

2. Az államigazgatásban keletkező adatok nyilvánosságáról

Cseres-Gergely Zsombor - Scharle Ágota

Ebben a kötetben a kormányzati működést támogató, közpénzből létrehozott adminisztratív nyilvántartásokra illetve az ezekből származó adatbázisokra épülő kutatásokat mutatunk be. Az adminisztratív adatok ilyen célú felhasználása Magyarországon még szórványosnak mondható, de ez a gyakorlat változóban van. A bevezetőben röviden összefoglaljuk a másodlagos felhasználás melletti és elleni érveket, és áttekintjük az adathozzáférés hatályos szabályozását.

Az államigazgatás adminisztratív nyilvántartásainak hozzáférhetősége: érvek és ellenérvek

Hatósági vagy szolgáltatási feladataik ellátása érdekében az államigazgatás intézményei sokféle adatot gyűjtenek, illetve működésük során sokféle adat keletkezik. Az adatok egy része az intézménnyel kapcsolatba került emberekre vonatkozik: a járulékot vagy adót fizető vállalkozókra, az orvosi ellátásban részesülő betegekre, az útlevelet kiváltó magyar állampolgárokra, stb. Más részük gazdasági szervezetekre vonatkozik: például társasági adót fizető, építési engedélyért folyamodó, vagy munkaerőigényt bejelentő vállalkozásokra. Az egységes ügymenet érdekében és az ellenőrizhetőség követelménye miatt az állami intézmények ezeket az adatokat valamilyen rendszerben - többnyire elektronikus formában – hosszabb-rövidebb ideig megőrzik.

Ezek a nyilvántartások elsődleges céljukat tekintve az állami intézmények feladatainak ellátását szolgálják, de rengeteg olyan értékes információt is tartalmaznak, amit más célokra is fel lehet használni. Ilyen másodlagos cél lehet például, hogy a nyilvántartásokból kormányzati döntéseket alátámasztó hatásvizsgálatok készüljenek, hogy civil szervezetek is ellenőrizhessék egyes kormányzati szervek működését, vagy, hogy tudományos kutatásokra vagy akár piackutatásra is lehessen használni az adatokat. Az alapvető összesített, átlagolt adatok természetesen ma is hozzáférhetők: itt nem ezekről, hanem a részletesebb vizsgálatokhoz szükséges egyedi adatokról lesz szó. A következőkben azt tekintjük át röviden, hogy milyen érvek szólnak az ilyen másodlagos felhasználások mellett és ellen.

Érvek az adminisztratív nyilvántartások másodlagos felhasználása mellett

Az államigazgatási adatok elérhetősége mellett közvetlen és közvetett hasznok, a közjósághoz közelítő jelleg, illetve a hozzáféréssel járó költségek sajátos szerkezete szól. Vegyük most sorra ezeket röviden!

Egy termék vagy szolgáltatás akkor közjóság, ha egy gazdasági szereplő ellátása után minden további gazdasági szereplő is költségmentesen hozzájuthat¹ és fogyasztásukból nem zárhatók ki egyes szereplők. Az első tulajdonság adatok esetében általában teljesül, így az adat potenciális közjóság. A második tulajdonság már csak bizonyos adatok esetében teljesül. A nyilvánosan közzétett - például a szabadon hozzáférhető internetes honlapokról letölthető – adatok közjavak, míg a keletkezésük helye és/vagy eljárási szabályok alapján elzártan kezelt adatok nyilvánvalóan nem azok. A nyilvános adatok esetében gyakorlatilag költségmentesen, míg a zártan kezelt adatok esetében kisebb-nagyobb költségek árán lehet biztosítani a hozzáférést.

Az adatkezelő számára a legnagyobb költséget az adatbázis létrehozása, az adatok összegyűjtése jelenti, amit közpénzből finanszíroz. Az adatbázis fenntartásának egy adategységre jutó költsége (határkölség) minimális, az adatok másolásának, módosításának, és átadásának költsége szintén relatíve csekély. A nagy adatbázisok ugyanis jellemzően elektronikus formában léteznek, és a módosítások könnyen előállítható, és gyorsan végrehajtható számítógépes programok segítségével elvégezhetők. A költségvetési szervek adminisztratív célú adatbázisainak többsége azzal az előnnyel is rendelkezik, hogy relatíve pontos információkat tartalmaz, mivel az adatkezelő alapfeladata az adatok ellenőrzése is, és munkájának természetes része az adatok folyamatos tisztítása, az esetleges adatbeviteli hibák javítása. Cseres-Gergely és Csorba (2006) egy egyszerű közgazdasági modellben összeveti az adatszolgáltatásból származó társadalmi hasznosságot, illetve az adatbázis fenntartásának és az adatok átadásának költségét. Az ebből levonható következtetés szerint az adatátadás társadalmilag optimális ára az adatbázis fenntartásához szükséges határkölséggel egyenlő, ami gyakorlatilag nulla.

