára vonatkozóan kell benyújtania, akkor a „Lépjön velünk kapcsolatba!” alkalmazás felületén található „Adatkérés” ügykategória űrlapján keresztül küldheti azt meg, majd az adatkérés teljesítésének formájára vonatkozó egyeztetést követően tudja a szükséges dokumentációt pótolni.

Az ügyintézés hatékonyságához kiemelten fontos, hogy az adatot igénylő minél pontosabban határozza meg kérése paramétereit. A lehető legjobban szűkítse le az adott témakörön belül a keresett információt, adja meg a pontos időszakot és azt, hogy milyen bontásban kéri az adatokat. Ezek az információk jelentősen meggyorsíthatják a válaszadást, mivel egy jól körülírt adatkérés esetében nem szükséges további egyeztető kérdések megküldése.

Az adatkérés teljesíthetőségének vizsgálata. A pontos adatigény birtokában a teljesíthetőség vizsgálata két alappillérre támaszkodik: az adatkérés szakmai, illetve adatvédelmi teljesíthetőségére. A szakmai vizsgálat során az Információs Szolgálat megvizsgálja, hogy a kért adat rendelkezésre áll-e, esetleg előállítható-e a KSH belső

rendszereiben tárolt adatok alapján. Amennyiben az adatigény teljesíthető ugyan, de a nyilvánosságra hozandó állománnyal kapcsolatban minőségi kifogások emelhetők, úgy a minőségi korlátokról megfelelő részletezettségben, teljes körű tájékoztatást kap az adatkérő, aki a kapott információk birtokában mérlegelheti, hogy továbbra is kitarthat-e az igénye mellett vagy eláll tőle.

Azokban az esetekben, amikor az igény azért nem teljesíthető, mert a kért adattal nem rendelkezik a KSH és nem is állítható elő a belső rendszerekből, az adatkérő tájékoztatása kiter arra is, hogy a keresett témáról vagy adatról hol és milyen formában található még érdemi információt. Ez lehet a témához szorosan kapcsolódó egyéb, a KSH belső rendszereiből teljesíthető adatokról történő tájékoztatás, vagy akár az adatkérőnek valamely más, a hivatalos statisztikai szolgálathoz tartozó adatgazdászervezethez történő irányítása. Statisztikai adatokat ugyanis nemcsak a KSH gyűjt és kezel, hanem a szolgálat részeként, a statisztikáról szóló 1993. évi XLVI. törvényben rögzített egyéb szervek is.

A KSH lehetőséget ad arra is, hogy egyedi statisztikai szolgáltatásként akár célzott felvételek végrehajtásával állítsa elő azokat az adatokat, melyekre az adatkérőnek igénye van. A szolgáltatás iránti igényét az adatkérőnek szintén az Információszolgálathoz kell benyújtania.

A teljesíthetőség ugyanakkor nemcsak az adat elérhetőségének függvénye, hanem a kiadhatóság adatvédelmi szempontjainak is. Vagyis, amennyiben az igény olyan adatokra irányul, melyek nyilvánosságra hozatalával adatvédelmi aggályok merülnének fel, a teljesíthetőség vizsgálata kiegészül az adatigény teljesítési folyamatába integrált adatvédelmi vizsgálattal is, amely jogi és módszertani szempontból elemzi az adatigényt és határozza meg a felfedés elleni védelmi eljárásokat, illetve a megfelelő adathozzáférési csatornába történő terelést. A vizsgálat célja az, hogy a KSH a törvényi kötelezettségeinek megfelelően, biztosítsa azt a legfontosabb alapelvet a statisztikai adatokra irányuló igények teljesítése során, hogy egyedi adatokat nem hoz nyilvánosságra, illetve minimalizálja azt a kockázatot, hogy a statisztikai céllal kezelt adatok megismerésén keresztül egyedekre vonatkozóan (például háztartásokra, vállalkozásokra, üzletekre stb.) új, eddig nem ismert egyedi szintű információ kerüljön nyilvánosságra.

Ugyanakkor néhány esetben, szigorú szabályok betartása mellett – összhangban az európai statisztikákról szóló 223/2009/EK európai parlamenti és tanácsi rendelettel, az információs önrendelkezési jogról és az információszabadságról szóló 2011. évi CXII. törvénnyel, valamint a statisztikáról szóló 1993. évi XLVI. törvénnyel, és a Központi Statisztikai Hivatal adatvédelmi politikájával – a KSH hozzáférést biztosít egyedi adatokhoz is.⁴ Amennyiben az adatvédelmi vizsgálat végül kimondja, hogy az

⁴ NAGY E. [2015]: A statisztikai adatvédelem és -hozzáférés szabályai az Európai Unióban. *Statisztikai Szemle*. 93. évf. 11–12. sz. 1051–1069. old., illetve LAKATOS M. [2015]: Az adatvédelem és a statisztika kapcsolatának jogi szabályozása. *Statisztikai Szemle*. 93. évf. 11–12. sz. 1017–1050. old.

adatkérés adatvédelmi okok miatt nem teljesülhet, további egyeztetések szükségesek az adatkérővel abból a célból, hogy az adatkérés módosításával, vagy adathozzáférés megfelelő csatornájának kiválasztásával az igény teljesíthetővé váljék.

Igény szerinti további egyeztetések. A teljesíthetőséget illetően tehát szakmai és adatvédelmi jellegű problémák merülhetnek fel egy-egy adatigény alkalmából. Természetesen ez nem azt jelenti, hogy az adatkérés egyáltalán nem teljesíthető, sokkal inkább azt, hogy a kért adattartalommal, vagy a kért bontásban adatvédelmi okokból nem adhatók át az adatok. Ilyen esetekben két lehetséges opció merülhet fel. Az egyik az egyeztetések eredményeképpen az adatkérés tartalmának olyan módosítása, amely során az összeállítandó adatállomány nyilvánosságra hozatala adatvédelmi aggályt nem vet fel. A másik megoldás során – az adatkérővel egyeztetve – a kért adatállomány kutatásra előkészített formában kerül előállításra, így a teljesíthetőség nem az adattartalom megváltoztatását, hanem a tudományos célú adathozzáférések számára nyitott, és az adatkérő számára a legalkalmasabb adathozzáférési csatornába történő terelését jelenti.

A tudományos célú adatigények teljesítése során is alapvető szempont, hogy az adatkérő által megfogalmazott igény és a statisztikai adatokhoz történő biztonságos hozzáférést biztosító adatvédelmi megoldások optimális egyensúlyba kerüljenek.⁵

A statisztikai adatok tudományos célú felhasználása szigorú feltételek mellett történik. Amennyiben az adatkérő elfogadja azt, hogy igénye kizárólag biztonságos környezetben, megfelelő garanciák mellett teljesíthető, akkor igényét a megfelelő adathozzáférési csatornához rendelt igénybejelentő adatlapon kell bejelentenie, melyben a részletes kutatási cél mellett közölnie kell magára mint kutatóra, valamint az esetleges kutatói intézményi háttérre vonatkozó információkat is.

A tudományos kutatásokat megelőzően ugyanis a KSH úgynevezett kutatói akkreditációt végez, melynek keretében vizsgálja a kutató és – amennyiben van – a kutatással foglalkozó intézményi háttér kutatói szempontoknak való megfelelését. Szintén a vizsgálat tárgya az adatkérés szakmai tartalmának vizsgálata (annak vizsgálata, hogy az adatkérés teljesíthető-e, tehát a KSH rendelkezik-e a kért információval) és a kutatás tudományos céljának elbírálása kötött, nyilvánosan is hozzáférhető szempontrendszer szerint. A kutatás engedélyezésének feltétele a sikeres kutatói akkreditáció. A kutatói akkreditáció folyamatáról és szempontrendszeréről a honlapon részletes tájékoztatót érnek el a kutatók.⁶

A kért adatállomány összeállítása. A véglegesen kialakított adatigény ismeretében a kért adatállomány az adathozzáférési csatornától függően kerül összeállításra. Az egyedi kérésre összeállított táblázatos adatok aggregálva, felfedés elleni védelemmel

⁵ VERECZKEI Z. [2015]: Adat-hozzáférési igények és adatvédelmi kötelezettségek egyensúlyáról. *Statisztikai Szemle*. 93. évf. 11–12. sz. 1089–1111. old.

⁶ Részletesen lásd: http://www.ksh.hu/kutato_i_akkreditacio

ellátva; anonimizált mikroadat-kiadás esetén megfelelően anonimizálva, míg biztonságos környezetben a kutatásra előkészített mikroadat-állományok esetében közvetlen azonosításra alkalmatlan formában. Minden átadásra vagy kutatásra előkészített adatállományhoz elkészül a változókkal kapcsolatos információkat, illetve az állomány egészére vonatkozó metaadatokat tartalmazó leírás is, melyet az adathozzáféréssel egy időben kap meg az adatkérő.

Az adathozzáférés biztosítása. Az adott adathozzáférési csatorna nemcsak azt határozza meg, hogy milyen a módszertani, adatvédelmi szintje, hanem azt is, hogy az adatállomány milyen jogi és informatikai biztonsági garanciák mellett használható.

A táblázatos adatállományok átadásával egy időben kerül megküldésre a „felhasználói feltételek” dokumentuma is, amely rögzíti a felhasználó jogait és kötelezettségeit a kapott állomány további kezelését illetően.

A biztonságos környezetben történő kutatásra előkészített mikroadat-állományokhoz való hozzáférés erős jogi korlátok mellett lehetséges, ezért az adott kutatáshoz kapcsolódóan az adathozzáférés feltételeit, a kutató és a KSH jogait és kötelezettségeit a KSH és a kutató közötti kutatási szerződésben, illetve a szerződés mellékletében szereplő titoktartási nyilatkozatban szükséges rögzíteni.

A kutató ezt követően a szerződésben rögzített biztonságos környezetbeli adathozzáférési csatornán érheti el a közvetlen azonosításra alkalmatlan mikroadat-állományokat. A KSH a kutatói munka során teljes szakmai, módszertani és informatikai támogatást biztosít a kutató részére.

A kutató a kutatási cél szem előtt tartásával szabadon használhatja a rendelkezésére bocsátott adatállományt, azokból igény szerinti kutatási eredményeket hozhat létre. A kutatási eredményeket azonban csak akkor viheti ki a biztonságos környezetből, ha azok a kiadást megelőző adatvédelmi ellenőrzésen átesnek.

*

A felhasználók statisztikai adatokba vetett bizalmát mi sem bizonyítja jobban, mint hogy évről évre egyre több megkeresés érkezik telefonon, illetve írásos formában a KSH-hoz. 2014-ben közel 27 ezer igényt kezelt az Információs Szolgálat, amelyből 46 esetben biztonságos környezetben történt az adathozzáférés. Ez az utóbbi szám 2015. október végén már 50 folyamatban lévő kutatást, illetve további négy elbírálás alatt lévő igényt jelentett.

Alexin Zoltán

PhD, a Szegedi Tudományegyetem egyetemi adjunktusa

E-mail: alexin@in.u-szeged.hu

A statisztikai táblázatokban található kis előfordulási számokkal kapcsolatos angol bírósági ítélet

A személyes adatok védelméhez és az információszabadsághoz való jog más alapvető jogokkal összehasonlítva nagyon fiatal. Az Európai Unióban általánosan csak 2009. december 1-jétől van érvényben, amikor hatályba lépett az Európai Unió Alapjogi Chartája.¹ Az egyes tagállamokban nem alapvető jogként már korábban is létezett, de két-három évtizednél nagyobb múltra így sem tekinthet vissza. A fogalmak és definíciók finomítása fontos kérdés, és ebben nagy szerepe van az alapvető kérdések tisztázására irányuló próbapereknek. Az ilyen ügyek jelentőségét nem a perben vitatott sérelem adja, hanem az, hogy az ítélet eddig megválaszolatlan, alapvető jogi kérdést tesz tisztába, esetenként radikálisan megváltoztatva a korábban elterjedt közfelfogást.

Az alapvető jogok értelmezésével kapcsolatos perek viteléhez komoly szaktudás szükséges. Először is azért, mert meg kell találni a megfelelő jogesetet, vagy esetleg mesterségesen létre kell hozni, majd kiválasztani azt a jogi lehetőséget, hogy ebből egy pert lehessen kreálni. Az ilyen ügyek nem érnek véget a rendes bíróságokon, hanem el kell vinni őket a fellebbezési lehetőségek végső határáig, az alkotmánybíróságig vagy valamely nemzetközi bíróságig. Sokszor a pereket polgárjogi aktivizmusnak állítják be, de valójában jó nevű nemzetközi jogászok, esetenként egyetemi tanárok állnak az adott ügy mögé felperesnek, és egész életművüket, tudományos munkásságukat ütköztetik a napi ítélkezéssel, jogalkalmazással. *Richardson és szerzőtársai (Richardson et al. [2012])* könyvükben olyan peres ügyeket ismertetnek, amelyeket a XIX. század végén indítottak személyes adatok védelme körében elszenvedett sérelmek ügyében Nagy-Britanniában vagy gyarmatain. Természetesen akkor még a személyes adat fogalma nem létezett, ezért a szokásos polgári jog alapelvei, például a kártérítéssel kapcsolatos jog alapján építették fel a pert. Az okozott kár kötelező megtérítésének elve a római jogból ered, és a mai napig része a polgári jognak.

¹ <http://eur-lex.europa.eu/legal-content/HU/TXT/?uri=CELEX:12012P/TXT>

Napjainkban az Európai Unió egyes tagállamainak statisztikai hivatalai kiterjedt adatgyűjtést folytatnak, amelynek során személyes, sőt időnként szenzitív információkat is összegyűjtenek (például a magyarországi népszámlálások a fogyatékosokra vonatkozó adatok esetében). Nemzeteken átnyúló szakmai egyetértés van abban, hogy a hivatalok által publikált adatoknak olyanoknak kell lenniük, hogy ne tegyék lehetővé azoknak a személyeknek az azonosítását, akikre vonatkoznak. A kétezres években több olyan eset került napvilágra, amikor ártalmatlannak látszó publikált adatok alapján sikerült konkrét személyeket azonosítani (Alexin [2014]). Az Egyesült Államok Szabványügyi Hivatala 2015 októberében elérhetővé tett egy, a körültekintő anonimizálásra vonatkozó szabványt (Garfinkel [2015]).

Az azonosítható személyes adatokat is tartalmazó statisztikai táblázatokban a kis értékeket nem tanácsos publikálni. A kis értékeket tartalmazó cellák üresen hagyása az egyik eszköz arra, hogy meggátoljuk az adatok mögött található személyek azonosítását, de ezen kívül még több, alternatív eszköz is van (például a kvantálás, barnardizáció, kerekítés stb.). Az Egyesült Államok Szövetségi Statisztikai Módszertani Bizottságának 22-es számú munkanyagában ezek ajánlásként szerepelnek.² Az Egyesült Királyság Nemzeti Statisztikai Hivatala először 2005-ben adott ki egy útmutatót az angol abortuszstatisztika adatvédelmi kérdéseivel kapcsolatban, amit aztán 2011-ben az ismertetésre kerülő ítélet nyomán megújítottak.³

A következőkben azt a jogi esetet ismertetjük, amellyel kapcsolatban az Egyesült Királyság és Wales közigazgatási ügyekben illetékes táblabíróságán született jogerős bírói döntés. A perben az igen érzékeny személyes adatokból kapott, aggregált statisztikai adatokról mondta ki a bíróság, hogy azok szabadon publikálhatók, annak ellenére, hogy ezt az Egészségügyi Minisztérium kockázatosnak tartotta.

Az Egyesült Királyság és Wales abortuszról szóló törvénye⁴ alapján az elvégzett abortuszokról jelentést kell írnia a végrehajtó orvosnak, és azt elküldenie a tisztifőorvosnak egy nyomtatványon⁵. Az évente elvégzett beavatkozásokról minden évben statisztikai jelentést készít a Nemzeti Statisztikai Hivatal. Abban az esetben, ha a terhesség megszakítására azért került sor, mert a magzat súlyos fejlődési rendellenességet mutatott, akkor a nyomtatványon meg kellett adni ezt a rendellenességet. Az egyes rendellenességek részletes előfordulási adatait az éves jelentés tartalmazta (lásd például a 2002-es jelentés 17. oldalán)⁶. Jogszabályváltozás miatt 2003-tól az éves jelentést már az Egészségügyi Minisztérium készítette el, azonban úgy találták, hogy a részletes, a különböző rendellenességeket és az előfordulásuk számát tartalmazó táblá-

² <http://fcs.m.sites.usa.gov/files/2014/04/spwp22.pdf>

³ <http://www.ons.gov.uk/ons/guide-method/best-practice/disclosure-control-of-health-statistics/confidentiality-guidance.pdf>

⁴ Abortion Act 1967, <http://www.legislation.gov.uk/ukpga/1967/87/contents>

⁵ https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/204074/HSA4-form.pdf

⁶ <http://webarchive.nationalarchives.gov.uk/20040726025155/http://publications.doh.gov.uk/public/sb0323.pdf>

zat publikálása veszélyes lehet, és ezért ezt kihagyták a jelentésből. A táblázatban az előfordulási számok több esetben nem érték el a tízet. A legfrissebb, 2014-es statisztika elérhető az interneten,⁷ egy Excel állományban található a táblázatok.

Az eset előzményeit az angol Adatvédelmi Hivatal weboldalán található határozat (ICO [2008]) és a jogerős bírósági ítélet⁸ alapján foglaljuk össze.

A ProLife Alliance egy olyan civil szervezet, amely figyelemmel kíséri a terhességmegszakításokat, különösen a 24 hetesnél idősebb magzatok esetén.⁹ A 2001-es táblázatban szerepelt egy ajak- és szájpád hasadék fejlődési rendellenesség miatt megszakított 24 hetesnél idősebb terhesség. A társaság egy tagja haladéktalanul rendőrségi feljelentést tett a londoni rendőrkapitányságon illegális abortusz gyanúja miatt. A rendőrség megkapta az tisztifőorvostól a kitöltött nyomtatványt, és elkövette azt az indiszkréciót, hogy nyilvánosságra hozta azt, hogy melyik illetékes vidéki rendőrkapitánysághoz továbbították az ügyet. Ott viszont csak egyetlen kórház volt a közelben, és rövid időn belül az újságírók meg is találták az abortuszt elvégző orvost. Addig zaklatták folyamatosan, amíg be nem ismerte. A rendőrség később bűncselekmény hiányában ejtette az ügyet. Az abortuszstatisztika alapján az Egészségügyi Minisztérium maga is végzett helyszíni ellenőrzéseket egy-egy esetben, de szabálytalanságot nem találtak.