Az adatok másodlagos felhasználása tudományos eredmények és jobb kormányzati döntések

¹ Azaz a használók között nincs rivalizálás: egy szereplő fogyasztása következtében a többi szereplő által igénybe vehető mennyiség nem csökken.

formájában is közvetlen hasznokkal járhat. A tudományos kutatásban nyilvánvaló, hogy a nagyszámú és pontos adatokon végzett empirikus vizsgálat alapján megbízhatóbb következtetésekre lehet jutni, mint kevés és bizonytalan adatból.² A kormányzati döntéselőkészítésben a megalapozottabb döntések jelentik a közvetlen hasznot. Hatásvizsgálatra a kormányzati működés két pontján is szükség van. Egyrészt egy szándékolt intézkedés hatását minél pontosabban kell tudni előre jelezni, hogy látható legyen melyik, és milyen erősségű intézkedés éri el legkisebb ráfordítással a kívánt hatást. Az intézkedés után pedig az új helyzet és a beavatkozás előtti állapot összehasonlításával kell az intézkedés tényleges hatását mérni. Mindkét esetben adatokkal kell rendelkezünk azokról az egységekről, akiket az intézkedés céloz. Az államháztartás korlátos erőforrásait tekintve az előzetes és utólagos mérés is kulcsfontosságú.

Végül, mindkét adathozzáférés esetén jelentkezik a használat által megvalósuló minőségellenőrzés, mint közvetett haszon. A zárt adatbázisok, függetlenül attól, hogy azok eredendően statisztikai, vagy államigazgatási céllal jöttek létre, csak a vonatkozó feladatok elvégzéséhez szükséges ellenőrzéseken esnek át, ami szükségszerűen nem lehet teljeskörű. Az adatbázisok kutatási célú felhasználása, éppen az államigazgatási munkától eltérő, az egyedi helyett az általánosra irányuló figyelem miatt, rejtett hibákat tárhat fel, amelyek kijavítása mindenkinek alapvető érdeke.

Magyarázatra szorulhat még az a kérdés, hogy a kutatók miért nem elégednek meg a nyilvántartásokból származó összesített adatokkal. Az egyedi (egyénre vagy vállalatra vonatkozó) adatok felhasználását a elsősorban a kutatási és hatásvizsgálati módszerek fejlődése, a többváltozós statisztikai (ökonometriai) modellek alkalmazása teszi szükségessé. Mikroszintű adatok kellhetnek a szabályozási javaslatok előzetes hatásvizsgálatához és a közkiadások, vagy egy adott szabályozás eredményességének utólagos értékeléshez is. Például annak vizsgálatához, hogy a különféle jóléti juttatások együttesen mennyire csökkentik a gyermekszegénységet, olyan adat kell, amiből kiderül, hogy ugyanaz a szülő hányféle és mekkora támogatást kap, milyen más jövedelme van, és hány gyermeket nevel. Ezt csak egy olyan adatállomány biztosíthatja, amelyben egyéni szintű adatok vannak, és azt is tudjuk, hogy kik élnek egy háztartásban. Vagy, ha az egyszülős családok szegénységi

² További szempont, hogy könnyen elérhető adatok hiányában a kutatók más országok adatait használják, és így saját országuk gazdaságának vagy társadalmának működéséről a szükségesnél kevesebb empirikus elemzés születik. Megfordítva, ha könnyen elérhetőek és jó minőségűek az adatok, a hazai kutatók könnyebben állják a versenyt a nemzetközi mezőnyben is.

rátáját akarnánk csökkenteni, mikroszintű adatok segítségével tudjuk megállapítani, hogy ezt egy meglévő támogatás növelésével, vagy új támogatással érdemes-e megoldani. Összesített adatokból nem lehet ugyanis megállapítani, hogy egy adott háztartás hányféle és mekkora támogatást kap, és milyen más jövedelme van.

Nemzetközi ajánlások és jó gyakorlatok az adminisztratív adatok másodlagos felhasználásában

Az állami adminisztratív nyilvántartások másodlagos felhasználását több nemzetközi szervezet is szorgalmazza, mint a kormányzati hatékonyság, illetve a kutatások ösztönzésének egy lehetséges eszközét. Egy 2007 júniusában szervezett nemzetközi konferencián az Európai Bizottság, az OECD, a Világbank és az ENSZ képviselői közös nyilatkozatot adtak ki, melyben többek között arra hívják fel a kormányokat, hogy minden területen segítsék a tényeken alapuló döntéshozatali kultúra elterjedését, és leszögezik, hogy a transzparens és felelős kormányzati döntéshozatal jelentősen hozzájárul a jólét növekedéséhez (Istanbul... , 2007). Az OECD 14 December 2006. december 14.-i C(2006)184 tanácsi javaslata a közfinanszírozású kutatási adatok hozzáférhetőségéről az adathozzáférés és adatmegosztás előmozdítását szorgalmazza a kutatók, kutatási és egyéb intézmények között a tagországok nemzeti jogrendjének figyelembevételével.

A kormányzati intézkedések és szabályozás rendszeres hatásvizsgálatát az Egyesült Államokban vezették be először. Az OECD 1995-ben útmutatót adott ki, amelyben úgy határozza meg a hatáselemzést, mint “információ alapú elemzési módszer, mely segít felbecsülni a tervezett szakpolitikai eszközök (törvények, rendeletek, stb) várható költségeit, következményeit és mellékhatásait.” Az Európai Bizottságban 2003 óta, az Egyesült Királyságban 2000 óta végeznek egységes rendszerben szabályozási hatásvizsgálatokat. Az 1999-ben elindított brit kormányzati modernizációs programban a tényeken alapuló döntéshozatal különösen nagy hangsúlyt kapott: egy önálló módszertani központot is létrehoztak a módszertani fejlesztéssel foglalkozó tudományos kutatóhelyek együttműködésének koordinálására (Bullock–Mountford–Stanley, 2001).

A nemzetközi gyakorlatban az adathozzáférést érintő szabályozás nyíltsága és jellege sokféle, azonban a legtöbb nyugat-európai és angolszász országban a hazainál nyíltabb intézményrendszer működik (Cseres-Gegely – Csorba, 2006). A hatásvizsgálatok beépülése a

döntéshozatali folyamatba együtt járt az adathozzáférésre vonatkozó szabályozás megújításával. Az Európai Parlament és a Tanács 2003/98/EK. számú irányelve (2003.11.17.) rendelkezik a közsféra információinak további felhasználásáról, és előírja, hogy a közigazgatási szerveket ösztönözni kell a birtokukban lévő adatok és más dokumentumok elsődleges céljukon túli felhasználásának biztosítására.³ Az irányelv minimum szabályokat állapít meg a tagállamok közigazgatási szervei birtokában lévő dokumentumok (ideértve az adatokat is) további felhasználására, illetve a felhasználás megkönnyítésének gyakorlati eszközeire vonatkozóan. Az irányelv szerint, amennyiben lehetséges, a dokumentumokat elektronikus úton kell elérhetővé tenni. Díj szedése esetén a dokumentumok rendelkezésre bocsátásából és további felhasználásának engedélyezéséből eredő teljes bevétel egy ésszerű nyereséghányaddal együtt sem haladhatja meg a dokumentumok gyűjtésének, előállításának, feldolgozásának és terjesztésének költségét. A díjaknak a vonatkozó elszámolási időszakban költségorientáltak kell lenniük, és azokat az érintett közigazgatási szervre vonatkozó számviteli elvekkel összhangban kell kiszámítani.

Érvek a hozzáférés kiterjesztése ellen

Személyes adatok

Az adminisztratív nyilvántartások másodlagos felhasználása, és különösen az egyedi adatok hozzáférhetőségének kiterjesztése elleni legsúlyosabb érv a személyes adatok, illetve az üzleti vagy adótitok védelme. A másodlagos felhasználáshoz ugyan nincs szükség a személyes azonosításra, de a gyakorlatban mégsem mindig egyszerű kizárni ennek lehetőségét.

Vegyünk egy példát: természetes személyre vonatkozó mikroszintű adat a következő „A” adatsor: 51 éves, egyetemi végzettségű, budapesti lakos. Ez az adatsor egy bizonyos személyre - az adatszolgáltatóra - vonatkozik, de alanyával nem hozható kapcsolatba: a három adatból nem lehet azonosítani, hogy ki volt az adatszolgáltató. Mikroszintű, és ezen belül személyes adat például a „B” adatsor: ugyanennek a személynek a neve és TAJ száma. Személyessé válna az „A” adatsor is, ha azt a TAJ számmal együtt közli valaki. Végül, személyes lehet a „C” adatsor is: szül. 1932, csillagász, II. kerületi lakos, ha csak egyetlen személyre igaz és a három adat más, nyilvános adatokkal együtt már lehetővé teszi az illető

³A magyar elektronikus információszabadságról szóló, 2005. évi XC törvény (továbbiakban Einfotv) indoklása is erre az irányelvre hivatkozik

(közvetett) azonosítását.

Az EU No.831/2002 számú tanácsi határozatában mikroszintű adatnak nevezi azt az adatot, ami egy természetes személyre, vagy egy jogi személyre vonatkozik (ld. 1. táblázat). Mindkét adatkört két további halmazra bonthatjuk. Az egyikben az alanyukkal kapcsolatba hozható mikroszintű adatok vannak: ez a statisztikáról szóló 1993. évi XLVI. törvény (továbbiakban Stt) fogalomhasználata szerint az „egyedi adat”, az egyénnel vagy szervezettel kapcsolatba hozható adat. A másik halmazban vannak az egy-egy alanyra vonatkozó, de személyes jellegétől megfosztott, pontosabban az alanyával nem kapcsolatba hozható mikroadatok. Ez utóbbi adatkör megegyezik a KSH Adatvédelmi szabályzata (KSH VII/2005.(SK)2.) szerinti elemi adattal, amely a közvetlen azonosítóktól meg van fosztva, és közvetett azonosításra sem alkalmas. Jellegében megfelel az EU No.831/2002 határozatában definiált „névtelen mikroadat”-nak (anonymised microdata) is, amely az egyén azonosításának lehetőségét minimalizáló, a lehető legkorszerűbb eljárással módosított egyéni adatokat takarja.