Mivel sem a 2003-as, sem a 2004-es évre készült statisztika nem tartalmazta a fejlődési rendellenességek részletes táblázatát, csak a rendellenességek együttes előfordulásának számát, ezért 2005. február 21-én az információs szabadság törvényre¹⁰ hivatkozva a ProLife Alliance kérte a korábbi formában a részletes adatokat az Egészségügyi Minisztériumtól. A minisztérium válaszában a publikálás veszélyességére hivatkozva megtagadta a táblázat elküldését. A ProLife Alliance tett még egy próbálkozást május 24-én és a véleményének felülvizsgálatára kérte fel a minisztériumot. A minisztérium időközben a Nemzeti Statisztikai Hivatal segítségét és egy szakmai útmutató elkészítését kérte, amely támpontot adna számára a statisztika elkészítéséhez. Az útmutató 2005 júliusában készült el. A minisztérium pedig egy részletesebb táblázatot állított össze ennek alapján, és azt publikálta is, de a 10-nél kisebb számokra cellaelnyomást alkalmazott. A panaszos 2006 januárban az adatvédelmi hatósághoz fordult, panaszt tett az Egészségügyi Minisztérium ellen, amiért az nem folytatta le a felülvizsgálatot, és nem adta ki neki az adatokat. A biztos kérte, hogy a panaszos várja meg a felülvizsgálat hivatalos eredményét. A minisztérium 2006. április 7-én kelt levelében adott tájékoztatást a felülvizsgálatról és elutasította a ProLife kérését, azzal, hogy egy kissé részletesebb táblázatot már közzé tett. Május-

⁷ <https://www.gov.uk/government/statistics/report-on-abortion-statistics-in-england-and-wales-for-2014>

⁸ http://www.informationtribunal.gov.uk/DBFiles/Appeal/i344/CO-13544-2009_HC_Judgment_20110420.pdf

⁹ A jogi szabályozás szerint, ilyen idős magzatok életét már csak akkor lehet elvenni, ha súlyos fejlődési rendellenesség mutatható ki náluk. Az ajak- és szájpádasadék sebészetiileg korrigálható, és így nem összeegyeztethetetlen az élettel.

¹⁰ Freedom of Information Act 2000, <http://www.legislation.gov.uk/ukpga/2000/36/contents>

ban a ProLife ismét az adatvédelmi hatósághoz fordult és arra kérte, hogy kötelezze a minisztériumot minden kért adat kiadására.

Az Adatvédelmi Hivatal határozata 2008. július 28-án született meg (*ICO* [2008]), amelyben helyt adott a ProLife kérésnek és kötelezte a minisztériumot a teljes táblázat közzétételére. A határozat ellen a minisztérium első fokon az Információs Közigazgatási Bírósághoz fordult, ez ismét a biztosnak adott igazat, és fellebbezés után került az ügy a Londoni Közigazgatási Táblabíróságra. Itt már jogerős ítélet született, amely ismét az adatvédelmi biztosnak adott igazat. További fellebbezés nem történt, a felek a döntésben megnyugodtak és elfogadták.

A Nemzeti Statisztikai Hivatal 2005-ös útmutatója szerint a nem biztonságos cellák az abortuszstatisztikában azok, amelyek értéke zéró, kivéve, ha logikailag más értéke is lehetséges; kisebb mint 5 és egy régióra (vagy Walesre) vonatkozik; kisebb mint 10 és egy tetszőleges, de a régiónál kisebb földrajzi területre vonatkozik; kisebb mint 10, ha különösen szenzitív adatra vonatkozik; kisebb mint 10, ha egy vagy két orvosra, illetve egy vagy két kórházra vonatkozik. A különösen szenzitív változók: a fiatal életkor (15 évesnél fiatalabb), a terhesség kora 24 hetesnél idősebb, az orvosi diagnózisok, a terhességmegszakítás módja a terhesség kora szerint. A nem biztonságos cellák esetén cellaelnyomást ajánlott alkalmazni.

Az előzményekhez tartozik egy másik ügy is: a skót Adatvédelmi Hivatal és az Egészségügyi Szolgáltatások Tervezésének Hivatala közötti per.¹¹ Ennek az ügynek a tárgya az volt, hogy egy parlamenti képviselő a hivataltól, az információs szabadság törvényre hivatkozva, kérte a választókerületében újonnan felfedezett gyermek (0-tól 14 éves korig) leukémiás betegek számát a 2002. és 2003. évekre vonatkozóan, életkor és nem szerinti bontásban. A közelben atomerőmű és használt fűtőelem feldolgozó üzem működik, amelyek egészségügyi kockázatot jelenthetnek. A kérést az iroda elutasította, hivatkozva az indirekt azonosíthatóság kockázatára, ezért a kérelmező a skót adatvédelmi hatósághoz fordult. Az Adatvédelmi Hivatal elismerte az adatok kockázatosságát, ezért az ún. barnardizáció elvégzésére kötelezte az irodát. A módszer lényege, hogy a kis számokat tartalmazó cellákat perturbálják: a zéró marad zéró, ha az érték 1, akkor véletlen módon 0-t vagy 1-et; ha 2–4 közötti akkor 0-t, +1-et vagy –1-et hozzáadnak. Ilyen módon a pontos adatokat a kérelmező nem ismerheti meg, de a tendencia így is megfigyelhető. A határozat ellen a hivatal fellebbezett arra hivatkozva, hogy a barnardizált adat továbbra is alkalmas lehet konkrét személyek azonosítására. Az ügy a Legfelsőbb Bíróságig jutott el, de a bírói tanács nem jutott egyezsége. Az ítéletet az elnök szava döntötte el, ugyanakkor a helyi szabályok szerint ilyenkor az nem kötelező. Az adatvédelmi hatóságot kötelezték egy megismételt eljárásra, amelyben az állapítsa meg, hogy a kért barnardizált adatok végül is személyes adatoknak tekinthetők-e. Ha az eredmény nem, akkor az adatokat ki kell adni a képviselőnek.

¹¹ <http://www.publications.parliament.uk/pa/ld200708/ldjudgmt/jd080709/comm.pdf>

Az Egyesült Királyság Adatvédelmi Hivatala hosszú idő után hozta meg a határozatát (majdnem két év elteltével) az abortuszstatisztikás esettel kapcsolatban. Megvárták, amíg a skót parlamenti képviselő ügye a Legfelsőbb Bíróságon véget ér. Az adatvédelmi biztos hangsúlyosan vette figyelembe az Nemzeti Statisztikai Hivatal által kibocsátott útmutatót, és alkalmazta azt a konkrét adatkérésre. Kötelezte az Egészségügyi Minisztériumot a teljes abortuszstatisztika átadására a ProLife Alliance számára cellaelnyomás nélkül.

Az Egészségügyi Minisztérium nem fogadta el ezt a döntést, és keresetet nyújtott be ellene az Információs Közigazgatási Bírósághoz. A minisztérium vitatta azt, hogy az adatok nem lennének személyazonosításra alkalmasak, a különösen szenzitív cellák esetében pedig hivatkozott a Nemzeti Statisztikai Hivatal útmutatójára, amely a különösen szenzitív cellák elnyomását ajánlotta. A minisztérium egy másik incidensre is hivatkozott, egy kilenc éves lány esetére. Itt arról volt szó, hogy újságcikk jelent meg arról, hogy abortuszt hajtottak végre egy kilenc éves lányon, akit meg is nevezett a cikk, azonban itt az újságíró az értesülését nem az abortuszstatisztikából szerezte. Az viszont igaz volt, hogy utóbb a statisztikát elővéve egyértelműen azonosítható volt benne. A biztos ezt sem értékelte úgy, hogy ez bizonyíték lett volna arra, hogy a kért adatok személyazonosításra alkalmasak lennének.

Az adatvédelmi biztos az útmutatót vette szintén alapul, és azzal érvelt, hogy a kért összesített adatok a teljes Egyesült Királyságra vonatkoznak (a régiónál nagyobb földrajzi egységre) – így az útmutató szerint nem tekinthetők nem biztonságos célaknak. A népszámlálási adatok szerint több mint 14,17 millió 10–49 éves angol nőre vonatkoznak az adatok. Az országban 420 olyan kórház van, ahol lehetőség van abortusz végzésére. Az útmutató alapján, tehát nincs szükség cellaelnyomásra.

Az angol adatvédelmi biztos azzal is érvelt, hogy a közérdekű adatok nyilvánosságra hozásához fontos érdekek fűződnek: az abortusztörvénynek történő megfelelés ellenőrzése; a közvélemény tájékoztatása arról, hogy hogyan alkalmazzák a törvényt; a szülésorvosok elszámoltathatóságának biztosítása; az Egészségügyi Minisztérium tevékenységének külső ellenőrzése; változások, trendek felismerése; egészségügyi tervezés, a fejlődési rendellenességek arányának követése. Ezek mind olyan érdekek, amelyek szembe állíthatók azzal a másik érdekkal, amely védi az abortuszon átesettek identitását.

A perben sok egyéb szempont is felmerült: például, hogy az Európai Parlament és a Tanács 95/46/EK adatvédelmi irányelvben¹² szereplő személyes adat definíció nem egyezik meg az angol adatvédelmi törvényben szereplővel; a Nemzeti Statisztikai Hivatal útmutatója nem jogi előírás, ezért jogilag nem kötelező azt alkalmazni; az abortusztörvény nem engedi meg személyes adatok közlését a beküldött nyomtatványokról; a statisztikai adatok azon a módon, hogy temérdek más adattal együtt,

¹² <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:hu:HTML>

azokkal összekapcsolva elvezethetnek egyes személyek azonosításához. Az ügy megoldása végül nem az volt, hogy a rendkívül kicsiny újra-azonosítási kockázatot olyan akadálnak tekintik, amely a közérdekű adatokhoz történő hozzáférést lehetlenné teszi. A kockázatok és a társadalmi haszon egyensúlya alapján döntött már az elsőfokú bíróság is, amelyet később teljes mértékben helyben hagyott a másodfokú táblabíróság is.

Az Egyesült Királyság Egészségügyi Minisztériuma kellő elszántsággal védi a személyekre vonatkozó egészségügyi adatait, és ha úgy látja, hogy a nyilvánosságra hozandó táblázatokban szereplő egészségügyi adatok legkisebb mértékben is azonosíthatók, akkor még bíróság előtt is védi az igazát.

Az itt bemutatott ítélet mérőföldkő lehet a további fejlődés irányának kijelölésekor. Miközben folyamatos kétségeink vannak abban a tekintetben, hogy a statisztikai táblázatokban szereplő kis számok mögötti személyek esetleg azonosíthatók, vagy a jövőben azonosíthatók lesznek, a mérleg másik serpenyőjében is fontos társadalmi célok vannak: az információszabadság, a jövőbeli egészségi kilátásaink, a káros jelenségek időben történő észlelése stb. Az ellentmondás az adatvédelem és az információszabadság között nem oldható fel mérlegelés nélkül.

Irodalom

- ALEXIN, Z. [2014]: Does Fair Anonymization Exist? *International Review of Law, Computers & Technology*. Vol. 28. Issue 1. pp. 21–44.
- GARFINKEL, S. L. [2015]: *De-Identification of Personally Identifiable Information*. National Institute of Standards and Technology. Washington, D.C. <http://dx.doi.org/10.6028/NIST.IR.8053>
- ICO (INFORMATION COMMISSIONER'S OFFICE) [2008]: *Freedom of Information Act 2000 (Section 50)*. Decision notice. 28 July. https://ico.org.uk/media/action-weve-taken/decision-notice/2008/443478/FS_50122432.pdf
- ICO [2014]: *Freedom of Information Legislation and Research Information: Guidance for the Higher Education Sector*. 30 April. <https://ico.org.uk/media/for-organisations/documents/1133/foi-legislation-and-research-guidance-for-the-higher-education-sector.pdf>
- MCHALE, J. V. – JONES, J. [2012]: Privacy, Confidentiality and Abortion Statistics: A Question of Public Interest? *Journal of Medical Ethics*. No. 38. pp. 31–34.
- MILLARD, C. – HON, W. K. [2012]: Defining 'Personal Data' in e-Social Science. *Information Communication and Society*. Vol. 15. No. 1. pp 66.
- RICHARDSON, M. – BRYAN, M. – VRANKEN, M. – BARNETT, K. [2012]: *Breach of Confidence: Social Origins and Modern Developments*. Edward Elgar Publishing Limited. Cheltenham.
- TUDOR, C. – CORNISHY, G. – SPICER, K. [2013]: Intruder Testing on the 2011 UK Census: Providing Practical Evidence for Disclosure Protection. *Journal of Privacy and Confidentiality*. Vol. 5. No. 2. pp. 111–132. <http://repository.cmu.edu/cgi/viewcontent.cgi?article=1101&context=jpc>

Beszélgetés Péterfalvi Attilával a Nemzeti Adatvédelmi és Információszabadság Hatóságának elnökével

Péterfalvi Attila a Veres Pálné Gimnázium angol tagozatán érettségizett, majd az Eötvös Loránd Tudományegyetem Állam- és Jogtudományi Karán szerzett diplomát. PhD-képesítését a Nemzeti Közzolgálati Egyetemen szerezte meg. 1986-tól az Államigazgatási Főiskolán (Nemzeti Közzolgálati Egyetemen) oktat. 2001 és 2007 között adatvédelmi biztosként működött, 2012 óta a Nemzeti Adatvédelmi és Információszabadság Hatóságának elnöke.

Kiindulásul beszéljünk arról, hogy milyen családi, társadalmi környezetből jöttél, hogy telt az ifjúságod, milyen tanulmányokat folytattál, melyek voltak kezdő lépéseid a munka világában?

Értelmiségi családból származom, rokonságom szinte minden tagja pedagógus, szüleim és testvéreim is azok. Iskoláimat a szokásos úton végeztem, a Veres Pálné Gimnáziumban érettségiztem, ezt követően az ELTE¹ Jogi Karára nyertem felvételt. A családom elvárásaival szemben már a kezdetektől jogász szerettem volna lenni, de nyilván nem hagytam érintetlenül a pedagógusi családi környezet sem, mert jogásként is arra törekedtem, hogy tanítsak. Minthogy szüleimnek Balatonalmádiban volt a nyaralója, sokat tartózkodtam ott, ezért az akkori kötelező pályázati rendszerben megpályáztam a Veszprémi Városi Tanács igazgatási osztályának a tanulmányi ösztöndíját, és mikor 1981-ben elvégeztem az egyetemet, ott kezdtem el dolgozni. Tudatosan választottam olyan munkakört, ahol polgári jogi kérdésekkel, például birtokvédelemmel foglalkozhattam. Néhány év után pályámat jogtanácsosként folytattam, majd 1986-ban megpá-

lyáztam a budapesti Államigazgatási Főiskola kihelyezett tagozatának egy tanársegédi állását, melyet el is nyertem, és így munkámat pedagógusként is végezhettem.

Mikor kezdted el érdeklődni az adatvédelem iránt, mikor volt módod intenzíven foglalkozni ezzel a témával?

Főiskolai oktatóként keltette fel az érdeklődésemet a polgári törvénykönyvnek a személyiségi jogokkal kapcsolatos része, miután *Máté Gábor*, akkori tanszékvezetőm is bátorított arra, hogy kutatási témául ezt a területet válasszam. Konkrét feladatot is kiszabott rám, ezen a területen, miután *Könyves Tóth Pál*, a KSH munkatársa révén tudomására jutott, hogy a Hivatalban készül egy adatvédelemmel, információszabadsággal kapcsolatos törvénytervezet. Kérte, hogy véleményezzem, és ettől kezdve jegyeztem el magam ezzel a témával, ekkor ismertem fel a jogi kodifikáció jelentőségét és kedveltem meg az ezzel kapcsolatos munkát. Ebben az időben főleg a Neumann János Társaság révén kerültem kapcsolatba azzal a lelkes társasággal, az említett Könyves Tóth Pálon kívül *Sólyom Lászlóval*, *Majtényi Lászlóval*, *Székely Ivánnal*, akik zászlójukra tűzték egy korszerű, európai normáknak megfelelő adatvédelmi törvény

¹ ELTE: Eötvös Loránd Tudományegyetem.

kidolgozását. Szakmai szempontból sűrű időszak volt ez. Meghatározó élményem, hogy 1991 őszén elnyertem egy NATO²-ösztöndíjat, melynek témája volt a személyiségi jogok, ezen belül a személyes adatok védelme, illetve az információszabadság. Érdekes, Kanadában inkább a személyes adatok védelméről szerezhettem jelentős tapasztalatokat, az Egyesült Államokban pedig az információszabadság kérdését tanulmányozhattam behatóbban. Időben nagyon jól jött ez a tanulmányút, mert itthon javában folytak az új adatvédelmi törvény előkészületi munkálatai és a Fidesz tanácsadójaként aktívan részt vehettem ebben a nagy jelentőségű munkában, megtapasztalhattam, hogy az országgyűlési viták során hogyan formálódik törvénné egy törvénytervezet, és ez a munka is sok tapasztalatot adott a kodifikációs munka mélyebb megismeréséhez. Meggyőződésem ugyanis, ahhoz, hogy egy jogterületet a gyakorlatban jól tudjunk kezelni, mindenekelőtt alapos, jól kimunkált kodifikációs munka eredményeként előállított jogszabály, törvény szükségeltetik. Ennek a törvénynek gyakorlati alkalmazását is nyomon követhettem, hiszen annak idején Majtényi Lászlónak, az első adatvédelmi biztosnak szakértője voltam, majd 2001 és 2007 között adatvédelmi biztosként igazán módomban volt az adatvédelmi törvény rendelkezéseit értelmezni.

Ehhez hozzátehetjük, hogy 2012-től, a NAIH³ elnökeként még több lehetőség van evvel a fontos témával foglalkozni. A 2014. évi országgyűlési beszámolóban-közleményben megemlékeztem az adatvédelmi biztos intézmény 20 éves évfordulójáról. Ez jó alkalom arra, hogy e hosszú időszak tevékeny résztvevőjeként röviden értékeljem ezt a korszakot.

² NATO (North Atlantic Treaty Organisation): Észak-atlanti Szerződés Szervezete.

³ NAIH: Nemzeti Adatvédelmi és Információszabadság Hatóság.

Azt hiszem, büszkék lehetünk arra, hogy már a rendszerváltozás előtt sokat foglalkoztunk ezzel a témával, és arra is, hogy térségünkben elsők között Magyarországon lépett hatályba az európai szinten kompatibilis első adatvédelmi törvény. Európai adatvédelmi ügyekért felelős tárgyalópartnereim közül többen értetlenkedtek, hogy miért született ilyen szigorú törvény. Ilyenkor mindig elmondom, hogy egy-egy jogi szabályozáskor sokszor nem lehet a történelmi körülményektől eltekinteni, hogy hazánkban viszonylag rövid történelmi időszak alatt, a XX. században milyen sokszor megsértették a személyiségi jogokat, így érthető volt az elvárás, hogy olyan törvény szülessék, mely maximálisan védi a személyes adatainkat. Az adatvédelmi törvénynek egy sor alapvető jogi megoldása kiállta az idő próbáját, gondolok például arra, hogy egyazon törvényben van szabályozva a személyes adatok védelme és a közérdekű adatok nyilvánossága. Megjegyzem, ez nem mindenhol van így, de a magyar gyakorlat bebizonyította, hogy hasznos állandóan ütköztetni ezt a két, látszólagosan ellentmondásos rendelkezést. Különösen fontosnak tartom, hogy a törvény fogalomrendszere a mindennapok gyakorlatába is beépült, amiben sokat segítettek az adatvédelmi biztos állásfoglalások, ajánlások, közlemények, melyek értelmezték a sokszor nem könnyű adatvédelemmel kapcsolatos fogalmakat és a törvény egyes rendelkezéseit. Ez a korszerű törvény segítette az adatvédelemmel kapcsolatos jogi kodifikációt más törvények megfogalmazásában is.