1. táblázat A statisztikai vagy adminisztratív adatgyűjtésből származó, egyénekre vagy vállalatokra vonatkozó adatokon belül megkülönböztethető adatfajták

Mikroszintű adat						
Természetes személyre vonatkozó adat			Szervezetre vonatkozó adat			
Személyes adat = azonosítható a személy, akire az adat vonatkozik [Avtv] = egyedi adat [Stt]		Nem személyes adat = nem azonosítható a személy	Anonim adat = nem azonosítható a szervezet, amelyre az adat vonatkozik	Nem anonim adat = azonosítható a szervezet, amelyre az adat vonatkozik = egyedi adat [Stt]		
Közérdekű személyes adat [Avtv]	Nem közérdekű személyes adat	Névtelen mikroadat [831/2002/EK] Anonimizált mikroadat =Elemi adat [KSH adatvédelmi szabályzata]	Adótitok által védett cégadat [Art]	Nyilvános (pl. cégbíróságnak bejelentett) adat	Közérdekű adat (ha költségvetési szerv kezeli)	Egyéb nem anonim adat

Megjegyzés: *hivatalos fordítás. Az eredeti angol: anonymised microdata. Az EU tanácsi rendelet a magyar gyakorlattól eltérően használja a „bizalmas adat” kifejezést: olyan adatot ért ezen, amelyből közvetlenül nem lehet azonosítani a személyt vagy szervezetet, és bizonyos feltételekkel megengedi az ilyen adatok kiadását. Stt: Statisztikai törvény, Avtv: Adatvédelmi törvény

Az EU határozat a magyarnál valamivel kevésbé szigorú, mivel csak a közvetlen azonosításra alkalmas adatok kiadását tiltja, és bizonyos feltételekkel megengedi olyan mikroadatok kiadását, amelyek a közvetett azonosítást még megengedik. Az Avtv ennél szigorúbb: mindaddig személyesnek tekinti az adatot, amíg a kapcsolat a személlyel helyreállítható (akár közvetlen,

akár közvetett módon).

Más, bizalmasnak vagy érzékenynek tekintett adatok (pl. orvosi titok, vallásra, etnikai hovatartozásra vonatkozó adatok, pénztártitok) kezelése nem okoz külön problémát, mivel az ezekre vonatkozó tilalom eleve csak olyan esetben értelmezhető, ahol az átadott adat természetes személyhez vagy gazdasági szervezethez kapcsolható (azaz nem anonim).

Költségek

Ha az adminisztratív adatbázisok létrehozásában és működésében nem érvényesülnek az elsődleges funkciójukon túli felhasználás szempontjai, akkor a nyilvántartások nem feltétlenül lesznek közvetlenül felhasználhatók hatásvizsgálatok vagy kutatások céljára. Gyakori eset, hogy ha az adatok tárolási rendje csak az alapfunkciókhoz igazodik, akkor más célú felhasználásra bonyolult és időigényes kigyűjteni az adatokat. Ezen alapul a másik gyakori ellenérv, miszerint a másodlagos felhasználás magas költségeket terhel az államigazgatási intézményekre, illetve túl sok emberi erőforrást köt le, ami már az elsődleges feladatok elvégzését is veszélyeztetheti.

Ez az érv kétségtelenül mérlegelendő, de a hozzáférés teljes megtagadását nem indokolja: legfeljebb azt, hogy a hozzáférés kiterjesztése fokozatosan történjen. Amint lehetőség nyílik a másodlagos felhasználásra, világossá válhatnak az ebből származó előnyök - nem csak az adatkérők, hanem az adatgazdák számára is -, és ez motivációt teremt azokhoz a pótlólagos beruházásokhoz, átalakításokhoz, amelyek megkönnyítik a nyilvántartások többcélú felhasználását. Ez pedig idővel a költségek csökkenéséhez vezet.

Indokolatlan üzleti előny

A közadatok széles nyilvánossága ellen szólhat az az érvelés, mely szerint az adatok közpénzen jöttek létre, így profítcélú felhasználás esetén az érintett vállalkozást közvetlenül finanszírozzák. Mivel a jelenlegi, törvényi szabályozás ilyen felhasználás elé vagy nem gördít semmilyen akadályt (elérhető anonim adatok), vagy egyáltalán nem tesz lehetővé (anonimizálást igénylő adatok), a kérdést nem tárgyaljuk részletesen. Annyit azonban megjegyzünk, hogy az adatok árazása közgazdaságilag egyáltalán nem nyilvánvaló. Erős érvek sorakoztathatók fel mellett, hogy (ha megkülönböztethetők) a profítcélú felhasználás

esetén az adatokhoz ne térítésmentesen, hanem bizonyos, a hozzáférés tényleges költségén túlmutató díj ellenében legyenek elérhetőek. Az egyik ilyen érv az, hogy egy bizonyos mértékig a profitcélú felhasználók képesek fizetni, a pótlólagos források bevonása és visszaforgatás pedig az adatok minőségének javulásához vezet, ami a közérdekű és profitcélú felhasználás hasznára válik.

A hatályos magyar szabályozás

A magyar szabályozásban alapvetően három törvény szabályozza az egyedi adatok hozzáférhetőségét, illetve néhány ágazati törvény is tartalmaz kiegészítő rendelkezéseket. A Statisztikai törvény határozza meg, hogy mely intézmények jogosultak statisztikai célra adatot gyűjteni vagy továbbítani, az Adatvédelmi törvény (Avtv) biztosítja a személyes adatok védelmét és a közérdekű adathoz fűződő állampolgári jogokat, és végül az Adathozzáférés törvény (Ahf) szabályozza az állami szervek adatkérési jogait és adatátadási kötelezettségeit. A munkaerőpiaci kutatások szempontjából néhány nagy állami szerv, mint az Egészségbiztosítási Pénztár (OEP), a Nyugdíjfolyósító (ONYF), az Államkincstár (MÁK) és az APEH nyilvántartásai a legfontosabbak. Az ezek működését meghatározó ágazati törvényekben is vannak adatkezelést érintő rendelkezések, így például a 2003. évi XCII. törvény az adózás rendjéről (Art) tartalmaz az adótitok fogalmát és kezelését érintő szabályokat. A továbbiakban csak az adminisztratív adatokra általánosan alkalmazható Adatvédelmi és Adathozzáférési törvény lényegét vázoljuk.

A közérdek és a magánszféra védelme

A hatályos szabályozás nem tiltja általában a mikroszintű adatok felhasználását, de az alanyával kapcsolatba hozható adatok kezelésére már többféle korlátozás van érvényben: a személyes adatok kezelését az Avtv, az egyedi adatokét az Stt., az adótitokét az adózás rendjéről szóló 2003. évi XCII. törvény (Art.) korlátozza (ld. 1 táblázat). Az Stt., illetve az államháztartásról szóló 1992. évi XXXVIII. törvény (Áht.) biztosít bizonyos adathozzáférési jogokat államigazgatási szakértők, illetve kutatók számára, de ezt korlátozza az Adatvédelmi törvény, amely alapesetben tiltja a személyes adatok kezelését. Az Avtv a természetes személlyel kapcsolatba hozható adatot személyes adatnak hívja, és főszabály szerint kizárja a személyes adatok kiadását vagy összekapcsolását. Az Avtv a személyes adatok, illetve a közérdek védelmének konfliktusában jellemzően a személyes adatok védelmének ad

prioritást, a közérdekből fakadó lehetséges korlátozások közül pedig egyedül a közszereplőket és közhivatalnokokat érintő korlátozásra utal explicit módon. A konkrét személlyel kapcsolatba nem hozható mikroszintű adatok kezelését azonban az Avtv nem korlátozza.

Az Avtv szerint személyes adatokat akkor lehet kezelni (továbbadni, összekapcsolni), ha ehhez az érintett hozzájárul, vagy törvény megengedi. A személyes adat az adatkezelés során mindaddig megőrzi e minőségét, amíg kapcsolata az érintettel helyreállítható. Személyes adatnak minősülnek a természetes azonosítók, mint például a név, születési év, lakcím, életkor, valamint minden olyan, az adott egyénre vonatkozó adat, amit ilyen azonosítókkal együtt adnak át. Külön törvény (1996. évi XX. tv.) véd egyes mesterséges személyes azonosítókat, mint adószám, TAJ szám, személyi szám: ezeket még korlátozottabban lehet kiadni, összekapcsolásra használni.

A hatásvizsgálatokhoz szükséges adatok biztosítása tehát azt a problémát veti fel, hogy a közérdek és az egyéni érdek (a személyes vagy bizalmas adatok védelme) hogyan egyeztethető össze. Közérdek, hogy a közpénzeket a lehető legnagyobb hatékonysággal költjük el, és lehetővé tegyük az ezt segítő hatásvizsgálatokat. Közérdek az is, hogy a közpénzen létrehozott adatvagyon adminisztratív célok mellett más, a társadalmi jólétet növelő célokra (hatásvizsgálatra és kutatásra) is hasznosítható legyen. Ezzel egyenragú szempont a bizalmas, és különösen a személyes adatok védelme, vagyis az, hogy minden állampolgárnak joga van a személyéről szóló információkról rendelkezni.

Ezt a problémát oldja fel az Adathozzáférés törvény, ami az információs önrendelkezési jog, illetve a bizalmas adatok védelmének sérelme nélkül biztosítja az adatvagyon hasznosulását és a mikroszintű adatok hozzáférhetőségét. Ennek lényege, hogy hatásvizsgálatok céljára csak az anonimizált mikroszintű adatokat (az 1. táblázat szürkével jelölve) lehet átadni, de az adatkezelő kérésre köteles az anonimizálást elvégezni. A törvénynek négy, egymással szorosan összefüggő eleme van: (1) előírja azt, hogy a közérdekű adatot az adatkezelők kötelesek kiadni, akkor is, ha azok mikroszintű adatok,⁴ míg azonosításra alkalmas (és nem közérdekű) adataikat kérésre anonimizálniuk kell, (2) a potenciális felhasználók (államigazgatási szakértők) számára pedig törvényi felhatalmazást ad az adatok igénylésére. (3) Az adatok összekapcsolhatóságához fentiekén túl egy technikai megoldás, a hash