Igen, én is fontosnak tartom az adatvédelmi biztos működés során keletkezett dokumentumokat, hiszen ezek az állásfoglalások, ajánlások mintegy szociológiai lenyomatát képezték a magyar társadalom elmúlt évtizedekben lezajló változásának, e téma iránti érzékenységének. Ezért kifogásolom kissé, hogy a NAIH

honlapján ezeknek a dokumentumoknak egy része felhasználóbarát módon nem rendezett, nem kereshető.

Valóban jelentős értéket képviselő dokumentumokról van szó, hiszen ezekben az állásfoglalásokban, ajánlásokban nagyon érdekes és sokoldalú információ van az adatvédelem témájáról. Sőt, ennél még tovább mennék és fontosnak tartom az európai szintű adatvédelemmel kapcsolatos állásfoglalások, ajánlások rendszerezett formában történő hozzáférhetővé tételét. Sajnos ehhez komoly informatikai fejlesztések és emberi erőforrások szükségesek, és tekintve hivatalunknak igen szűkös költségvetését csak bizonyos aktuális feladatok megoldása után tudunk erre megfelelő összeget biztosítani, ehhez még hozzátenném, hogy jelenlegi hivatali informatikai kapacitásunk egy családi internetes elérhetőség színvonalán van. Azt tervezem, hogy hivatali működésem második felében mindenképpen megoldjuk ezt a problémát, akár pályázat útján is szerzünk pénzt, és megfelelő formában elérhetővé tesszük ezeket a dokumentumokat.

Visszatérve az 1992-ben hozott adatvédelmi törvényre, ha ez a törvény jól betöltötte a feladatát, miért volt szükség arra jelentős változtatásra, melyet a 2011. évi CXII. törvény az információs önrendelkezési jogról és információszabadságról (a továbbiakban: Infotv.) szóló jogszabály jelentett. E törvény mennyire kompatibilis az európai uniós elvárásokkal? Nem estünk-e át a ló másik oldalára, amikor erős hatósági jogosítványokat kapott a Hivatal? Nem tompul-e a „nép ügyvédjének” szerepe? Nem vált e túlságosan bürokratikussá az adatvédelem ügyének kezelése?

Gondolom, a terjedelmi korlátok csak rövid választ engedélyeznek, hiszen egy tanulmányt lehetne írni az indokokról. Csak azt tudom

mondani, hogy már adatvédelmi biztostként is szembesültem azzal, hogy a jogszerűtlen adatkezelések visszaszorítására nem elegendők az ún. ombudsman-típusú eszközök (felszólítás, ajánlás), hanem keményebb intézkedés szükséges, mely magába foglalja a hatóságként fellépés lehetőségét. A hatósági fellépés terjedő félben levő kívánalmát egyébként az európai uniós jogszabályok is tartalmazzák, ezért a jogalkotó úgy döntött, hogy átalakítja az adatvédelemmel kapcsolatos ombudsmani rendszert és létrehozza a NAIH-t. A Hatóság vizsgálatára vonatkozó szabályok azonban továbbra is alapvetően ombudsman-típusúak, és ha az nem vezet eredményre, vagy ha olyan súlyúak, akkor lehet megindítani az adatvédelmi hatósági eljárást, mely akár komoly pénzbírságot is eredményezhet. Ehhez még hozzá kell tennem, hogy jó lenne, ha a bírságolási jog nemcsak az adatvédelem, hanem az információszabadság tekintetében is meglenne. Az elmúlt évek tapasztalatai alapján mondhatom, hogy a jelentős összegű bírságolási lehetőség komoly visszatartó erőt is jelent. Emellett egy sor új típusú rendelkezést is tartalmaz az Infotv., például létrehozta a belső adatvédelmi felelősök konferenciáját, amelyben a Hatóság szervezésében alkalom nyílik az adatvédelemmel kapcsolatos gyakorlat összehangolására és a jogalkalmazási tapasztalatok megosztására. Jelentősen megváltozott a Hivatal szerepe a titokfelügyeleti intézmény tekintetében, és a jogalkotó e tekintetben igen erős jogosítványokkal ruházta fel a NAIH-t. Egyébként az Infotv. a látszat ellenére harmonizál a régi adatvédelmi törvénnyel, számos rendelkezést olykor szó szerint átvesz a korábbi jogszabályból, ugyanakkor néhány ponton – az adatvédelmi biztosi tapasztalatokat is figyelembe véve – pontosít. Például jobban körül írja azt, hogy a közfeladatot ellátó szerv vezetőjének mely személyes adata nyilvános, ezek egy részét felsorolja, illetve utal más törvényre. Azt hiszem, a bürokratizálódás

veszélye mindig fenyeget, ezért igyekszem a NAIH-t úgy irányítani, hogy a hivatal zökkenőmentesen fogadja be a bejelentéseket és az azzal kapcsolatos ügyeket minél szakszerűbben és határidőre intézzék el. Az adatvédelmi biztosi gyakorlatomhoz hasonlóan, igyekszem a nyilvánosságot is mind többször igénybe venni a hivatal munkájának megismertetése céljából, és felhívni azokra a témákra a figyelmet, melyek a nyilvánosságot érdekelhetik, illetve amit jómagam fontosnak tartok.

Az adatvédelmi biztosi állásfoglalásokban megtalálható, hogy egyes állami intézmények részéről időről időre felmerül az igény, hogy a személyi azonosítók adatait (például személyi, társadalombiztosítási és adóazonosító) együtt kezeljék, melyet ezek az állásfoglalások következetesen elutasítottak. A Sólyom László vezette Alkotmánybíróságnak az 1991-es személyi számmal kapcsolatos döntése is erről szólt. A civil adatvédők is általában az osztott adatkezeléseket támogatják. Az elektronikus személyi igazolvány bevezetése nem okoz-e gondot ezeknek az elveknek érvényesítésekor?

Azt gondolom, hogy az 1991. évi alkotmánybírói döntésnek, főleg szellemiségében, ma is érvényes üzenete van, vezetésében az, hogy univerzális személyi azonosító használata továbbra sem lehetséges. Azonban 1991-től az informatika fejlődése igen nagy lendületet vett, és ennek köszönhetően – technikai szempontból – a kapcsolati kódok használata egyre kevésbé ütközik akadályba. Ez így van szerte az Európai Unióban, azonban az adatvédelem és adatbiztonság kérdését szorosan együtt kell kezelni, hiszen a technikai fejlődés miatt a biztonsági kérdések egyre inkább előtérbe kerülnek. Ez érvényes az elektronikus személyi igazolvány és a mögötte levő különböző személyi azonosítók adatkezelésére is. Anélkül, hogy a részletekbe belemennék, jelzem, hogy az elektronikus kártyán levő

adatokhoz nem lehet egy leolvasóval univerzálisan hozzáférni, csak célhoz kötötten, megfelelő jogszabályi előírások által osztottan lehet az információkat lekérdezni. Ez azt jelenti, hogy az adatok közül mindenki azt láthatja, ami rá tartozik. E számítástechnikai megoldásnál sem csökkenhet a védelmi szint, tehát meg kell teremteni az adatbiztonságot az adatokat tartalmazó kártya és a leolvasó esetében is.

Az 1990-es évek ombudsmani beszámolóiban még a nagy állami adatkezelők (például NAV⁴, egészség-, nyugdíjbiztosító, rendőrség) adatkezelése volt problémás, később viszont a nagy gazdasági erőfölénnyel rendelkező magánszervezetek adatkezelésének hiányossága okozott gondot (bankok, biztosítók, telefon-szolgáltatók). A korábbi országgyűlési beszámolóiban kiemelten kezelték ezt a kérdést, a NAIH által jegyzett beszámolóiban viszont ez a téma háttérbe szorult. Mi ennek az oka?

Teljesen érthető volt, hogy kezdetben a nagy állami intézmények adatkezelési gyakorlata állt előtérben, hiszen a pártállami körülmények között szocializált állami alkalmazottaknak is alkalmazkodniuk kellett a rendszerváltás után az adatvédelmi törvény által megkívánt előírásokhoz. Maguknak a nagy állami adatkezelőknek is ki kellett alakítaniuk belső szabályozóikat, „szolgáltatóbarátibbá” kellett tenni informatikai rendszerüket. Menet közben megszülettek a jelentős mennyiségű személyes adatok kezelését szabályozó ún. szektorális törvények, melyek kihatottak a magánszervezetek, a betelepült multinacionális cégek (melyek anyacégeire más országok sokszor megengedőbb szabályai voltak érvényesek) adatkezelési gyakorlatára, és bizony ez a gyakorlat nem mindig volt kompatibilis a jogszabályi környezettel. Időközben jelentősen fejlődött az informatika, és főleg az internet

⁴ NAV: Nemzeti Adó- és Vámhivatal.

használata hatott az adatkezelés körülményeire. Miután az idő múlásával az állami szervek jelentős része megtanulta a törvény általi elvárások teljesítését, érthető, hogy ezek után a magánszervezetek adatkezelésével kapcsolatos beadványok többségbe kerültek. Mára már mindkét szektor átesett a tanulási folyamaton, és inkább egyéb szempontok szerint lehet strukturálni ezt a témát. 2012 óta ez már a NAIH országgyűlési beszámolóiban is jelentkezik. Először is csökkentettük a terjedelmet, mert több száz oldalas anyagot nehéz kommunikálni. Fontosnak tartjuk, hogy a közvéleményt statisztikai adatok segítségével is tájékoztassuk munkánkról, a beszámolóban hangsúlyosabban szerepet kapjanak a nemzetközileg is fontosabb ügyek, például kiemelten kezeljük az új technológiák által generált adatvédelmi problémákat. Az új területeken jelentkező adatkezelési témákat csokorba rendezve taglaljuk, például a termékbemutatók, vagy a követeléskezelés kérdését. Természetesen a beszámolóban a nagy állami és magán adatkezelőkkel kapcsolatos beadványok értékelése is szerepel abban az esetben, ha új típusú problémák felvetéséről van szó.

Az ombudsmani és a NAIH éves beszámolóiban statisztikák olvashatók a beérkezett beadványok jellegéről. Már a kezdetektől jellemző volt, hogy az információszabadsággal foglalkozó beadványok jóval kisebb (10 százalék körüli) arányban szerepelnek, mint a személyes adatok védelmével kapcsolatos témák. Sokan ezt a magyar polgárok közélet iránti közömbösségével magyarázzák. Igaz ez? Ha igaz, akkor a közérdekű adatokhoz történő hozzáférés fizetősé tételenek nagyobb lehetősége, kiterjesztése nem fogja-e ezt az aktivitást visszavetni?

Azért némileg emelkedett a közérdekű adatok miatti beadványok aránya, jelenleg 12

százalék körüli. A rendszerváltás után volt egy illúzió, hogy demokratikus körülmények között majd az állampolgárok jelentős része állandó érdeklődést mutat a közérdekű adatok iránt. Ezt az érdeklődést némileg helyettesítik a médiában megjelenő tudósítások a közérdeklődésre számot tartó témák napirenden tartásáról, taglalásáról. Megint csak a számítástechnikai fejlődésre utalnék, hiszen az internet korában, megfelelő jogszabályi előírások és technikai lehetőségek folytán ma már nagyon sok információ közvetlenül megtalálható a különböző állami intézmények honlapján. Egyébként azt a viszonylag állandósult arányt úgy is lehet magyarázni, hogy azért ilyen alacsony ez az érték, mert ezen a területen minden rendben van, de talán ezt túlzás lenne kijelenteni. Több javaslatot tettünk a közérdekű adatok nyilvánosságra hozásának érdekében, például azokban a témákban, melyek iránt jelentős az állampolgárok érdeklődése, megelőzve e témák iránti nagyszámú kérelmeket, az adott állami szerv ezzel kapcsolatos információkat közvetlenül az internet segítségével hozza nyilvánosságra. Külön felhívtuk a figyelmet arra, hogy ha egy állami szerv kiad egy közérdekű információt, valós tartalmúnak, tényszerűnek és pontosnak kell lennie. Valószínű, hogy ha az elmúlt húsz év alatt a közérdekű adatok iránti igény ilyen állandóságot mutat, akkor a közeljövőben sem fog ez erősen változni. Ezért is gondolom azt, hogy a legutóbbi jogszabályváltozás a közérdekű adatok fizetősé tételenek megváltozásáról erre a témára nem fog lényeges befolyást gyakorolni. Érzékeltetve a problémát, a saját gyakorlatunkból is tudok példát mondani. Amikor egy igénylő azt kérte, hogy válogassuk ki a hozzánk érkezett beadványokra adott állásfoglalásaink közül azokat, amelyek megfelelnek az általa megadott szempontoknak, sajnos olyan feltételeket adott meg, amelyek nem szerepeltek az ügyviteli nyilvántartásunkban, ezért sok

aktát kellett külön, egyenként átnézni és értékelni, melyek a megadott szempontoknak megfeleltek és ez jelentős költséggel és munkaráfordítással járt. A jogalkotó szándéka tehát többek között e probléma kezelésére irányult. Ezzel kapcsolatban azt tudom mondani, hogy miután friss jogszabály-módosításról van szó, várjuk meg gyakorlati megvalósulását. Azt meg tudom ígérni, hogy ha a NAIH azt fogja tapasztalni, hogy az adatgazdák ezt a lehetőséget az információszabadság korlátozására próbálják felhasználni, akkor jogkörünkben eljárva, az esetleges káros tendenciák megakadályozására akár törvénymódosítást is kezdeményezünk

Ha már szó volt az információszabadságról szóló törvény legutóbbi módosításáról, milyen további indokok voltak a változtatásokat illetően? A kormány elfogadta a NAIH minden módosító javaslatát?

Az Infotv. módosításának egyik célja, hogy jobban harmonizáljon a készülő európai uniós rendelettel, például, hogy teljesüljön az a kívánalom, mely azonos feltételek melletti adatkezelést jelent a tagállamoknál, valamint, hogy harmadik országban is azonos módon történjék az adatszolgáltatás. Továbbá olyan új szabályok kerültek be a törvénybe, melyek alaposabban figyelembe veszik a közfeladatot ellátó szervek érdekeit, ügyelve az információt kérő személyek érdekeire is. A NAIH el nem fogadott javaslatai közé tartozik például, hogy az adatvédelmi nyilvántartás erre külön létrehozott honlapra történő bejelentkezéssel alapuljon, és ne Hatóságunknak kelljen e nyilvántartás minden részletét kezelnie. Már volt szó arról, hogy az információszabadság megsértése miatt szükség lenne bírságolási jogra, de legalább arra, hogyha valamely állami szerv a kötelezően nyilvánosságra hozandó információt nem teszi közzé, akkor ezt a jogsértést el

lehessen marasztalni ily módon is. A készülő európai uniós adatvédelmi rendelet szól az ún. incidenskezeléssel kapcsolatos nyilvántartásról, jó lett volna, ha ezt az adatvédelmi intézményt már most bevezetjük.

Milyen új technikai eszközök okozzák a legtöbb gondot a személyes adatok védelme tekintetében? Ezeket a gyors technikai változásokat hogyan lehet a jog eszközeivel követni, gondolva például a legújabb ilyen eszközre, a magáncélú drónok alkalmazására.

A technika – egyébként örömteli – fejlődése igen megnehezíti adataink védelmét. Ez a technikai fejlődés összefügg a terrorizmus elleni küzdelemmel is, ami számos technikai vívmányt igénybe vesz. Ilyenek például az egyre inkább terjedő biometrikus rendszerek, melyek biztonságosabban tudnak személyeket azonosítani, mint a hagyományos azonosítók. Ezek közé tartozik az ujjlenyomat, az arcképfelismerés, a vénaszkenner vagy az íriszazonosítás. A részleteket mellőzve azt mondhatom, hogy ezeknek az eszközöknek az alkalmazásakor különösen fontos a szükségesség, arányosság, alkalmasság hármas kritériumrendszerének vizsgálata, mert könnyen belátható, hogy egy sportrendezvény beléptető rendszerének esetében szükségtelen arcfelismerő technikát alkalmazni, ugyanakkor nemzetbiztonsági érdekből előfordulhat ezen eszköz használata. Ugyancsak probléma, hogy a biometrikus azonosítók működhetnek univerzális azonosítókként, ami – mint említettem – tilos az alkotmánybírói határozat alapján. Ezért nem szabad ezeket a biometrikus azonosítókat egy központi nyilvántartásba szervezni, hanem csak oly módon, mely a magánszemélyek érdekeit maximálisan figyelembe veszi. A NAIH ajánlást adott ki a drónokkal kapcsolatos adatkezelésről, külön konferenciát szerveztünk e téma megtárgyalására. Már definiálni

sem könnyű ezt az eszközt, lényegében egy pilóta nélküli ön- vagy távirányítóval működő repülő tárgyról van szó, melyben a – főleg az erre a tárgyra szerelt – kiegészítők által megvalósuló adatgyűjtések, -kezelések, -tárolások jelentik az igazi problémát. Hatóságunk külön törvényt javasol ezeknek az adatvédelmi szempontból igen kényes technikai eszköznek a használatáról, ugyanis nem kielégítő az utalás más, hasonlóknak tűnő, szabályozási tárgyra (például köztéri kamerázásra). E törvénynek rendelkeznie kell a drónok állami, kereskedelmi és kiemelten a magáncélú felhasználásáról. Ebben az esetben is fontos a megfelelés az említett hármas kritériumnak, és különösen fontos, hogy ezek az eszközök (valószínűleg) a Légügyi Hatóság feladataként legyenek regisztrálva. Ugyanakkor általános jellegű probléma, hogy a gyorsan változó technikák kihívását a jogalkotás hogyan követi. A drónokkal kapcsolatos konferencián is több felszólaló, például a jelenlevő bírók is jelezték, hogy nem lenne szerencsés, ha minden egyes új technikai eszköz kapcsán új törvény születne. Inkább olyan rugalmas jogi szabályozásra lenne szükség, mely a részleteket mellőzve, a bírói gyakorlatra bízna a vitás kérdések elrendezését.

Történelmileg igen rövid idő alatt terjedt el az internet, melyen sok személyes adattal találkozhatunk. Legutóbb a luxemburgi székhelyű Európai Bíróság megerősítette az egyénnek a „felejtéshez” való jogát. Mi a véleményed erről, és hogyan lehet ezt a gyakorlatban megvalósítani?

A NAIH közleményben üdvözölte ezt a bírósági megközelítést, mert ebből az örök időkre szóló nyilvánosság tilalma származik, mely azt jelenti, hogy a személyiségi jogok védelme előbbre való az üzleti szféra érdekeinél és a technológia fejlettségéből következő lehetőségeknél. A Hatóság egyetértett az Európai

Bíróság azon érvelésével is, amely szerint az egyén, a közéletben játszott szerepével összefüggésben, a felhasználók tájékoztatásának korlátozását nem kérheti. A keresőmotorokra a magyar hatósághoz is folyamatosan érkeznek panaszok, ezek azt kifogásolják, hogy a találati lista egyes elemeinek eltávolítására tett kérelmeket nem teljesítették. Az ilyen panaszok elbírálása szoros európai együttműködésben zajlik, és a következő években az esetjog gyarapodására lehet számítani.