⁴ A közérdekű adatok az Avtv alapján nyilvánosak, kiadásukat bárki igényelheti.

függvény segítségével történő titkosítás is szükséges (ez megszünteti a személyességet), és egy előírás, ami kötelezi az adatgazdát arra, hogy ezt a megoldást alkalmazza, és együttműködjön az adat kiadásában. (4) Az információs önrendelkezési jog védelme érdekében előírja, hogy adatösszekapcsolásra irányuló kérelmeket be kell jelenteni az erről nyilvántartást vezető szervnek, illetve, hogy mikroszintű adat nem adható át abban az esetben, ha a személyesség nem szüntethető meg.

Az adatok anonimizálását többféle módon is el lehet végezni. Az első lépés minden esetben a mesterséges és természetes azonosítók törlése. Bizonyos adatbázisok esetében azonban az adatokból is vissza lehet következtetni a személyre vagy vállalkozásra: például kisebb települések esetén a nem, az életkor, a település neve, és a személy foglalkozása is elegendő lehet az azonosításhoz. Ilyen esetben az adatok pontosságának csökkentésével (például település helyett kistérség, életkor helyett korcsoport) lehet megszüntetni a közvetett azonosítás lehetőségét. Ezt többféleképpen is el lehet végezni, ezért az Ahf azt is előírja, hogy a módszer kiválasztásánál a kutatási célt is figyelembe kell venni (2.§(2)). Például az átképzési programok eredményességének vizsgálatához pontos adat kell a foglalkozásról: ekkor a lakóhelyre vonatkozó adatot lehet nagyobb bontásban megadni. A távolsági közlekedés kihasználtságának vizsgálatához viszont legalább kistérségi adatok szükségesek: ekkor például a foglalkozásra vonatkozó adatot kell nagyobb bontásban megadni.

Főszabályként mindkét törvény úgy rendelkezik, hogy maga az adat ingyenes, felszámítani csak az adatátadás költségét lehet (Avtv 20.§, Ahf 3.§). Az átadás határidejét illetően az Avtv szerint a közérdekű adatokat kezelő szerv az igény tudomására jutását követő legrövidebb idő alatt, legfeljebb azonban 15 napon belül tesz annak eleget (20. § (2)). Az elektronikus információszabadságról szóló, 2005. évi XC törvény 3.-6. § pedig a közérdekű adatokra elektronikus közzétételi kötelezettséget ír elő, egységes közadatkereső rendszerben (2005. XC. 7.-8. §). Az Ahf az 1.§ az Avtv-vel összhangban állapítja meg a közérdekű adat kiadásának 15 napos határidejét, és ennek térítésmentességét. Az adatok átadásának megtagadására értelemszerűen ugyanez a határidő vonatkozik.

Nyilvántartások összekapcsolása

Felmerülhetnek olyan kutatási vagy hatásvizsgálati feladatok, amelyeket csak több adatbázis összekapcsolásával lehet megválaszolni. Például a háztartásoknak juttatott állami támogatások

nettó értékét, vagy annak változását csak úgy számolhatjuk ki, ha tudjuk, hogy milyen támogatásokban részesül az adott háztartás és milyen adókat és járulékokat fizet. Ennek kiszámolásához APEH, OEP, ONYF, MÁK adatokat kell egyszerre vizsgálni, és ehhez a négy intézmény nyilvántartásainak összekapcsolására van szükség.

Személyes adatok esetében megnehezíti az összekapcsolást, hogy Magyarországon három nagy adminisztratív nyilvántartási rendszer létezik. A három rendszer elkülönítése azt célozta, hogy az eltérő állami feladatokkal kapcsolatos személyes adatokat csak indokolt esetben, törvényi felhatalmazással, és ellenőrizhető módon lehessen összekapcsolni. A korábbi, egységesen használt személyazonosító jel helyett ma (1) a személyiadat- és lakcímnnyilvántartásban a személyazonosító jelet, (2) az adózók azonosításában az adóazonosítót, (3) a társadalombiztosítási rendszerben pedig a társadalombiztosítási azonosító jelet (TAJ) használjuk.

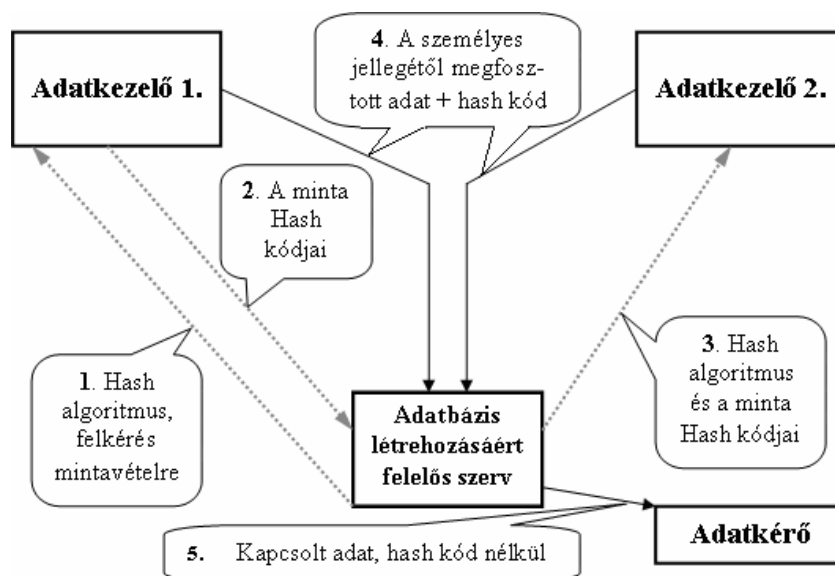
Ezt a három azonosítót a személyazonosító jel helyébe lépő azonosítási módokról és az azonosító kódok használatáról szóló 1996. évi XX. törvény hozta létre, és egyszersmind rendelkezett arról, hogy ezek az azonosítók személyes adatok [5. § (2) bekezdés], valamint arról, hogy mely szervek és hogyan használhatják ezeket az azonosítókat. A különböző azonosítót használó szervek egymással meghatározott célból, úgynevezett kapcsolati kód segítségével cserélhetnek adatokat [3. § d) pont és 26. §]. Ez a kapcsolati kód nem alkalmas két adatkezelőnél gyűjtött adat kutatási és hatásvizsgálati célú összekapcsolására. E kapcsolati kód csupán idősoros vizsgálatokat tesz lehetővé, vagyis azt, hogy ugyanannak a személynek ugyanattól az adatkezelőtől kapott adatait az adatkérő hozzá tudja rendelni a korábbi adatokhoz. Ez természetesen nem véletlen: a jogalkotó épp ezzel akarta elkerülni a különböző forrásokból származó személyes adatok összekapcsolását.

Nem személyes adatok összekapcsolása viszont – ha azok együttesen sem válnak személyessé, tehát együtt sem köthetők egy konkrét természetes személyhez – nem problematikus. Ennek módszere szintén valamiféle kapcsolati kód használatát igényli. A megoldást itt egy olyan kapcsolati kód biztosítja, amely az adatkérő számára nem összekapcsolható az adatok alanyával, tehát nem személyes adat, de amely segítségével az egyes adatok összekapcsolhatók.

Az Avtv szerint személyes adatok csak akkor kapcsolhatók össze, ha a) annak megfelelő

jogalapja van, vagyis az érintett ahhoz hozzájárult, vagy törvény azt megengedi, és b) az adatkezelés feltételei minden egyes személyes adatra nézve teljesülnek [Avtv. 8. § (1) bekezdés]. Az Ahf 4. és 5. §-a ezért olyan kapcsolási módot ír elő, amely biztosítja, hogy az adatkérő számára az adat nem lesz összekapcsolható az adatok alanyával, tehát nem személyes adat. A kapcsolást végző intézmény (az adatbázis létrehozásáért felelős szerv) is személyességüktől megfosztott formában kapja meg az adatokat, de egy olyan kóddal együtt, amelynek segítségével az egyes adatkezelők adatait össze tudja kapcsolni (ld. 1. ábra). Az anonim kapcsolást egy ún. egyirányú algoritmus teszi lehetővé, amely kapcsolati kódot képez, ugyanabból az adataból mindig ugyanazt, amelyből viszont nem lehet visszanyerni az eredeti adatot. Az ilyen tulajdonságú kódok egyik leggyakrabban alkalmazott formája a *hash függvény*.

1. ábra Két adatbázisból vett minta összekapcsolása anonim kapcsolati kóddal



Forrás: 2007/CI törvény indoklása

Például annak vizsgálatához, hogy a táppénz szabályozása hogyan hat a munkakínálatra, az Országos Egészségbiztosítási Pénztár táppénz adatait és az Országos Nyugdíjbiztosítási Főigazgatóság járulékbevallásra (munkavállalásra) vonatkozó adatait kell összekapcsolni. Mindkét szervezet a TAJ alapján azonosítja az egyéneket, így ebből képeznék a kapcsolati kódot.

Konkrét példán szemléltetve a folyamat a következő (lásd 2.1. ábra). Az adatbázis létrehozásáért felelős szerv (1) meghatároz egy hash függvényt és ezt megküldi az egyik

adatkezelőnek. Ha az 1. adatkezelő az OEP, akkor az OEP mintát vesz a nála tárolt adatokból, és minden, a mintában szereplő TAJ számra lefuttatja a hash algoritmust, ami a mintaelem-számmal megegyező számú hash kódot eredményez. A (2) második lépésben kizárólag ezeket a hash kódokat küldi meg az adatbázis létrehozásáért felelős szervnek, aki azokat (3) továbbítja a 2. adatkezelőnek a hash függvénnyel együtt. Ha a 2. adatkezelő a példában az ONYF, ő a saját adatbázisában szereplő TAJ számokon lefuttatja a hash algoritmust, majd az adatbázis létrehozásáért felelős szervtől megkapott hash kódok alapján kiválogatja a mintát. Ezután, (4) mindkét adatkezelő elküldi a mintát és a hozzájuk tartozó hash kódokat az adatbázis létrehozásáért felelős szervnek, aki a hash kódok alapján a két különböző adatbázisból származó adatokat összekapcsolja, majd a hash kódokat törli. Utolsó lépésként (5) az adatkérő megkapja a személyes jellegétől megfosztott, kapcsolt mintát.