Az országgyűlési beszámolókból csak elvéve van szó a statisztikával kapcsolatos adatvédelmi problémákról, ami véleményem szerint arra utal, hogy ezen a területen hagyományosan erős az adatvédelem iránti elkötelezettség. Hogyan látod a statisztika és az adatvédelem kapcsolatát, milyen tapasztalataid vannak erről?

Többször kerültem a statisztika közelébe, legnagyobb élményem ezzel kapcsolatban, hogy figyelemmel kísérhettem a 2001. és a 2011. évi népszámlálás munkálatait. Ekkor éreztem azt, hogy milyen jelentősége van annak, hogy az adatszolgáltatók, a lakosság bízson a statisztikában, abban, hogy adataikat csak statisztikai célra használják. Mindkét esetben adatvédelmi szempontból jó megoldásnak tartottam, hogy a személyes adatok név nélkül kerültek felvételre, és a népszámlálási adatok feldolgozásának teljes folyamatában kiemelten figyeltek az adatvédelem szempontjaira. Mindkét alkalommal jó kapcsolat alakult ki a KSH⁵ vezetésével, és kölcsönösen tájékoztattuk egymást a felvetődő kérdésekről. Az állam az adófizetők pénzéből gyakran költ jelentős összeget egy-egy statisztikai felvételre, ezért jogos az igény, hogy a felvett információk minél szélesebb körben hasznosuljanak, már csak azért is, mert ezek az adatok

⁵ KSH: Központi Statisztikai Hivatal.

közérdekűnek minősülnek. Azonban a felvett személyes adatokat védeni kell, tehát oly módon kell a statisztikai adatokat nyilvánosságra hozni, hogy az adatszolgáltató személye ne legyen azonosítható. Tudom, hogy nem könnyű ezt a kettős elvárást teljesíteni, adatvédelmi biztosként egy korábbi állásfoglalásomban, 2003-ban javasoltam, hogy az akkori Oktatásügyi Minisztérium egyik kutatócsoportja a 2001. évi népszámlálás nemzetiségi adatainak kutatását a KSH számítógépparkja, a KSH szakemberei által ellenőrzött módon kerüljenek feldolgozásra, így biztosítva ezen igen érzékeny személyes adatokat tartalmazó adatállomány védelmét. Azt is fontosnak tartom, hogy a statisztikai célból felvett adatok csak ebből a célból legyenek hasznosítva, minden más – például adminisztratív célú – felhasználás nem tolerálható.

Beszélgetésünk harmadik résztvevője egy Luca névre hallgató, igen barátságos szájkás

szőrű tacskó, úgy látom, szereted a kutyákat. Szabad idődben mivel foglalkozol, milyen kedvteléseid vannak?

Igen, nagy kutyabarát vagyok, mindig volt kutyám, most történetesen Luca, aki valóban emberbarátnak mondható. Mint említettem, a Balaton mellett felnőve, már fiatal koromtól eljegyeztem magam a vitorlás sporttal, úgyhogy nyáron, amint szabadidőm engedi, sokat vitorlázom, nagyon szeretem a Balatont és vidékét. A téli tájat is szeretem a hófödte hegyeket, így a téli sportok közül a sielés a hobbim, amikor tehetem, felcsatolom a sílécet.

Köszönöm a beszélgetést, jó egészséget kívánok!

Lakatos Miklós,

a KSH szakmai főtanácsadója

E-mail: Miklos.Lakatos@ksh.hu

A Nemzeti Adatvédelmi és Információszabadság Hatóság 2014. évi beszámolója

A NAIH¹ 2014. évi beszámolója a hagyományt teremtve követi a korábbi két év beszámolójának struktúráját, bár kisebb változtatásokat végrehajtottak, például a NAIH nemzetközi szerepvállalásainak taglalása ezúttal a beszámoló végére került. Már az eddigi NAIH-beszámolókból is a hatóság elnöke köszönti az olvasót, így van ez a jelen beszámoló esetében is. Dr. Péterfalvi Attila ebben a köszöntőjében megemlékezik az adatvédelmi biztosi intézmény létrejöttének húsz éves évfordulójáról. Méltatja az elmúlt húsz év

eredményeit, megemlékezik az első évek nehézségeiről, az azt követő konszolidáció éveiről, majd arról, hogy ez az intézmény azonban 2011-re fokozatosan „kinőtte” kereteit. A klasszikus ombudsmani szerepkör, mely elsősorban a meggyőzésre épül, egyre inkább hatósági jogosultságokkal bővült, végül a jogalkotó a magyar ombudsmani modell átalakítása során úgy döntött, hogy az adatvédelem országgyűlési biztosi intézményét nevében, szervezetében és jogállásában is hatósággá alakítja át.

A 2014. évi beszámoló elején a Hatóság működésével kapcsolatos statisztikai adatokat találjuk. E statisztikák egyik legfontosabb in-

¹ NAIH: Nemzeti Adatvédelmi és Információszabadság Hatóság.

formációja, hogy a NAIH ügyei miként oszlanak meg információs ágak szerint. Ezt a megoszlást befolyásolja az, hogy a NAIH az adatvédelmi nyilvántartási bejelentéseket és az azokhoz kapcsolódó konzultációs ügyeket (691 ügy) az információs ágak szerint az egyéb kategóriába sorolja be. Az ügyek információs jogok szerinti megoszlása a következő: adatvédelmet érinti 1599 (53%), információszabadságot érinti: 386 (13%), mindkét alapjogot érinti: 78 (2%), egyéb, a Hatóság más feladatkörébe tartozó ügy: 967 (32%). A közérdekű adatok megismeréséhez való alapjogot tehát a vizsgált ügyek 15 százaléka, összesen 464 ügy érintette. Ez az adat a 2013-ban érkezett 420 ügyhöz képest mind számszerinti, mind aránybeli növekedést jelez. Az adatsor alapján megállapíthatjuk, hogy az állampolgárok továbbra is jóval érzékenyebbek a személyes adataik kezelésével kapcsolatos, mint a közérdekű adatok nyilvánosságát érintő ügyekben.

A 2014. évi beszámoló külön fejezetet szentel a technológiai fejlesztések információs alapjogi hatásainak bemutatására. Terjedelmi okokból e fejezetet nem lehet minden részletében ismertetni, csak jelezzük, hogy a beszámoló összefoglalja, a legújabb technológiák az ún. okos eszközök (például mobiltelefonok) főbb jellemzőit, az internetalapú ún. felhőszolgáltatások veszélyeit, e technológiai eszközök állami befolyásolásának lehetőségeit, illetve korlátait. Az interneten az adatok kezelése, feldolgozása egyre kevésbé rendelhető meghatározott földrajzi helyszínhez. Teret hódítanak az ún. felhőszolgáltatások, amelyek egyaránt lehetővé tehetik az adatok tárolását, feldolgozását, illetve az alkalmazás-szolgáltatást. A szolgáltatásnak nem állják útját országhatárok, mert az adatok „bárhol lehetnek a felhőben”. Ha nem határozható meg a tényleges adatkezelés helye, akkor kérdésessé válik, hogy az adatkezelés melyik állam joghatósága alá tartozik és milyen jogi előírások vonatkoznak

a felhőben kezelt adatok védelmére. A felhőszolgáltatást rendszerint nem egy vállalat nyújtja, hanem sok, különböző szolgáltató együttműködése szükséges hozzá. Ilyen körülmények között a személyes adatok kezeléséért való felelősség is megoszlik a szolgáltatást nyújtók között. A technológiai fejlesztéseket nemcsak a felhasználók igényei lendítik előre, hanem az is, hogy a világszerte jelenlévő informatikai nagyvállalatok számára felértékelődtek azok az információk, amelyeket az általuk gyártott eszközök és az általuk nyújtott szolgáltatások révén a felhasználókról összegyűjthetnek. Ugyanis a felhasználókról összegyűjtött személyes adatok felhasználhatók például piackutatásra, reklámüzenetek személyre szabásához, a meglévő szolgáltatások optimalizációjához, termékek és szolgáltatások tervezéséhez, vagy adatkereskedelemre.

A beszámoló ezt követően kiemeli néhány a technológiai fejlesztéssel kapcsolatos témát, ezek egyike a biometria alkalmazásával foglalkozik, melyben a NAIH állásfoglalást is kiadott. A beszámoló ezzel kapcsolatos legfontosabb pontjait idézzük:

„A személyazonosításhoz használt biometrikus index olyan adat, amely közvetlenül, véglegesen, megváltoztathatatlanul és letagadhatatlanul kapcsolódik az adatalanyhoz. Ezért a személyazonosításra alkalmas biometrikus index a természete szerint egységes személyazonosító kód. Valamely univerzális azonosító kód, akár a biometrikus indexadat alkalmazásának általánossá válása ellentétes lenne az osztott információs rendszerekre vonatkozó, az Alkotmánybíróság által kimunkált alkotmányossági követelménnyel és utólag megkérdőjelezné olyan, a magyar információs alapjogvédelem fejlődése szempontjából fontos eredményeknek a létjogosultságát, mint az ágazat specifikus azonosítók rendszere és az összerendelési nyilvántartás.

Egyes biometrikus személyazonosítási technológiák alkalmazása nem igényli az érintett közreműködését, ezért rejtett megfigyelést tesz lehetővé. Az automatizált arcfelismerésen alapuló azonosítás és a hasonló biometrikus technológiák kiterjedt, tömeges alkalmazása ahhoz vezetne, hogy a települési közterületek, a nyilvánosság számára nyitva álló egyéb közterületek, és a tömegközlekedési eszközök teljesen megszűnnének a magánélet színterei lenni. Mindenkinél azzal kellene számolnia, hogy titokban folyamatosan, automatikusan megfigyelhetik, ha kilép a lakásból. Ilyen helyzet kialakulása nyilvánvalóan sértené a magánélet tisztelgetéséhez és személyes adatok védelméhez fűződő jogot. A biometrikus technológiák alkalmazása nem teremtheti meg rejtett, tömeges megfigyelés bevezetésének előfeltételeit.” (23. old.)

„Olyan jogi kereteket kell megteremteni, amelyek garantálják, hogy a biometrikus azonosításon alapuló titkos információgyűjtés, vagy egyéb, kötelező jellegű megfigyelés

- nem ölthet tömeges méreteket;
- még pszeudonim (álneves) formában sem lehet készletező;
- szigorúan csak törvényben meghatározott, legitim cél érdekében történhet.” (24. old.)

A beszámoló szól a NEK² koncepciójáról és törvénytervezetéről, melyet a NAIH is véleményezett. Megállapítja: a kártyabirtokosok adatainak biztonsága szempontjából kedvező, hogy a NEK-ben a kártyakibocsátás engedélyezése szabályozott eljárás keretében történik. A kártyaelfogadók a kártyán tárolt információk közül csak azokhoz férhetnek hozzá, amelyet az adott kártya tranzakcióhoz a kártyafelhasználó hozzáférhetővé tesz. A rendszerbe tartozó kártyák használata során a kártyák érvényességére vonatkozó informáci-

² NEK: Nemzeti Egységes Kártyarendszer.

ók a központi nyilvántartásból online lekérdezhetők, azonban a központi nyilvántartás e lekérdezéseket nem naplózhatja. A központi kártyanyilvántartás kártyánként elkülönítve tartalmazza a kártyákra és a kártyabirtokosok személyére vonatkozó azonosító adatokat, valamint azt, hogy valakinek milyen kártyái vannak, azt csak néhány, a törvényben meghatározott szervezet – például a bíróságok vagy a nemzetbiztonsági szolgálatok – ismerhetik meg a törvényben meghatározott feladataik ellátásához.

A következő téma, melyet a NAIH beszámolója fontosnak tartott kiemelni, a kapcsolati kódok kérdése. A problémát az okozza, hogy a kapcsolati kód képzésére – a személyiadat- és lakcímnnyilvántartás kapcsolati kódjai kivételével – nincsenek általános érvényű előírások. Ezért 2014-ben a Hatóság több, kapcsolati kódot használó adatkezelőnél vizsgálta, hogy esetükben a kapcsolati kód képzésének módja összhangban van-e személyes adatok védelmére vonatkozó előírásokkal. A beszámoló megállapítja, hogy:

„Jelenleg nincs jogszabály által kizárva az, hogy több adatkezelő, eltérő célú adatkezelések esetében ugyanazokból a személyes adatokból, egyazon eljárással képzett kapcsolati kódot használjon. Kifejezett jogszabályi tiltás hiányában az is gyakorlattá vált, hogy az adatkezelők adott kapcsolati kód használatának időtartamát nem korlátozzák adott adatkezelési műveletre, vagy bizonyos ésszerű időszakra, hanem változatlan kapcsolati kódokat használnak az adatkezelés teljes időtartama alatt. Fontos rámutatni arra, hogy ha a fentiek értelmében egy személyhez több, különböző célú adatkezelésben ugyanaz az állandó kapcsolati kód tartozik, akkor a kapcsolati kód ezen adatkezelésekben olyan azonosítónak válik, amely megteremti a technikai feltételeket a különböző célú adatkezelések összekapcsolásához.” (26. old.)

A NAIH azért tartja ezt kiemelten fontos témának, mert tart attól, hogyha különböző célú adatkezeléseket azonos kapcsolati kóddal rendezik össze, akkor fennáll egy univerzális kapcsolati kód kialakításának veszélye, melyet az Infotv.³ kifejezetten tilt.

A beszámoló III. fejezete a hatósági ügyekkel foglalkozik, mely szól az adatvédelmi hatósági eljárásról, más hatóságokkal történő együttműködésről, a bírósági eljárás alá vont ügyekről és néhány kiemelt vizsgálatról. Ez utóbbiak közé tartoznak a termékbemutatók. A beszámoló jelzi, hogy a Hatósághoz több éve érkeznek olyan panaszbeadványok, melyek kifejezetten termékbemutatót szervező, lebonyolító, illetve terméket forgalmazó cégek adatkezelését kifogásolják, így jól látható a Hatóság számára, hogy egy hosszú évek óta fennálló problémáról van szó, mely a társadalom jelentős részét érinti. A beszámoló megállapítja, hogy:

„A termékbemutatókra csalogatandó emberekhez több úton jutnak el: telefonon, interneten, levélben, nyomtatott sajtóban megjelentetett hirdetésekben, úrlap elhelyezésével és közösségi helyeken szórólapokat osztogatva, valamint adatbázisokat vásárolva gyűjtik a magánszemélyek személyes adatait és építik fel saját adatbázisukat. A termékbemutatókkal kapcsolatban felmerült főbb jogsértések lényege, hogy a cégek nem feltétlenül jogszerű forrásokból szerzik be a meghívandó személyek nevét, címét, telefonszámát, több szempontból is valótlan tájékoztatást nyújtanak a meghívottak/résztvevők/vásárlók részére, túl sok személyes adatot rögzítenek, tárolnak, továbbá a tisztességesség követelményébe ütközik az adatkezelési tevékenységük.” (34. old.)

³ Infotv.: Az információs önrendelkezési jogról és információszabadságról szóló 2011. évi CXII. törvény.

A NAIH hatósági jogkörében eljárva jelentős összegű bírságokat szabott ki és a tapasztalatokat egy jelentésben foglalta össze, amelyben felhívja mind az e tevékenységi körrel foglalkozó vállalkozások, mind az érdeklődő magánszemélyek, mind pedig a jogalkotó figyelmét arra, hogy milyen nehézségek láthatóak, és milyen helytelen gyakorlat alakult ki, továbbá javaslatot tesz a felmerülő problémák megoldására.

A NAIH a követeléskezeléssel kapcsolatban is végzett célzott vizsgálatokat és azt tapasztalta, hogy a követeléskezelést végző cégek az adósok egyezségkötésre való hajlandóságának és a fizetőképességének felmérése céljából az érintettek vagyoni és jövedelmi viszonyait, továbbá a munkahelyi és családi körülményeit is felméri. A Hatóság eljárási tapasztalatai szerint a követeléskezelést végző cégek által kezelt adatkör az adatkezelés céljához képest általában túlzott mértékű. A Hatóság szerint a követeléskezeléssel üzletszerűen foglalkozó pénzügyi intézmények követeléskezelési gyakorlata alapján azonban általánosan megállapítható, hogy a megbízóik révén birtokukba került személyes adatokkal nem csupán technikai műveleteket hajtanak végre, hanem azokat önállóan, a megbízóiktól gyakorlatilag független módon felhasználják és azokra támaszkodva a megbízási szerződéseik keretei között döntéseket hoznak. A NAIH e témában is ajánlást tett, melyben a követeléskezeléssel összefüggő adatkezelések kapcsán minimumkövetelményeket fogalmazott meg, és intézkedési javaslatot tett a jogalkotó szervek számára.

Az országgyűlési adatvédelmi beszámolók állandó témája a direkt marketing célú adatkezelések anomáliái. A Hatóság első ízben vizsgálta a direkt marketing szakterületet, ezen belül az adatbázist építő és azok hasznosításával foglalkozó társaságokat. Ugyanis az érintettek beadványaikban gyakran sérelmezik,

hogyan olyan társaságoktól kapnak kérést reklámüzenet és keresik meg őket különböző csatornákon ajánlatokkal, melyek részére nem adtak felhatalmazást adataik kezelésére.

„A Hatóság határozataiban a vizsgált cégeket adatvédelmi bírság megfizetésére kötelezte, továbbá felszólította őket adatkezelési gyakorlatuk és adatkezelési szabályzatuk jogszabályoknak megfelelő átalakítására, valamint hogy a módosult szabályozásról és adatkezelési gyakorlatokról adott tájékoztatást követően kérjék az összes regisztrált hozzájárulását, illetve hozzájárulásának megerősítését, a meg nem erősített regisztrációk esetében pedig gondoskodjanak a felhasználók adatainak dokumentált törléséről.” (48. old.)

A beszámoló IV. fejezete a NAIH jogalkotással kapcsolatos tevékenységét ismerteti, jelzi a feladat komolyságát, mert az esetleges problémák kiszűrését, már egy szakszerűen összeállított jogszabállyal is meg lehet oldani.

A beszámoló V. fejezete azokról az ombudsmani típusú vizsgálati ügyekről szól, melyek az adatvédelem témáját érintik.

A technikai fejlődéssel is összhangban van, amikor a rendőrség az elektronikus hírközlési szolgáltatóktól egy konkrét földrajzi helyet ellátó állomás (vagyis egy konkrét földrajzi helyszínhez kapcsolható bázisállomás) egy meghatározott időszakra vonatkozó teljes adatforgalmának átadását kéri. Nyilvánvaló, hogy a rendőrségnek a nyomozati munkához igen fontosak a konkrét földrajzi helyszínre vonatkozó hívásadatok, azonban – tekintettel arra, hogy igen kényes személyes adatokról van szó –, a rendőrségnek is körültekintően kell eljárni, ezért a NAIH véleménye szerint:

„A rendőrségnek a megkeresésben pontosan meg kell határoznia, hogy milyen adatokra terjed ki az adatigénylés, illetve hogy mi az

adatigénylés célja. Emellett a megkeresésben meg kell jelölni azt is, hogy a személyes adat kérése miért tekinthető elengedhetetlenül szükségesnek. Amennyiben a későbbiekben kiderül, hogy a megkeresett szervezet olyan személyes adatokat is az eljáró hatóság rendelkezésére bocsátott, amelyek a megkeresés célja szempontjából irrelevánsak, úgy a hatóságnak azokat törölnie kell.” (67. old.)

A pénzügyi szektor adatkezelésének legnagyobb problémája továbbra is az, hogy a szerződéskötés során az állampolgárok vagy nem kapnak tájékoztatást – vagy félrevezető tájékoztatást kapnak – arról, hogy mely adat-szolgáltatás kötelező, mely adatok megadása szükséges és elengedhetetlen feltétele a szolgáltatásnyújtásnak, s melyek azok, amelyeknél az ügyfelet valós választási lehetőség illeti meg, mint például a mobiltelefonszám, e-mail cím megadása esetében. A NAIH leszögezi, hogy:

„A pénzügyi szolgáltatóknak a különböző célú adatkezeléseiket egyértelműen el kell határolniuk. A közvetlen üzletszerzés, piackutatás, bankcsoporton belüli adattovábbítás esetében lehetőséget kell teremteniük arra, hogy az egyes adatkezelési célokról külön-külön nyilatkozhassanak ügyfeleik, élve az információs önrendelkezési jogukkal.” (69. old.)

Már sok éve probléma, hogy az állampolgárokat kérést SMS-üzenetekkel, illetve telefonhívásokkal zaklatják. A NAIH beszámolója ezzel a témával is foglalkozik. Megállapítja, hogy a jogszabály szerint az előfizető számára külön költség nélkül biztosítani kell azt a jogot, hogy kérésére kimaradjon a nyomtatott vagy elektronikus névjegyzékből, illetőleg az előfizetői névjegyzékben feltüntessék, hogy személyes adatai nem használhatóak fel közvetlen üzletszerzés, tájékoztatás, közvélemény- vagy piac-

kutatás céljára, illetőleg azt is kérheti, hogy lakcímét csak részben tüntessék fel az előfizetői névjegyzékben. Továbbá az emberi beavatkozás nélküli, automatizált hívórendszer csak akkor alkalmazható közvetlen üzletszerzés, tájékoztatás, közvélemény- és piackutatás céljára, ha ehhez az előfizető előzetesen hozzájárult. A NAIH jelzi, hogy a hatályos jogszabályi rendelkezések azt az elvet követik, hogy az előfizetői névjegyzékben szereplő személyek adatait fel lehet használni közvetlen üzletszerzésre, kivéve akkor, ha a névjegyzékben való szereplés ellen tiltakoznak, vagy nyilatkoznak arról, hogy személyes adataikat közvetlen üzletszerzésre nem lehet felhasználni. Ez utóbbival kapcsolatban az elektronikus hírközlési szolgáltatók azt a gyakorlatot követik, hogy egy paragrafjelet helyeznek el a természetes személy neve mellett, ezzel jelezve, hogy telefonszámát nem lehet közvetlen üzletszerzésre felhasználni. Végül e témával kapcsolatban a NAIH összefoglalóan így nyilatkozik:

„A hatályos törvényi rendelkezések alapján az állampolgár elektronikus hírközlési szolgáltatója (vezetékes telefonszolgáltató, mobiltelefon-szolgáltató) felé tehet olyan nyilatkozatot, amelyben kéri, hogy adatai ne jelenjenek meg a névjegyzékben, vagy tüntessék fel mellette, hogy azok közvetlen üzletszerzés céljából nem használhatóak fel. Arra azonban nincs gyakorlati lehetőség, hogy az érintett valamennyi telemarketinggel foglalkozó gazdasági társaság felé általános nyilatkozattal éljen annak érdekében, hogy az ilyen tevékenységet folytató cégek egyike se léphessen vele kapcsolatba a telefonszámán keresztül. Ezt a szándékot juttatja kifejezésre a § jel feltüntetése a telefonkönyvben, amelyet minden szereplő köteles lenne tiszteletben tartani.” (72. old.)

A beszámoló e fejezete kiemelten foglalkozik még a társasházak dokumentumainak megismerhetőségi feltételeiről, mely szintén

visszatérő témája az országgyűlési beszámolóknak, valamint a hulladékgazdálkodó közszolgáltatók adatkezeléséről.

Miután a 2014. év választási év volt, a NAIH beszámolója mind a személyes adatok védelmével, mind a közérdekű adatok nyilvánosságával kapcsolatban tartalmaz tényfeltáró megállapításokat. A NAIH többek között az ajánlások gyűjtésével kapcsolatban mindenekelőtt felhívta az adatkezelők figyelmét arra, hogy a választópolgárok magánszféráját maximálisan kötelesek tiszteletben tartani, így tilos az ajánlásért bármely ellentételezés adása vagy elfogadása. Az ajánlóíveken szereplő adatokat a jelöltállításól eltérő célra felhasználni jogellenes, ezért az ajánlóívek másolása (sokszorosítása, fénymásolása), illetőleg duplikált adatbázisok létrehozatala az azokon szereplő információk alapján nem megengedett. Az ajánló személye nem hozható nyilvánosságra, bizalmasan kezelendő. A kampány célú adatkezelésekkel kapcsolatban a NAIH azt a követelményt fogalmazta meg, hogy amennyiben a központi névjegyzék, illetőleg a polgárok személyiadat- és lakcímnnyilvántartásának felhasználásával juttatnak kampányanyagokat a választópolgárok részére, akkor az érintetteket e tényről tájékoztatni kell. Továbbá a nyilvános terjesztésű telefonkönyvből azon magánszemély előfizetők telefonszámait lehet leválogatni és felhívni, akik hozzájárultak annak közzétételéhez, és előfizetői szerződésükben a közvetlen üzletszerzés, tájékoztatás, közvélemény- vagy piackutatás célját szolgáló megkeresések fogadását sem zárták ki. Nem megengedett ugyanakkor az olyan hívórendszer alkalmazása, amely a számgenerálás elvén alapszik, mert ennek használata sértheti a választópolgárok magánszféráját.

A beszámoló VI. fejezete szintén a vizsgálati ügyeket taglalja, de ezt ezúttal az információs szabadság oldaláról teszi. A 2014. évi választásokkal kapcsolatban érkeztek beadvá-

nyok a NAIH-hoz, mely ügyekben a NAIH általában az átláthatóság biztosítása érdekében foglalt állást. Például egy jegyző beadvánnyal fordult a NAIH-hoz, amelyben azzal kapcsolatban kért állásfoglalást, hogy nyilvánosságra hozható-e a helyi önkormányzati képviselők és polgármesterek általános választásán indulni szándékozó jelöltek lakcíme a nyilvántartásba vételükről szóló határozatban. A NAIH megállapította:

„Az egyéni szavazólapon a törvény értelmében csak a jelölt nevét kell feltüntetni, a lakcímet már nem. A Ve.⁴ tehát egyértelműen megjelöli azon adatkört, amely tekintetében a nyilvánosságnak mindenkor érvényesülnie kell: ez kizárólagosan a jelöltek neve. A NAIH ezért megállapította, hogy a független jelöltek lakcímeinek, mint személyes adatoknak, a megfelelő jogalap nélküli nyilvánosságra hozatala az illetékes választási bizottság internetes honlapján, sértette az érintettek személyes adatok védelméhez fűződő alapjogát.” (82. old.)

A fejezet még szól a kampányfinanszírozásról, melynek lényege, hogy a kampányfinanszírozás nagy része közérdekű, illetve közérdekből nyilvános adatnak minősül.

A korábbi országgyűlési beszámolók állandó témája volt az állami, illetve önkormányzati tulajdonú gazdasági társaságok adatainak nyilvánossága. E téma azért fontos, mert a korrupció elleni küzdelem egyik fontos területéről van szó. A NAIH beszámolója ezzel kapcsolatban néhány általános érvényű megállapítást tesz:

„Mind a NAIH, mind a Transparency International Magyarország azt tapasztalta,

⁴ 2013. évi XXXVI. törvény a választási eljárásról. http://net.jogtar.hu/jr/gen/hjegy_doc.cgi?docid=A1300036.TV

hogy a nyilvánossággal egyáltalán foglalkozó állami, önkormányzati cégek még azt az alapvető elvárást is vitatják, amely szerint a közpénzzel való gazdálkodás egyenértékű a minimálisan korlátozott nyilvánossággal. A NAIH már 2012-ben kifejtette a nyilvánosság érdekében tett előremutató jogalkotási lépés jelentőségét, ezt azonban több esetben nehezen tudta elfogadtatni a közpénzzel gazdálkodó gazdasági társaságok vezetőivel, jogtáncosaival.

A nyilvánossággal szembeni gyanakvás, hátráltató kötelezettségként való felfogása nem sokat változott. Az előremutató alaptörvényi szabályozás követése a várt, remélt változásokat egyelőre nem igazolta ... A nyilvánosság kiterjesztése tényleges megvalósításának tehát elsődleges feltétele az, hogy a közpénzből gazdálkodó, közpénzből alapított gazdasági társaságok elfogadják, hogy ők az információszabadság, mint alapjog kötelezettjei.” (84. old.)

„A közpénzekkel való gazdálkodás átláthatóságának, a tisztességes közélet megteremtése, mint közérdek, valamint az információszabadság, mint alapjog érvényesülése érdekében szükséges tehát a cégek üzleti érdekének, üzleti titkainak, mint magántitoknak a korlátozása. A nemzeti vagyonnal gazdálkodó cégek vezetőinek is ugyanolyan mértékben viselniük kell e feladatukkal összefüggő adataik nyilvánosságát, mint más közfeladatot ellátó személyeknek közfeladatuk ellátása során. A nemzeti vagyonnal gazdálkodó gazdasági társaságok működésének, gazdálkodásának nyilvánosságát garantáló jogszabályi alapok tehát megvannak, azok gyakorlati érvényesülése érdekében azonban még sok teendő van mind a NAIH, mind a kötelezettek részéről.” (87. old.)

Ezen túl a beszámoló szól a Paksi Atomerőművel kapcsolatos vizsgálat és jogszabályvéleményezés eredményeiről, az információ-

szabadság korlátairól (a nagy tömegű adatigénylések teljesíthetőségéről), az önkormányzatokkal kapcsolatos vizsgálatokról.

A beszámoló VII. fejezete a minősített adatok megismerésével kapcsolatos eljárásokat ismerteti, a VIII. fejezete a nemzetközi ügyeket és a társadalmi kapcsolatokat tárgyalja, a IX. fejezet pedig a NAIH költségvetéséről számol be.

A NAIH 2014. évi beszámolója már egy stabilnak mondható szervezet munkáját ismer-

teti, azonban az adatvédelem és információszabadság témája iránti alaposabb érdeklődést ez a viszonylag kis terjedelmű beszámoló bizonyára nem tudja kielégíteni. Ehhez ajánlom tanulmányozni a NAIH honlapján elérhető részletes dokumentumok (például állásfoglalások, ajánlások, közlemények) szövegét is.

Lakatos Miklós,

a KSH szakmai főtanácsadója

E-mail: Miklos.Lakatos@ksh.hu

Hírek, események

Szervezeti változás. *Vukovich Gabriellának*, a KSH elnökének döntése értelmében 2015. november 1-jei hatállyal a hivatal Elnöki főosztályán létrejött a Statisztikai koordinációs osztály. Ezzel a főosztály struktúrája a következőre módosult: Elnöki titkárság, Személyügyi és oktatási osztály, Nemzetközi kapcsolatok osztálya, Pályázati programok osztálya és Statisztikai koordinációs osztály.

Kinevezés. *Vukovich Gabriella*, a KSH elnöke 2015. november 1-jei hatállyal *Kárpáti Józsefet* nevezte ki az Elnöki főosztály Statisztikai koordinációs osztálya vezetőjének.

Közszolgálati jogviszonyban töltött ideje alapján jubileumi jutalomban részesült 2015. október vagy november hónapban 25 éves szolgálatért: *Bódiné Vajda Györgyi* (Tájékoztatási főosztály); 30 éves szolgálatért: *Bauer Károly* (Gazdálkodási főosztály), *Mike Sándor Zoltánné* (Szegedi főosztály), *Székely Gáborné* (Népesedési és szociális védelmi statisztikai főosztály), *Szigetvári Katalin* (Pécsi főosztály); 35 éves szolgálatért: *Patakiné Sárosi Zsuzsanna* (Vidékfejlesztési, mezőgazdasági és környezeti statisztikai főosztály); 40 éves

szolgálatért: *Krupincza Edit* (Debreceni főosztály), *Nyitrai József* (Győri főosztály), *Hazay Árpád* (Pécsi főosztály).

Az ENSZ¹ Fenntartható Fejlődési Csúcstalálkozóját 2015. szeptember 25-e és 27-e között rendezték New Yorkban, melyen a 193 tagállam egyhangú döntéssel elfogadta a „Világunk átalakítása: a fenntartható fejlődés 2030-ig szóló programja” elnevezésű globális cselekvési keretet. A 2000-ben jóváhagyott millenniumi fejlesztési célokra épülő megállapodás az első egyetemes, átfogó cselekvési menetrend. 17 fenntartható fejlődési cél és 169 konkrét ún. célpont kitűzésével a fenntartható fejlődés három (környezeti, társadalmi és gazdasági) dimenziója közötti egyensúly kialakítására törekszik. Nemcsak a fejlett országok „hazai” teendőit részletezi, hanem olyan feladatokat is, amelyek a kevésbé fejlettek támogatását célozzák. A csúcstalálkozón részt vevő magyar delegációt *Áder János* köztársasági elnök vezette. A program több idegen nyelven is olvasható a http://www.un.org/ga/search/view_doc.asp?sy

¹ ENSZ: Egyesült Nemzetek Szervezete.

mbol=A/RES/70/1&referer=/english/&Lang=F honlapon.

Globális partnerség jött létre az ENSZ vezetésével a fenntartható fejlődési adatkért. A kormányok, fejlesztési ügynökségek, civil szervezetek, magáncégek és alapítványok részvételével 2015. szeptember 28-án, az ENSZ Fenntartható Fejlődési Csúcstalálkozó-ját követően indult kezdeményezés a fenntartható fejlődési célok megvalósulását szolgálja, a politikai kötelezettségvállalások ösztönzését és fenntartását, a prioritások és normák összehangolását, a kapacitások növelését, az innováció előmozdítását és a bizalom kiépítését segítve. A hároméves induló szakaszban a partnerség működését egy ideiglenes irányító csoport vezeti, illetve egy titkárság koordinálja; fogadó szervezete pedig az ENSZ Alapítvány. A kezdeményezéshez többek között az ISI² is csatlakozott.

„Adatok a fenntartható fejlődésért” mottóval adott ki nyilatkozatot az Amerikai Statisztikai Szövetség, a Királyi Statisztikai Társaság, az ISI és közel 50 más statisztikai szervezet (köztük az MST³ is) a Statisztika Világnapja alkalmából. A közös állásfoglalás aláírói az ENSZ által nemrégiben elfogadott fenntartható fejlesztési célok megvalósulásának érdekében további erőfeszítésekre szólítottak fel, és hangsúlyozták, a nemzeti statisztikai rendszerek központi szerepet töltenek be az új feladatokkal járó adatkihívások kezelésében.

A 2015. évi közgazdasági Nobel-díjat Angus Deaton, a Princeton Egyetem közgazdaságtan és nemzetközi kapcsolatok professzora nyerte el a fogyasztással, szegénységgel és jóléttel foglalkozó elemzéseiről. „A jólétet és a

szegénység csökkentését elősegítő gazdaságpolitika kialakításához először az egyének fogyasztási döntéseit kell megértenünk. Ezt Angus Deaton mindenki másnál jobban segítette. Az egyéni döntések és az aggregált eredmények összekapcsolásával kutatása segített átalakítani a mikro- és makróökonómiát, valamint a fejlesztés-gazdaságtant.” – fogalmazott a Svéd Királyi Tudományos Akadémia a bejelentéskor. Elismert kutatásai során három központi kérdésre keresett választ. Mennyit költenek a fogyasztók a különböző árucikkekre? Mennyit költenek el, és mennyit takarítanak meg a társadalmi jövedelmekből? Hogyan mérjük és elemezzük legjobban a jólétet és a szegénységet? Az 1945. október 19-én, Edinburghben született közgazdász nevéhez fűződik az ún. Deaton-paradoxon is: a fogyasztás szintje kevésbé változékony, mint amennyire az aktuális jövedelem ingadozik.

A Szerb Statisztikai Hivatal szakértőivel találkoztak a KSH munkatársai 2015. szeptember 21-én és 22-én. Az IPA⁴ 2012 támogatói program keretében zajlott látogatáson a vendégek betekintést kaptak a statisztikai népesség-nyilvántartás témakörébe. Tájékoztatót kaptak a 2011. évi népszámlálás többszörös adatgyűjtési módszeréről, különös tekintettel az internetes adatgyűjtésre, valamint a KSH címregiszteréről, a geokódok használatáról és a térképalapú népszámlálási adatközlésekről.

A Moldáv Statisztikai Hivatal szakértői delegációját fogadta a KSH 2015. szeptember 29-e és október 1-je között külkereskedelemsziszatika témában. Először *Laczka Éva* gazdaságstatisztikai elnökhelyettes köszöntötte a vendégeket, és bemutatta számukra a külkereskedelem magyar gazdaságban betöltött

² ISI (International Statistical Institute): Nemzetközi Statisztikai Intézet.

³ MST: Magyar Statisztikai Társaság.

⁴ IPA (Instrument for Pre-accession Assistance): előcsatlakozási támogatási eszköz.

szerepét, a KSH történetét, valamint a magyar intézmény előtt álló kihívásokat. Megnyitója után *Szabó Péter* főosztályvezető is üdvözölte a delegációt, majd a Külkereskedelmi makroelszámolási osztály munkatársai nyújtottak betekintést a magyar szolgáltatás-külkereskedelmi statisztikába.

Elnöki szintű delegáció érkezett a Kínai Népköztársaság Statisztikai Hivatalából
2015. október 1-jén a KSH-ba. *Vukovich Gabriellának*, a KSH elnökének betegsége miatt a találkozót *Németh Zsolt* társadalomstatisztikai elnökhelyettes nyitotta meg, majd *Wang Baoan*, a kínai hivatal elnöke köszöntötte a vendéglátókat. A magyar szakértők több témát is bemutattak a küldöttségnek: a hazai statisztikai rendszer tevékenységét, a KSH jogi és működési környezetét, koordinációs felelősségét, az adatgyűjtés-módszertani rendszert, a nemzeti számlák készítésének történetét, eredményeit és dilemmáit, továbbá a hivatali minőségbiztosítási keretrendszert és az általános statisztikai folyamatmodell alkalmazását.

A Cseh Statisztikai Hivatal szakértőit fogadták a KSH-ban 2015. október 8-án és 9-én. A találkozó során a külföldi delegáció áttekintést kapott a háztartási költségvetési felvétel és a változó életkörülmények adatfelvétel integrálásának módszertani lépéseiről, valamint az így kialakult háztartási költségvetési és életkörülmény adatfelvétel felépítéséről, feldolgozásáról és felhasználási lehetőségeiről.

Az Európai Rendőrakadémia Európai Közös Mesterprogramjának „Európai rendőri tevékenység” című képzése részeként látogattak el EU⁵-s tagállamok magas rangú rendőri vezetői 2015. október 14-én, a KSH-ba. A Nemzeti Közzolgálati Egyetem és a hivatal együttműködése keretében létrejött

⁵ EU: Európai Unió.

rendezvényt *Németh Zsolt*, a KSH társadalomstatisztikai elnökhelyettese nyitotta meg. A résztvevők az Európai Statisztikai Rendszeréről, az adat-előállítás folyamatáról, az uniós adatbázisok használatáról és a bűnügyi statisztikákhoz való hozzáférési lehetőségekről kaptak tájékoztatást, továbbá épületséta során megismerkedhettek a hivatal történetével és működésével is.

A világ statisztikusai 2015. október 20-án ünnepelték a Statisztikai Világnapot.
Mint már arról egy korábbi számunkban beszámoltunk, az ENSZ közgyűlése 2015. június 3-án fogadta el azt a magyar betérjesztést, hogy 2015. október 20-a (és a továbbiakban minden ötödik évben ugyanez a nap) a Statisztika Világnapja legyen. Az ünnepi rendezvények kiváló alkalmat nyújtottak a statisztikai közösségek számára, hogy bemutassák eddig elért munkájukat és eredményeiket. A KSH, a világnap alkalmából 2015. szeptember 18-án indított eseménysorozatának részeként, 2015. október 20-án és 21-én szakmai konferenciát rendezett „Innováció és kooperáció – a hivatalos statisztika megújításáért” címmel a Magyar Tudományos Akadémia székházában. Ennek keretében felvázolta a hivatalos statisztika átalakítására vonatkozó fejlesztési terveit az új európai uniós statisztikai rendelettel összhangban. A tervezett fejlesztések célja válaszkérés napjaink olyan kihívásaira, mint a modernizáció és a bővülő felhasználói csoportok változó igényei. Az ülés napirendjén fontos témák szerepeltek, például a statisztika jövője, a hivatalos statisztika megújításának lehetőségei, a kutatók és a hivatalos statisztika, valamint a nemzeti statisztikai hivatalok közötti együttműködés és a központi banki statisztikák. Az előadások mellett kerekasztal-beszélgetésekre is sor került, melyek keretében a résztvevők számos (a statisztika fenntarthatóságával, a mikroadatok cseréjével, az emberi

erőforrással, a statisztikusok képzésével és a kommunikációval kapcsolatos) kérdést megvitatathattak. A rendezvény online közvetítését a KSH honlapján az érdeklődők a világ minden táján nézhették. (A konferenciáról részletes beszámolót a következő számunkban olvashatnak.)

„Miben élünk? – Lakásvizonyok, 2015” elnevezéssel hajtott végre adatfelvételt a KSH 2015. október 10-e és november 30-a között a lakosság lakáskörülményeinek felmérésére. Ezt megelőzően 1999-ben és 2003-ban került sor hasonló, átfogó vizsgálatra. A véletlenszerűen kiválasztott, majd felkeresett 20 ezer lakásban élő személyek önkéntesen, név nélkül válaszolhattak a feltett kérdésekre. A felmérés témakörei a következők voltak: a lakások lakói; az épület jellemzői; a lakás jellemzői, helységei, felszereltsége; épületállapot, épület- és lakásfelújítás; lakásköltségek, lakáshitelek; lakástörténet, lakásmobilitás; lakáspiaci helyzet és lakásváltoztatási tervek. Az adatfelvétel eredményeit 2016 első félévében teszi a hivatal honlapján közzé, majd részletes elemzések is készülnek.

A KSH is részt vett kiállítóként az Országos Mezőgazdasági és Élelmiszeripari Kiállításon és Vásáron 2015. szeptember 23. és 27. között. A hivatal standjához látogatók tájékoztatást kaptak az intézmény munkájáról, fejleszthették statisztikai jártasságukat, és tesztelheték a 2016. évi gazdaságszerkezeti összeírás kérdőívét.

„Térökonometria a gyakorlatban” címmel szervezett a Miskolci Egyetem Gazdaságtudományi Kara a KSH-val és a *Területi Statisztika* folyóirattal közösen konferenciát 2015. október 7-én, az egyetemen. A rendezvényen a következő előadások hangzottak el: 1. Kistérségi vándorlások modellezése: a demográfiai és egy lehetséges térbeli megkö-

zelítés (*Bálint Lajos*, a KSH NKI⁶ tudományos munkatársa), 2. A gravitációs modell kalibrálása (*Dusek Tamás*, a Széchenyi István Egyetem egyetemi docense), 3. Monetáris impulzusok területi hatásai (*Kocziszky György*, a Monetáris Tanács tagja), 4. A konvergencia lokális szintű mérése (*Kotosz Balázs*, a Szegedi Tudományegyetem főiskolai docense), 5. A fejlesztési támogatások területi dimenziói a turizmus példáján – egy térökonometriai elemzés tanulságai (*Major Klára*, a HÉTFA Kutatóintézet kutatási igazgatója), 6. Munkaerőpiaci hátrányok területi vonatkozásai Észak-Magyarországon (*Síposné Nádori Eszter*, a Miskolci Egyetem egyetemi adjunktusa), 7. A saját jogú nyugdíjak alakulásának területi sajátosságai (*Szilágyi Roland*, a Miskolci Egyetem egyetemi docense és *Varga Beatrix* egyetemi docens, a Miskolci Egyetem Üzleti Információgazdálkodási és Módszertani Intézetének intézetigazgató-helyettese), 8. Az EU-keretprogramokban való részvétel és a regionális innovációs teljesítmény Kelet-Közép-Európában (*Varga Attila*, a Pécsi Tudományegyetem egyetemi tanára és *Sebestyén Tamás*, a Pécsi Tudományegyetem egyetemi adjunktusa). A rendezvényen átadták a *Területi Statisztika* folyóirat 2014. évi „legnívósabb tanulmányainak” díjait is.

Négynapos programsorozatot indított 2015. november 2-án, Szegeden a KSH, az SZTE GTK⁷, az Erste Bank, a Csongrád Megyei Kereskedelmi és Iparkamara, illetve az Árkád Szeged Bevásárlóközpont a takarékosági világnap alkalmából. Ennek „nyitányaként” egy kerekasztal-beszélgetésre került sor az egyetemen „Válságok társadalmi és gazdasági hatásai” címmel, melynek házigazdája

⁶ NKI: KSH Népszégtudományi Kutatóintézet.

⁷ SZTE: Szegedi Tudományegyetem Gazdaságtudományi Kara.

Kovács Péter, az SZTE GTK általános dékánhelyettese, tanszékvezető egyetemi docens, résztvevői *Kovács Árpád*, a Költségvetési Tanács elnöke, *Kovács Levente*, a Bankszövetség főtitkára, *Krisán László*, a KAVOSZ Vállalkozásfejlesztési Zrt. vezérigazgatója, *Sándor György*, a KSH elnöki főtanácsadója, *Szapáry György*, a Magyar Nemzeti Bank korábbi alelnöke, valamint *Hári Tibor*, a Baptista Szeretetszolgálat munkatársa voltak. A moderátori szerepet *Holka László*, a KSH vezető-főtanácsosa töltötte be. Az eszmecsere elején Sándor György tartott ismertetőt az 1929-es válság demográfiai következményeiről, valamint a 2008. évi krízis háztartásokra gyakorolt hatásairól. Majd a meghívottak a válságok természetéről, bankéletről és kisvállalkozásokra kifejtett hatásáról, a költségvetési fegyelem szerepéről, a rászorulóknak számának gyarapodásáról, az adakozási kedv megcsappanásáról, illetve a takarékoskodás és a pénzügyi kultúra kapcsolatáról beszélgettek. Mindannyian egyetértettek abban, hogy napjainkban az egyre gyorsabb lefolyású válságok miatt megingó bizalom csak lassan áll helyre.

Ugyancsak a programsorozat részeként, az ország több pontjáról – Budapestről, Debrecenből, Egerből, Győrből, Hatvanból, Hódmezővásárhelyről, Kiskunfélegyházáról, Paksról, Szegedről, Szentgotthárdról és Zalaegerszegről – érkezett csapatok szálltak versenybe a STAT WARS II. középiskolai vetélkedőn, 2015. november 4-én. Statisztikai, pénzügyi, logikai

és ügyességi feladatok útján ismerkedtek a takarékoskodás kérdéskörével, illetve a gazdasági világválságok társadalmi és gazdasági hátterével. A 16 csapat megmérettetését a Zalaegerszegről érkezett Csány László Közgazdasági Szakközépiskola diákjai nyerték a Győri Deák Ferenc Közgazdasági és Informatikai Szakközépiskola, valamint a Budapesti Gazdasági Szakképzési Centrum II. Rákóczi Ferenc Közgazdasági Szakközépiskola tanulói előtt. A lakosság statisztikai jártasságának fejlesztését szolgálta, hogy a bevásárlóközpont látogatói november 7-éig az ott kiállított plakátok statisztikatörténeti, demográfiai, pénzügyi és gazdasági információira épülő nyereményjátékban vehettek részt.

„Öngyilkosságok a XX. század elején, Magyarországon” címmel került sor 2015. november 3-án *Bálint Lajos*nak, a KSH Népeségtudományi Kutatóintézet tudományos munkatársának az előadására az MST Statisztikatörténeti Szakosztályának szervezésében, a KSH Könyvtár Thirring Lajos-termében.

„Számok világa mindenkinek” címmel rendezett programot az MST, a KSH Könyvtár és a Pannon Egyetem Georgikon Kara a Szentendrei Skanzenben, 2015. november 7-én és 8-án. A szervezők, kapcsolódva a Szent Márton Újborfesztyvál és Libator rendezvényéhez, a Statisztikai Világnapról és a társaság 25 éves jubileumáról emlékeztek meg.

Folyóiratszemle

Martin, D. – Cockings, S. – Harfoot, A.:

Földrajzi keretrendszer fejlesztése népszámlálási munkahelyadatokhoz

(Development of a Geographical Framework for Census Workplace Data.) – *Journal of the Royal Statistical Society*. 2013. Vol. 176. Part 2. pp. 585–602.

A mai nemzetközi gyakorlatban azt figyelhetjük meg, hogy a legtöbb nemzeti statisztikai hivatal a népszámlálásból származó munkahelyi adatokat lakóhelyi körzetek szerinti bontásban közli. Ez azonban torzításokhoz vezet, mivel a mai modern társadalmakban a lakóhely és a napközbeni aktivitás szintere általában különbözik. A kereskedelmi és az egyéb foglalkoztatási központokban kevesen laknak, míg a jellemzően külvárosi lakóövezetekben kevesen dolgoznak. Mindebből egyenesen következik, hogy a lakóhelyi körzetek szerinti részletes munkahelyi statisztikák publikálása gyakran adatvédelmi korlátokba ütközik. Sok lakóhelyi körzetben ugyanis olyan alacsony a munkavállalók száma, hogy nem éri el azt a nemzeti statisztikai hivatal által figyelembe vett küszöbértéket, ami felett az adatok már közölhetők lennének. Továbbá az a probléma is felmerülhet, hogy bár egy lakóhelyi körzet elegendő számú munkavállalót tartalmaz, viszont ezekre részletes statisztikák nem állnak rendelkezésre.

A szerzők e jelenséget a 2001-es angol és walesi népszámlálási adatokon illusztrálják. Ebben a népszámlálásban a közölt adatok legkisebb földrajzi egysége az ún. outputkörzet volt. Létrehozásuk az automatizált zónakialakítás módszerével történt, és a lakóhelyhez kapcsoló-

dott. Az outputkörzeteket úgy képezték (összesen 175 434-et), hogy a címetek poligonokkal kötötték össze, oly módon, hogy az egy irányítószám alá tartozó tömbök egy poligont alkotnak. Adatvédelmi okokból egy körzet legalább 40 háztartást és 100 személyt, átlagosan pedig 297 személyt tartalmaz. Az outputkörzetek képzéséhez nem használták fel azokat az irányítószámokat, melyeken kizárólag munkahelyek találhatóak. Ezekben ugyanis gyakran kizárólag nagyvállalatok vagy vállalatcsoportok vannak, lakóhelyek szinte egyáltalán nincsenek. A 2001-es népszámlálásban az automatizált zónakialakítás módszerével létrehozott outputkörzetek konzisztens lakóhelyi népességet eredményeztek. Ezek az outputkörzetek azonban elfedik azt a munkahelyi népességben tapasztalható területi változékonyságot, amit a nagy foglalkoztatási körzetekben dolgozók okoznak. A szerzők megvizsgálták a munkahelyek outputkörzetek szerinti eloszlását, és egy szélsőségesen balra ferde eloszlással szembesültek. Ez azt jelenti, hogy a kevés munkavállalót tartalmazó outputkörzetek dominálnak. Ez magyarázza, hogy a 2001-es angol és walesi népszámlálásból az outputkörzetek szintjén mindössze négy egyváltozós (kor, szocio-ökonomiai besorolás, társadalmi státus, munkahelytől vett távolság) statisztikát publikáltak adatvédelmi okokból; miközben a munkaszociológusok számára fontos munkahelyi iparágat, a foglalkozás típusát és a munkába járás módját nem közölték. Ez az eredmény tehát nem más, mint a lakóhelyek és a munkahelyek éles földrajzi elkülönülésének empirikus bizonyítéka.

A szerzők megemlítik egy amerikai és egy ausztrál példát, amikor nem csak lakóhelyi

körzetekre vonatkozóan közöltek munkahelyi adatokat. Ezek azonban inkább a szállítás modellezésére alkalmasak, nem fedik le az egész országot, nincsenek integrálva a statisztikai tájékoztatási rendszerekbe, illetve esetükben utólagos manuális feldolgozás is szükséges.

Összességében tehát elmondható, hogy a részletes munkahelyi statisztikák hiánya jelentős mértékben megnehezíti a foglalkoztatási szerkezet és az ingázás elemzését, ami új földrajzi körzetek kidolgozását tette indokolttá.

A szerzők célja tehát olyan földrajzi körzetek kialakítása, amelyek tükrözik a munkahelyek elhelyezkedését, így az adatvédelmi akadályok is elhárulnak a részletesebb munkahelyi adatok közlése előtt. Ezeket a földrajzi entitásokat munkahelyi körzetnek nevezték el.

A munkahelyi körzetek létrehozását a 2001-es angol és walesi népszámlálásból kiválasztott öt mintaterületen (City of London, Tower Hamlets, Nottingham, Southampton, Suffolk Coastal) tesztelték. Ezek összesen körülbelül 800 ezer személyt tartalmaztak.

A munkahelyi körzetekkel szemben a következő elvárásokat támasztották:

1. Hierarchikusan tagolódnak az output-körzetek által megszabott struktúrába. Ezt a kritériumot úgy lehet teljesíteni, ha a munkahelyi körzetek az outputkörzetek összevonása vagy feldarabolása révén jönnek létre. Amennyiben a kétféle földrajzi struktúra között átfedés lenne, az újabb adatvédelmi problémákat vetne fel.

2. Minden munkahelyi körzet tartalmazzon legalább egy minimális számú munkahelyi népességet. Ez az outputkörzetek esetében leirtakkal analóg módon biztosítható: minden munkahelyi körzetnek legalább 100 munkavállalót kell tartalmaznia.

3. Minden munkahelyi körzet tartalmazzon egy minimális számú irányítószámot. Míg a munkahelyi népesség adatvédelme könnyedén biztosítható a 2. kritériumban leírt módon, a

vállalatok azonosításának megakadályozása nem ennyire triviális. A népszámlálásban ugyanis a vállalatok explicit módon nem jelennek meg. Mivel a népszámlálási adatok külső vállalati regiszterrel sem kapcsolhatók össze, ezért a vállalatok védelmére egy újszerű megoldást kellett kidolgozni. Könnyen belátható, hogy ha a munkaerőre vonatkozó adatokat a legkisebb közölhető egységre, az irányítószámok szintjén publikálnák, akkor annak ellenére, hogy sok esetben egy irányítószám nem azonosít be egyetlen vállalatot, bizonyos esetekben ez mégis előfordulhatna. Azért, hogy az irányítószámok egyedi jellemzőit ne lehessen felfedni, egy munkahelyi körzetnek legalább 3 irányítószámot kell tartalmaznia, és bármely irányítószámot elhagyva egy adott körzetből, legalább 100 munkavállalónak kell maradnia a munkahelyi körzetben.

4. Adott munkahelyi körzet iparági struktúrája lehetőleg minél homogénebb legyen. Ez a kritérium akkor teljesül, ha a területen belüli korreláció értéke magas.

5. A munkahelyi körzetek alakjának minél kompaktabbnak kell lenniük, amit úgy mérnek, hogy a kerület négyzetét elosztják a munkahelyi körzet területével.

A munkahelyi körzetek tehát az outputkörzetekből jönnek létre úgy, hogy ahol a munkahelyi népesség nem éri el a küszöbértéket (a 100 főt), a szomszédos outputkörzeteket összevonták, míg ahol túl nagy volt a munkahelyi népesség (legalább 625), azokat pedig szétदारolták. A cél a 250 munkavállalót tartalmazó munkahelyi körzetek létrehozása volt. Az öt mintaterület munkahelyi népességre vonatkozó leíró statisztikáit (minimum, maximum, átlag, szórás) összevetették az output- és a munkahelyi körzetek szintjén. Azt tapasztalták, hogy az utóbbiak kialakításával sikerült sokkal egyenletesebbé tenni a munkahelyi népesség eloszlását: így már minden körzet tartalmazott

munkavállalókat, miközben a kiugró maximum értékek jelentősen csökkentek (például City of Londont tekintve a legtöbb munkavállalót tartalmazó outputkörzet majdnem 80 ezres értéke a munkahelyi körzet esetében alig több, mint 15 ezer). A munkahelyek egyenletesebb eloszlását igazolja vissza az is, hogy az egy munkahelyi körzetre jutó munkahelyi népesség általában nőtt (kivéve City of London), miközben a szórás csökkent.

Az adatvédelmi kérdéseket Southampton példáján illusztrálták a szerzők. Elemezték azokat az irányítószámokat, amelyekhez a legnagyobb munkahelyi népesség tartozott, mivel várhatóan ezek alapján lehet a legnagyobb foglalkoztató vállalatokat beazonosítani. Ha megnézzük azt a 20 irányítószámot, amelyhez a legnagyobb munkahelyi népességszám kapcsolódott, majd ezt összevetjük annak az outputkörzetnek a munkahelyi népességével, amelybe az adott irányítószám tartozik, akkor csak négy esetben tapasztalható az, hogy az adott irányítószám alatt foglalkoztatottak száma közel volt (legfeljebb 100 fő az eltérés) az outputkörzet teljes foglalkoztatotti létszámához. Ebben az esetben azt lehet feltételezni, hogy az adott outputkörzet foglalkoztatottsága valószínűleg egyetlen vállalathoz köthető. A megfelelő népszámlálási rekordokból azonban az látszik, hogy ez a foglalkoztatottság igen sok iparág között oszlik meg, és a foglalkoztatói létszámok is nagymértékben szórnak. Vagyis hiába feltételezhető néhány esetben, hogy egy outputkörzet foglalkoztatotti létszáma egyetlen vállalathoz köthető, maga a vállalat közvetlenül nem azonosítható be a népszámlálásban szereplő egyéb adatok segítségével. A foglalkoztatottak a munkahelyük iparágát inkább a saját munkájuk alapján ítélik meg, és a foglalkoztatotti létszám is inkább a saját részlegük létszámát tükrözi, semmint az egész vállalatét. Meg kell továbbá jegyezni, hogy éppen ezek a változók (például a

foglalkoztatotti létszám), amelyek esetén a legnagyobb válaszmegtagadás figyelhető meg.

Összességében elmondható, hogy a munkahelyi népességet illetően a munkahelyi körzetek sokkal előnyösebb statisztikai tulajdonságokkal rendelkeznek, mint az outputkörzetek. Előbbiek létrehozásának fő célja tehát teljesült: ezekre vonatkozóan részletesebb munkaerő-adatokat lehet publikálni, mint a korábbi lakóhely alapú outputkörzetekre. A módszer általános jellegű: az automatizált zónakialakítás módszerével más térstatisztikai eloszlási problémák megoldása is lehetséges.

Kristóf Péter,

a KSH fogalmazója

E-mail: Peter.Kristof@ksh.hu

Brandt, M.:

Az európai mikrodadat-hozzáférés fejlesztése

(Improvement of Access to European Microdata.) – *Joint UNECE/Eurostat Work Session on Statistical Data Confidentiality*. 28–30 October 2013. Ottawa.

Az ESR¹ legnagyobb feladata, hogy ellássa az Európai Unió intézményeit, a tagországokat, valamint a nyilvánosságot megbízható információkkal a társadalomról, gazdasági környezetről és fejlesztésekről. Emiatt fontos a statisztikai adatok gyűjtése, feldolgozása és elemzése a tagországokban. A kérdőívek, adatfelvételek harmonizálása, az előkészítő folyamatok és az adatok szállítása az Európai Unió 28 tagállama számára bonyolult procedúra. Szükséges az információk összehangolása a tagállamok között, és az Európai Unióra vonatkozó összesített adatok előállítását. Ebből következik az is, hogy

¹ ESR: Európai Statisztikai Rendszer.

ezek az adatok értékes és egyedi források a nemzetközi kutatók számára. Hasznosságuk nem merül ki a standardizált aggregált táblák publikálásában. Sokkal több lehetőség nyílik a legjobb gyakorlatok, különbségek feltérképezésére a tagállamok között.

Az ilyen módon előállított aggregált adatok azonban nem elég részletesek a legtöbb kutatónak, fontosabb számukra, hogy az eredeti mikroadat állományhoz hozzáférjenek. Erre jelenleg csak egy lehetőség van: a kutatónak el kell utazni az Eurostat Kutatószobájába Luxemburgba. Pontosan ezt a problémát hivatott kezelni a DARA² elnevezésű ESR-projekt, melynek olyan rendszer létrehozása a célja, amely segítségével a nemzeti statisztikai intézeteknek és a tagállamoknak lehetőségük van távolról is hozzáférni az Eurostat adataihoz. A tapasztalatok azt mutatják, hogy a kiépített infrastruktúra nehezen használható, ellenben lehetőséget biztosít a kutatók számára.

A projekt 2011 októberében indult, kétéves időtartammal. Fő résztvevői voltak a Francia Közgazdasági és Statisztikai Iskolák Csoportosulása, a magyar Központi Statisztikai Hivatal, az Egyesült Királyság Nemzeti Statisztikai Hivatala, a Portugál Statisztikai Hivatal, a Német Szövetségi Statisztikai Hivatal és a Német Szövetségi Állami Statisztikai Intézet.

Az írás bemutatja a projekt előzményeit, tartalmát, eredményeit, valamint röviden a további lehetőségekre is kitér.

A DARA-projektet a „Decentralised Access to EU Microdata Sets” (decentralizált hozzáférés az európai uniós mikroadatokhoz) elnevezésű projekt előzte meg, melyben azt vizsgálták, lehetőség van-e a tagországok kutatószobáiban távoli hozzáférést biztosítani bizalmas adatokhoz. Így született meg az ajánlás, hogy hozzanak létre egy ilyen rend-

szert. A részletek kidolgozásáért, hogy miként jöjjön létre ez a biztonságos csatorna az Eurostat és a tagállamok között, már a DARA-projekt volt felelős, melynek tagjai első körben meghatároztak egy koncepciót, technikailag hogyan nézzen ki ez a rendszer. A megvalósítást a VIP-SICON³-projekt keretében biztosították. Az Eurostat és egy külsős cég volt felelős a mikroadatokhoz való hozzáférés kiépítéséért. Bár ez a rendszer több feladatot hivatott ellátni, elsődleges mégis az, hogy biztonságos megosztást és hozzáférést biztosítson az európai bizalmas adatokhoz a tagországokban. Ez a rendszer lehetővé teszi az adatok megtekintését, feldolgozását és elemzését.

Az eredeti terv az volt tehát, hogy létrehozzanak a tagállamok és az Eurostat, mint csomópont között egy hálózatot. Azonban a VIP-SICON-rendszer nem készült el a DARA-projekt végéig, ezért a munkacsoport kitalált egy alternatív megoldást, melynek keretében a fő csomópontot az Eurostat helyett Franciaország jelentette. Ez a megoldás a meglévő francia távoli hozzáférési rendszeren alapul, melyet átterveztek az európai rendszerre. Ez volt a bizonyítéka a megvalósíthatóságnak, ám nem használható tényleges feldolgozáshoz. A rendszerben erős felhasználói azonosítást hoztak létre, melynek lényege az adatok biztonsága, hogy a felhasználó ne tudjon adatot eltávolítani, lemásolni a rendszerből. Így a hozzáférés biztosításához öt szerepkört határoztak meg:

1. Kutató: hozzáfér a mikroadatokhoz és statisztikai szoftverekhez Windows környezetben.

2. Támogató munkatárs: a kapcsolatot biztosítja a kutató és más felhasználók között.

³ VIP-SICON (Vision infrastructure project on secure infrastructure for CONFidential data access): Vision infrastruktúra-projekt a bizalmas adatokhoz való hozzáférés biztosítására.

² DARA (decentralised and remote access to confidential data): decentralizált és távoli hozzáférés bizalmas adatokhoz.

3. Kimenetellenőr: a kutató által kinyert eredményeket látja el felfedés elleni védelemmel.

4. Adatmenedzser: az adatbázisokat és a metaadatokat készíti elő a kutatók számára.

5. Információtechnikai menedzser: a rendszer fejlesztéséért és a karbantartásáért felelős.

Egy felhasználóhoz több szerepkör, illetve egy szerepkörhöz több felhasználó is tartozhat. A rendszer legnagyobb előnye, hogy az adatok nem kerülnek ténylegesen a tagországokba, azok mindig egy biztonságos szerveren van-

nak, és csak a tagországok biztonságos kutatószobáiból érhetőek el.

A DARA-projekt bebizonyította, hogy a gyakorlatban is lehetőség van biztonságos, felhasználóbarát módon hozzáférni az európai mikroadatokhoz. Persze sok még a tennivaló: ki kell építeni az Eurostatban a megfelelő központot, valamint a tagállamoknak is létre kell hozniuk a fogadásra alkalmas környezetet.

Pécs Mária,

a KSH tanácsosa

E-mail: Maria.Pecs@ksh.hu

Kiadók ajánlata

CAREY, P. [2015]: *Data Protection. A Practical Guide to UK and EU Law. Fourth Edition.* (Adatvédelem. Gyakorlati útmutató az Egyesült Királyság és az EU jogszabályaihoz. Negyedik kiadás.) Oxford University Press. Oxford.

Az Egyesült Királyság egyik vezető adatvédelmi szakértője által írt értékes kézikönyv gyakorlati megoldásokat kínál az Egyesült Királyság és az Európai Unió adatvédelmi törvényével kapcsolatban felmerülő problémákra. A teljesen átdolgozott és kibővített negyedik kiadás újdonságának számít, hogy benne korábban még nem tárgyalt esetek és iránymutatások, új módszerekkel foglalkozó anyagok, valamint a viselkedésalapú hirdetésről és a magánélet védelmére vonatkozó hatásvizsgálatról szóló fejezetek szerepelnek. A szerző jelentős mértékben átdolgozta a CCTV-t¹ tárgyaló fejezetet is, és új információt közöl a közösségi médiáról, illetve a Leveson-jelentésről.

¹ CCTV (closed-circuit television): zárt rendszerű televízió.

A fejezetek az általános európai adatvédelmi rendelettervezet valószínű hatásával kapcsolatos szempontokat az Európai Bizottság által 2012-ben kiadott rendelettervezet és az Európai Parlament által közzétett módosítási javaslatok elemzésével sorakoztatják fel.

JAY, R. P. (ed.) [2015]: *Data Protection & Privacy 2016.* (Adatvédelem és -biztonság 2016.) Law Business Research Ltd. London.

Napjainkban, párhuzamosan az információk nélkülözhetlenné válásával a kereskedelmi és közéletben, az adatvédelemről szóló jogszabályok is egyre fontosabb szerepet töltenek be. A kötet a világ különböző tájain hatályban vagy előkészítés alatt levő, legfontosabb adatvédelmi és -biztonsági törvények közül többel is foglalkozik. Olyan témákat tárgyal, mint az adatvédelmi szabályok megsértése, a mentességek, a témát érintő egyéb jogszabályok, a PII-formátumok², a jogszerű

² PII (personally identifiable information): személyesen azonosítható információk.

feldolgozás, az értesítések, a biztonság, a biztonsági kötelezettségek és azok megsértése, a regisztrációs formalitások, a büntetések, az internethasználat és az elektronikus kommunikációs marketing.

CREMEAN, D. [2015]: *Freedom of Information: What You Need to Know*. (Amit az információszabadságról tudnunk kell.) LexisNexis Australia. Chatswood.

A kötet az ausztrál információszabadság tárgykörébe vezeti be az olvasókat, gyakorlati segítséget nyújtva számukra, illetve megismertette őket az ausztrál államszövetségi és állami információszabadsággal foglalkozó jogszabályokkal.

A könyv első részében a szerző teljes körű áttekintést ad a témáról, amit a „Formák és irányadó esetek” című második, illetve az információszabadságról szóló 1982. évi törvénnyel és az információvédelmi irányelvekkel foglalkozó részek követnek. Nélkülözhetetlen referenciamunka mindazok számára, akik kormányzati vagy hatósági szerveknél, egyesületeknél, egyetemeken foglalkoznak a témával.

KATZ, J. – LINDELL, J. [2015]: *Introduction to Modern Cryptography: Principles and Protocols*. (Bevezetés a modern kriptográfiába: alapelvek és protokollok.) Chapman & Hall/CRC. Boca Raton.

A kriptográfia fontos szerepet játszik az adatok védelmének, integritásának, valamint a számítógépes hálózatok biztonságának megőrzésében. A kötet a modern titkosításba nyújt pontos, ám könnyen érthető betekintést, a hivatalos meghatározásokra, hipotézisekre és bizonyításokra helyezve a hangsúlyt.

A szerzők először a szakterület alapelveit, többek között a biztonság modern számítástechnikai megközelítését ismertetik, amellyel sikeresen kezelhetők a tökéletes titoktartás

korlátai. Később, a titkos kulcsú titkosítási technika és az üzenethitelesítés átfogó tárgyalás után, bemutatják a blokkrejtjelező titkosító eljárás tervezési alapelveit (például a (szimmetrikus kulcsú) adattitkosítási és a magas szintű adattitkosítási szabványokat), valamint a blokkrejtjelező titkosító eljárás biztonságos kidolgozását is alacsony szintű primitívekből (elemi műveletekből). A kötet második része a nyílt kulcsú kriptográfiára összpontosít. Először bevezet az RSA-³, a Diffie–Hellman- és az El Gamal-féle, valamint az egyéb titkosítási rendszerek megértéséhez szükséges számelméletbe, majd a nyíltkulcsú kriptográfiát és a digitális aláírásokat mutatja be. Végezetül a véletlen orákulum modellt és alkalmazásait tárgyalja.

SUMMER, S. [2015]: *You: For Sale. Protecting Your Personal Data and Privacy Online*. (Eladó vagy. A személyes adatok és a magánszféra online védelme.) Elsevier/Syngress. Waltham.

Bármit, amit online teszünk, figyelemmel kísérik, naplózzák, elemzik, és gyakran „csomagolva” eladják a legmagasabb árat kínálóknak. Valahányszor felmegyünk egy honlapra, hitelkártyát használunk, végighajtunk az autópályán, vagy elmegyünk egy zárt láncú TV-kamera előtt, naplóznak és nyomon követnek minket. Minden nap több milliárd ember dönt úgy, hogy megosztja életének részleteit a közösségi médián, melyeket aztán hirdetőik számára forgalmazznak.

Edward Snowden számítógépes szakember, korábbi CIA-alkalmazott⁴ kijelentése, miszerint a kormányok – köztük az Egyesült Államoké és az Egyesült Királyságé is – saját állampolgáraik után szaglászhatnak, megrengette a világot. De úgy tűnik, senki sem veszi észre,

³ RSA: Rivest–Shamir–Adleman.

⁴ CIA (Central Intelligence Agency): Központi Hírszerző Ügynökség.

hogy mindez már évek óta zajlik, és nem fedjük fel a piszkos valóságot sem: az embereknek fogalmuk sincs arról, hogy a számítástechnikai cégek megszerzik, majd hatalmas nyereség fejében eladják ügyfeleik személyes adatait.

A kötet azoknak szól, akiket a magánszféra vállalati és kormányzati megsértése foglalkoztat. Először felvázolja, hogy pontosan minek teszi ki magát a legtöbb internet- és okostelefon-használó olyan elterjedt honlapok és alkalmazások miatt, mint a Facebook és a Google, majd bemutatja, miként védhetjük ellene magunkat. Ezeken túl jogi és kormányzati kérdéseket, jövőbeli tendenciákat is tárgyal.

A könyvben szereplő, vezető biztonsági szakértőkkel, fekete piaci adatkereskedőkkel, bűnüldöző és a témával foglalkozó egyéb csoportokkal készített interjúk új fényben tüntetik fel személyes adatainkat, illetve segítenek megérteni ezek értékét és a velük kapcsolatos veszélyeket.

KO, R. – CHOO, R. (eds.) [2015]: *The Cloud Security Ecosystem*. (Az adatfelhők biztonsági ökoszisztémája.) Elsevier/Syngress. Waltham.

A világhírű kutatók és szakértők tudására építő kötet átfogó módon, multidiszciplináris és nemzetközi perspektívából tárgyal a számítástechnikai felhők biztonságát érintő számos témát, az üzleti, jogi és nemzetközi környezetben indult legújabb kezdeményezések mellett technikai biztonsági megoldásokat is felsorakoztatva. Holisztikus megközelítésben írja le az adatfelhők biztonságával kapcsolatos kulcsfontosságú kutatási eredményeket és politikai fejleményeket. Olyan vezérgondolkodók és vezető kutatók írásait foglalja egybe, akik a téma technikai, jogi, üzleti és igazgatási szempontjaival foglalkoznak. Ők a különböző tudományágak közötti kapcsolatot és az adatfelhő-biztonsági tervezés, végrehajtás és intézkedéshozatal megvalósítását, jogi kihívásait

tárgyalva mutatják be az élvonalbeli kutatásokat.

TICHER, P. [2015]: *Data Protection and the Cloud: Are the Risks Too Great?* (Adatvédelem és adatfelhő: túl nagyok a kockázatok?) IT Governance Publishing. Ely.

A vállalkozások több mint 85 százaléka használja ki a felhőalapú számítástechnika nyújtotta előnyöket, amely azonban nem áll összhangban az adatvédelmi törvénnyel. A kötet ezzel a kérdéssel foglalkozik, miközben szakértői betekintést nyújt a számítástechnikai felhőalapú szolgáltatások használata során felmerülő jogi és gyakorlati adatvédelmi kockázatokba. Kiemeli, hogy a szervezeteknek milyen veszélyekkel kell szembenéniük, és ezek enyhítésére korrekciós intézkedéseket sorakoztat fel. A következő témákat tárgyalja: a személyes adatok biztonságának, integritásának és elérhetőségének védelme, adatvédelmi kötelezettségek, az adatkezelő és -feldolgozó kapcsolata, választás a felhőalapú szolgáltatást nyújtók közül, a felhők biztonsága (beleértve a kéttényezős hitelesítést, az adatok osztályozását és szegmentációját is), az adatok fokozott sérülékenysége adatátvitel során, saját eszközök használatával kapcsolatos problémák, adatátvitel külföldön, az Egyesült Államok „biztonságos kikötő” alapelvei és az EU jogszabályai, egyéb jogszabályok, jogi keretek és útmutatók (többek között az EU általános adatvédelmi rendelete, a felhőalapú számítástechnikai szabványok, az ISO 27001, a nyílt forráskódú webes alkalmazások biztonságával foglalkozó projekt), valamint a biztonsági kockázatok csökkentése. Az utóbbi számos összetett intézkedést kíván a végponttól végpontig terjedő biztonság megteremtésére. A számítástechnikai felhők nem oldják meg a biztonsági problémákat, hanem egy újabb kezelendő elemmel bővítik azt. A kötet ezért arról nyújt

információt, hogy miképp oldjuk ezt meg az adatvédelmi törvény nyolc alapelvének kielégítése mellett.

OTTENHEIMER, D. [2015]: *The Realities of Securing Big Data, 1st Edition*. (A big data védelmével kapcsolatos tények, 1. kiadás.) Wiley. Hoboken.

Napjainkban több adat gyűlik össze, nagyobb sebességgel és több formátumban, mint valaha. A hagyományos információbiztonság egyszerűen már nem képes lépést tartani ezzel az átalakulással, ami nehéz dilemma elé állítja

az informatikai vezetőket. Egyrészt a legkisebb cégektől kezdve a nagy kormányokig már mindenki kezdi kihasználni az óriási adathalmazokban rejlő lehetőségeket. Másrészt több bajt, mint hasznot fedezhetünk fel, ha a big data biztonságával kapcsolatos tényeket tekintjük. Még konkrét megoldások igénye nélkül is lehet, hogy fontolóra kell vennünk e probléma kezelését és az ellensúlyozást szolgáló ellenőrzést. A kötet annak meghatározásában segít az olvasóknak, hogy miképp és hol nyújtható legjobb védelem big data környezetben felfedéssel, zavarral és veszteséggel szemben.

Társfolyóiratok



A CSEH STATISZTIKAI HIVATAL
FOLYÓIRATA

2015. ÉVI 2. SZÁM

Pavlik, Z.: Az emberiség fejlődése demográfiai szempontból.

Šanda, R.: Szokásos tartózkodási hely és bejegyzett lakóhely a 2011. évi népszámlálásban.

Šprocha, B.: Házasságon kívüli születések és termékenység Szlovákiában.

Koudelka, P.: Termékenység és a férfiak szerepe a háztartásban. Segít-e a nemek közötti egyenlőség?

Fiatal Demográfusok 6. Konferenciája.

A fejlődő országok demográfiájának hat lényeges problémája – előadás.

Alena Šubrtová 80. születésnapja.

A lakosság szociális juttatásokból származó jövedelme 2000 és 2014 között.



A FRANCIA GAZDASÁGI ÉS PÉNZÜGYMINISZTERIUM, VALAMINT A STATISZTIKAI ÉS GAZDASÁGKUTATÓ INTÉZET FOLYÓIRATA

2014. ÉVI 477. SZÁM

Boutchenik, B. et al.: A családi háttér iskolai és szakmai életútra gyakorolt hatásának számszerűsítése.

Benoteau, I. – Rémy, V.: A Munkaügyi Minisztérium Kutatási, Tanulmányi és Statisztikai Igazgatósága által végrehajtott „Panel 2008” felmérés mint a Társadalmi Kohéziós Terv keretében támogatott munkaszerződéseknek az álláskereső szakmai életútjára gyakorolt hatását értékelő eszköz.

Benoteau, I. – Rémy, V.: Milyen kapcsolata van az államilag támogatott foglalkoztatás kedvezményezettjeinek a munkaerőpiaccal? Leírás a Panel 2008 adatai alapján.

Benoteau, I.: Milyen hatása van az államilag támogatott munkaszerződéseken alapuló foglalkoztatásnak a kedvezményezettek szakmai életútjára? Értékelés a Panel 2008 adatai alapján.



A KÖZPONTI STATISZTIKAI HIVATAL FOLYÓIRATA

2015. ÉVI 4. SZÁM

Jakobi Á.: A grid: aggregált és dezaggregált rácsmodellek a területi egyenlőtlenségek vizsgálatában.

Daróczy G. – Bálint L.: Áramlásterkép R-ben.

Bugya T. – Trócsányi A. – Pirisi G. – Fábrián Sz. Á.: A magyarországi mentőellátás térbeli hatékonyságjavításának modellezése – egy lehetséges térinformatikai alkalmazás segítségével.

Kothenz Gy.: Szegedi parkok életminőség-befolyásoló hatásainak elemzése látogatói vélemények alapján.

2015. ÉVI 5. SZÁM

Kincses Á.: A szomszédos országok migrációs forrásterületeinek a vizsgálata a magyar népszámlálási adatok felhasználásával 2001 és 2011 között.

Grünhut Z.: Innovativitás és kockázatvállalás.

Bozóti A.: Nyugat-Dunántúl és Dél-Dunántúl gyógy- és termálfürdővel rendelkező településeinek komplex versenyképességi vizsgálata.

Arató M.: Biztosítások területi kockázatainak meghatározásáról.

Czirfusz M. – Pósfai Zs.: Kritikus ponton? Önkormányzati lakásgazdálkodás a gazdasági válság után.



A SZLOVÁK STATISZTIKAI HIVATAL FOLYÓIRATA

2015. ÉVI 3. SZÁM

Šprocha, B.: A lakosság korösszetétele Szlovákia régióiban és előreszámítása 2030-ig.

Piscová, M. – Klobučký, R. – Bahna, M.: Idősek a társadalomban és a családban. Nézetek az idősekről statisztikák és összehasonlító felvételek alapján.

Káčerová, M. – Ondačková, J.: Szlovákia népességének elöregedése európai kontextusban.

Vaňo, B.: Lehet a népességöregedés folyamatát befolyásolni Szlovákiában?

Páleník, V.: Az idősek gazdasága (ezüstgazdaság) Szlovákiában és az ezzel kapcsolatos lehetőségek.

Dotlačilová, P. – Langhamrová, J.: Gompertz és Makeham továbbélési valószínűségi függvényének paraméterbecslése: a King-Hardy-módszer.

Cár, M.: Európai bizottsági jelentés a népesség elöregedéséről.

Katuša, M.: Ahogy a statisztikusok kezelik a népesség elöregedésének fogalmi kérdéseit.

Sedláková, D. – Štukovská, Z.: Az idősekről alkotott társadalmi képnek meg kell változnia – interjú.

Bušová, B.: Az idősek szociális ellátásának kérdése a szolgáltatást nyújtók szemszögéből.

Hrebeňárová, L.: Időskorúak képzése – a „University of the Third Age” (Időskorúak Egyeteme) elnevezésű nemzetközi kezdeményezés küldetése, kihívásai és lehetőségei.

Langhamrová, J.: Hosszú emberi életek – az emberiség sikere.

statistika

EKONOMICKO - STATISTICKÝ ČASOPIS

A CSEH STATISZTIKAI HIVATAL
FOLYÓIRATA

2015. ÉVI 2. SZÁM

Rolland, A.: Láncok, üzletek és hálózatok. A hivatalos statisztika és a közszolgálati érték létrehozása.

Dabbicco, G.: A közösségi szektor határai a nemzeti számlákban és a nemzetközi költségvetési számviteli standardokban.

Holý, D. – Stražilová, G.: A munkanélküliség veszélye és jövedelmi szintek társadalmi-gazdasági csoportok szerint – az ESeG-osztályozás⁵ bemutatása.

Mazziotta, M. – Pareto, A.: Időbeli változások mérésére szolgáló, nem kiegyenlítő kompozit indikátorok összehasonlítása – eset-tanulmány.

Hindls, R. – Hronová, S.: A Benford-törvény és alkalmazási lehetőségei a kormányzati statisztikában.

Laczka, É.: Az agrárstatisztika fejlődése a tervezéstől az adatközlésig.

WIADOMOŚCI STATYSTYCZNE

A LENGYEL STATISZTIKAI FŐHIVATAL
FOLYÓIRATA

2015. ÉVI 6. SZÁM

Panek, T.: Az életminőség-mérések hierarchikus modellje.

Bielak, R.: Életminőség a fejlesztési stratégiában. Méréssel kapcsolatos dilemmák.

Wilak, K.: Hibabecslések autokorrelációja munkaerő-felmérések esetén.

Wrzaszcz, W. – Zegar, J. S.: A mezőgazdasági üzemek gazdasági fenntarthatósága a 2010. évi mezőgazdasági összeírás adatai alapján.

Paradysz, S.: Iparosítás, iparlépítés és az újraiparosítás kezdete Lengyelországban.

Białynicki-Birula, P.: A globális pénzügyi válság hatása az egészségügyi kiadásokra a fejlődő országokban.

Baruk, J.: Innováció a közszolgáltatásban és hatása az Európai Unió vállalkozásaira.

Lengyelország társadalmi-gazdasági helyzete 2015 áprilisában.

Wirtschaft und Statistik

A NÉMET SZÖVETSÉGI STATISZTIKAI
HIVATAL FOLYÓIRATA

2015. ÉVI 4. SZÁM

Meinke, I.: A Német Statisztikai Hivatal általános felülvizsgálati politikája.

Kaus, W. – Mundil-Schwarz, R.: A népességszám és -szerkezet meghatározása a 2011. évi népszámlálás alapján.

Körner, T. – Marder-Puch, K.: A mikrocenzus és más munkaerő-statisztikák összehasonlítása.

Baumann, T.: A bűnügyi statisztikák nemzetközi összehasonlíthatóságának elvei.

Weber, T. – Beck, M.: Munkabéren felüli juttatások Németországban.

Junglewitz, G.: Az Európai Unión belüli kereskedelmi háromszögek megjelenése a külkereskedelem-statisztikában.

⁵ ESeG (European socio-economic groups): európai társadalmi-gazdasági csoportok.

Statisztikai Szemle

93. évfolyam

2015. év

Tartalom

TANULMÁNYOK

Hol tart a kelet-közép-európai kockázati- és magántőkepiac öt évvel a válság után? – <i>Karsai Judit</i>	1/5
A hazai társadalmi-gazdasági mutatók vizsgálata a káoszelmélet eszközével – <i>Nováky Erzsébet – Orosz Miklós</i>	1/25
A többtényezős termelékenység és a környezeti állapot változása a magyar mezőgazdaságban, az EU-csatlakozást követően – <i>Nagy Zsuzsanna – Baráth Lajos</i>	1/53
Külföldön dolgozó magyarok, Magyarországon dolgozó külföldiek – <i>Lakatos Judit</i> .	2/93
Csehország, Magyarország és Szlovákia termékenységi idősorainak összehasonlítása – <i>Berde Éva – Németh Petra</i>	2/113
Védjegyjoltalmak és a regionális innovációs aktivitás – <i>Szigethy László</i>	2/142
Rövid távú olajár-előrejelzések teljesítményének stabilitása – <i>Uliha Gábor</i>	3/189
Hogyan juthatunk (BPM)5-ről a (BPM)6-ra – módszertani váltás a szolgáltatás- külkereskedelmi statisztikában – <i>Regős Gábor</i>	3/225
A BUBOR-piac kiszáradásának jelei, avagy mi olvasható ki a 2006 és 2012 közötti BUBOR-jegyzésekből – <i>Fliszár Vilmos</i>	3/243
Javuló magyar gazdasági tendenciák bizonytalan világgazdasági környezetben – <i>Nagy Katalin – Palócz Éva – Vakhal Péter</i>	4/297
Behálózva, avagy a webalapú technológiák csoportos együttműködésre gyakorolt hatásának kvantitatív mérése – <i>Heidrich Balázs – Kása Richárd – Sándorné Kriszt Éva</i>	4/319
Hazautalt pénzek – nemzetközi áttekintés és a főbb mérési nehézségek – <i>Kajdi László</i>	4/353
Az időjárás véletlen hatásának szerepe a szezonális kiigazítás során, a hazai földgázfogyasztás példáján – <i>Mák Fruzsina</i>	5/417
Gondolatok az inflációról, több évtized távlatában – <i>Marton Ádám</i>	5/442
A területi fejlettség és egyenlőtlenségek lehetséges értelmezései – kritikai értékelés és kutatási eredmények I. – <i>Harcza István</i>	5/460
A területi fejlettség és egyenlőtlenségek lehetséges értelmezései – kritikai értékelés és kutatási eredmények II. – <i>Harcza István</i>	6/521
A szórakoztató, az irodalmi és a művészeti alkotások elszámolása a nemzeti számlákban – <i>Dienes Ferenc</i>	6/552
Miért választják a termelők a rövid ellátási láncokat? – <i>Benedek Zsófia – Fertő Imre</i>	6/580

Egészségesen várható élettartamok Magyarországon, 2011 – <i>Faragó Miklós</i>	7/633
A gyermeknevelés hatása a háztartások kiadási szerkezetére – <i>Neulinger Ágnes – Radó Márta</i>	7/662
Ikrek a családban – <i>Pári András – Drjenovszky Zsófia – Hegedűs Rita</i>	7/689
Adatforradalom és hivatalos statisztika – <i>Vukovich Gabriella</i>	8–9/745
A mintakoordináció szerepe a statisztikai adatgyűjtésekben – <i>Mihályffy László</i>	8–9/759
Strukturális egyenletek modelljének alkalmazása a Közös Agrárpolitika 2013-as reformjának elemzésére – <i>Kovács Attila</i>	8–9/801
A hazai egészségügyi intézmények kontrolling-rendszere – <i>Krenyácz Éva</i>	8–9/823
A magyarországi létminimum-számítás korszakai nemzetközi összehasonlításban – <i>Havasi Éva</i>	10/885
Felsőoktatásba jelentkezők preferenciáinak térbeli és időbeli szerkezete, teljesítményfüggése – <i>Kosztján Zsolt Tibor – Telcs András – Török Ádám</i>	10/917
Az adatvédelem és a statisztika kapcsolatának jogi szabályozása – <i>Lakatos Miklós</i> ..	11–12/1017
A statisztikai adatvédelem és -hozzáférés szabályai az Európai Unióban – <i>Nagy Eszter</i>	11–12/1051
Az információszabadság szerepe, jelentősége, aktuális kihívásai – <i>Révész Balázs – Buzás Péter</i>	11–12/1070
Az adat-hozzáférési igények és az adatvédelmi kötelezettségek egyensúlyáról – <i>Vereczkei Zoltán</i>	11–12/1089
Nyilvános mikroadatfájlok összeállításának főbb jellemzői, különös tekintettel az adatvédelmi szempontokra – <i>Kristóf Péter</i>	11–12/1112
Az EU-SILC- és az EU-LFS-adatok nyilvános mikro-adatfájljai – <i>de Wolf, Peter-Paul</i>	11–12/1140
A célzott adatsere módszere a térstatisztikában – <i>Nagy Beáta</i>	11–12/1152

MŰHELY

Hablicsekné Richter Mária PhD-értekezésének vitája – <i>Rózsa Gábor</i>	3/260
A bruttó hazai termék becslése a folyó ártól az implicit árindexig – <i>Hüttl Antónia – Pozsonyi Pál – Szőkéné Boros Zsuzsanna</i>	6/598
Többet, mást, hamarabb – amit a válság utáni gazdaságpolitikai döntéshozatal igényel a statisztikától – <i>Bod Péter Ákos</i>	7/713
Statisztika és szovjethatalom: a virágzás kora (1917–1927) I. – <i>Holka László</i>	8–9/844
Adat és információ – A statisztikai rendszerek feladatairól – <i>Vincze János</i>	10/943
Statisztika és valóság: a munkaerő-statisztikai módszertanválasztás hatása a gazdaságpolitikai narratívákra – <i>Hosszú Edmond – Romhányi Balázs</i>	10/959
Statisztika és szovjethatalom: a virágzás kora (1917–1927) II. – <i>Holka László</i>	10/976
A Központi Statisztikai Hivatalban kezelt statisztikai adatok gyakorlati hozzáférése – <i>Mátyás-Bodovics Edina</i>	11–12/1170
A statisztikai táblázatokban található kis előfordulási számokkal kapcsolatos angol bírósági ítélet – <i>Alexin Zoltán</i>	11–12/1077

FÓRUM

Interjú	
Beszélgetés Ormai Lászlóval – <i>Lakatos Miklós</i>	4/376
Beszélgetés Pongrácz Tiborné demográfussal – <i>Lakatos Miklós</i>	8–9/858
Beszélgetés Péterfalvi Attilával a Nemzeti Adatvédelmi és Információszabadság Hatóságának elnökével – <i>Lakatos Miklós</i>	11–12/1183
Beszámoló, értekezéslet, konferencia, nekrológ	
Mórocz Béla (1926–2014) – <i>Giczi Imréné</i>	1/74
Augusztinovics Mária (1930–2014) – <i>Király Júlia</i>	2/163
Papíron – laptopon – tableten – interneten... több csatornán – de egy cél felé – <i>Virágh Eszter</i>	2/166
Beszámoló az MTA Statisztikai és Jövőkutatási Tudományos Bizottsága Statisztikai Tudományos Albizottságának 2015. január 20-án tartott üléséről – <i>Hunyadi László – Kovács Péter</i>	3/265
Klinger András (1930–2015) – <i>Kamarás Ferenc – Rózsa Dávid</i>	5/487
Pódiumvita a települési-térségi társadalmi fejlettség méréséről alternatív megközelítésben – <i>Pári András</i>	6/610
A Nemzeti Adatvédelmi és Információszabadság Hatóság 2014. évi beszámolója – <i>Lakatos Miklós</i>	11–12/1190
Hírek, események	1/75, 2/168, 3/268, 4/385, 5/490, 6/613, 7/725, 8–9/864, 10/993, 11–12/1197

SZAKIRODALOM

Könyvszemle	
Rózsa. D. (szerk.): Portrék a magyar statisztika és népességtudomány történetéből – <i>(Hajnal Béla)</i>	3/271
Csaba L.: Európai közgazdaságtan – <i>(Szabó Katalin)</i>	4/391
Fazekas K. – Neumann L. (szerk.): Munkaerőpiaci tükör 2013 – <i>(Hárs Ágnes)</i>	5/495
Folyóiratszemle	1/78, 2/171, 3/279, 4/397, 5/502, 6/617, 7/729, 8–9/871, 10/998, 11–12/1202
Kiadók ajánlata	1/85, 2/180, 3/284, 4/405, 5/510, 6/622, 7/736, 8–9/877, 10/1005 11– 12/1206
Társfolyóiratok	1/86, 2/182, 3/286, 4/408, 5/512, 6/624, 7/738, 8–9/879, 10/1007, 11–12/1209

Névmutató

(A Statisztikai Szemle 2015. évi számaiban megjelent írások szerzői.)

Alexin Zoltán	11–12/1077	Marton Ádám	5/442
Baráth Lajos	1/53	Mátyás-Bodovics Edina	11–12/1170
Benedek Zsófia	6/580	Mihályffy László	8–9/759
Berde Éva	2/113	Nagy Beáta	11–12/1152
Bod Péter Ákos	7/713	Nagy Eszter	11–12/1051
Buzás Péter	11–12/1070	Nagy Katalin	4/297
Dienes Ferenc	6/552	Nagy Zsuzsanna	1/53
Drjenovszky Zsófia	7/689	Németh Petra	2/113
Faragó Miklós	7/633	Neulinger Ágnes	7/662
Fertő Imre	6/580	Nováky Erzsébet	1/25
Fliszár Vilmos	3/243	Orosz Miklós	1/25
Giczi Imréné	1/74	Palócz Éva	4/297
Harcza István	5/460, 6/521	Pári András	6/610, 7/689
Havasi Éva	10/885	Pozsonyi Pál	6/598
Hegedüs Rita	7/689	Radó Márta	7/662
Heidrich Balázs	4/319	Regős Gábor	3/225
Holka László	8–9/844, 10/976	Révész Balázs	11–12/1070
Hosszú Edmond	10/959	Romhányi Balázs	10/959
Hunyadi László	3/265	Rózsa Dávid	5/487
Hüttl Antónia	6/598	Rózsa Gábor	3/260
Kajdi László	4/353	Sándorné Kriszt Éva	4/319
Kamarás Ferenc	5/487	Szigethy László	2/142
Karsai Judit	1/5	Szökéné Boros Zsuzsanna	6/598
Kása Richárd	4/319	Telcs András	10/917
Király Júlia	2/163	Török Ádám	10/917
Kosztján Zsolt Tibor	10/917	Uliha Gábor	3/189
Kovács Attila	8–9/801	Vakhal Péter	4/297
Kovács Péter	3/265	Vereczkei Zoltán	11–12/1089
Krenyác Éva	8–9/823	Vincze János	10/943
Kristóf Péter	11–12/1112	Virágh Eszter	2/166
Lakatos Judit	2/93	Vukovich Gabriella	8–9/745
Lakatos Miklós	4/376, 8–9/858, 11–12/1017, 11–12/1183, 11–12/1190	de Wolf, Peter-Paul	11–12/1140
Mák Fruzsina	5/417		

2015-ben a következő lektorok segítették bírálatukkal a *Statisztikai Szemle* szakmai színvonalának megőrzését. Munkájukat ezúton is köszönjük!

Balatoni András
Besenyei Lajos
Bruckner Józsefné
Dóry Tibor
Elek Péter
Erhart Szilárd
Farágó Miklós
Füstös László
Hablicsekné Richter Mária
Hunyadi László
Juhász Anikó
Lakatos Judit
Láng Dóra
Lindnerné Eperjesi Erzsébet

Mák Fruzsina
Marton Ádám
Melegh Attila
Molnár György
Nemes Nagy József
Pellényi Gábor
Romhányi Balázs
Rózsa Gábor
Sivák József
Sugár András
Szép Katalin
Szívós Péter
Vita László
Zemplényi Antal