Az adatbázis létrehozásáért felelős szerv szerepe kettős: egyfelől biztosítja, hogy két külön adatbázist személyes adatoktól mentesen lehessen összekapcsolni. Másfelől az adatkérések és adatszolgáltatások nyilvántartásával elősegíti az Adatvédelmi Biztos ellenőrző tevékenységét.

Adminisztratív adatok másodlagos felhasználása a gyakorlatban

Bár a jogalkotásról szóló 1987. évi XI. tv rendelkezik a jogszabályok hatályosulásának vizsgálatáról,⁵ részletes hatásvizsgálatok viszonylag ritkán készülnek. A mikroszintű adatok elemzése szakértelmet, számítógépes kapacitást és speciális számítógépes programokat igényel; ma a költségvetési szervek többségénél ezek nem állnak rendelkezésre, ezért kevesen élnek az Ahf által biztosított lehetőségekkel. Az Európai uniós forrásokhoz kapcsolódó követelmények miatt, illetve az elemzési kapacitások javulásával várható, hogy a másodlagos felhasználás iránti igény jelentősen bővülni fog.

A központi államigazgatáson belül adatkérések és hatásvizsgálatok terén legnagyobb tapasztalattal a Pénzügyminisztérium Gazdaságpolitikai főosztályának Közgazdasági kutató osztálya rendelkezik. Ez az egység 2003-mas megalakulása óta készít hatásvizsgálatokat, a költségvetési kiadások eredményességét vagy új szabályozási javaslatok hatásait illetően. Ezekhez a vizsgálatokhoz már az Ahf 2007. évi hatályba lépése előtt is kért adminisztratív

⁵ Ennek további erősítését célozza a jogalkotásról szóló törvény készülő tervezete, mely szerint (34. §) a szakminiszter minden jogszabály előkészítése során megfelelő részletességű előzetes hatásvizsgálat elvégzésével felméri a szabályozás várható következményeit, és legalább ötvenként lefolytatja a feladatkörébe tartozó jogszabályok utólagos hatásvizsgálatát.

nyilvántartásokból származó (Apeh, MÁK, OEP, ONYF) adatokat, és elsőként kezdeményezte két nyilvántartás (Apeh és ÁFSZ Bértarifa) összekapcsolását az Ahf alapján. Ezekről az adatkérésekről a további fejezetekben lesz bővebben szó. Az összekapcsolást az Ahf végrehajtási rendeletében megjelölt felelős szerv, a Neumann Kht végzi.⁶

Hivatkozások

A döntéselőkészítéshez szükséges adatok hozzáférhetőségének biztosításáról szóló 2007. évi CI törvény indoklása, <http://www.parlament.hu/irom38/03029/03029.pdf>

Istanbul... (2007): Istanbul Declaration. A Measuring the Progress of Societies: World Forum on Statistics, Knowledge and Policy konferencián részt vevő nemzetközi szervezetek (European Commission, OECD, Organisation of the Islamic Conference, United Nations, United Nations Development Programme, World Bank) által kiadott nyilatkozat. Isztambul, június 30.

Bullock, H.–Mountford, J.–Stanley, R. (2001): Better Policy-Making. Policy Studies Directorate, Centre for Management and Policy Studies, Cabinet Office, London. 2001 November, <http://www.policyhub.gov.uk/docs/betterpolicymaking.pdf>.

Cseres-Gergely Zsombor – Csorba Gergely (2006): Múkincs vagy működő tőke? : Gondolatok a kutatási célú adatok hozzáférhetőségéről. – In: Közgazdasági szemle, 53. évf. 10. sz. 2006. 902–918.

⁶ A Neumann Kht internetes portálján keresztül bonyolítja le az adatkéréssel kapcsolatos feladatok egy részét, és kereshető szerkezetben közli az állami intézmények listáját, illetve az általuk szolgáltatott adatokat is (lásd: <http://www.kozadat.hu/kereso/dpresults.html?institutionId=122>).

Függelék

Útmutató adminisztratív adatok kutatási célú igényléséhez költségvetési szervek számára

1. Az adat jellegének ellenőrzése. Személyes adatról van szó? Ha nem, az adatokhoz feltétlenül hozzáférést kell kapnunk, a kérdés inkább az, hogy miként jutunk az adatokhoz.

1b. Milyen állapotban vannak az adatok? Ha elektronikus formában van, és már kiadták valakinek ugyanebben a formában, akkor költségmentesen megkaphatjuk, ha az eredeti adatgazdánál, vagy az NDA-nál rendelkezésére állnak. Szigorúan véve az előbbtől további teendő nélkül el is kérhetjük, de célszerű az adatgazdának szólni.

1c. Ha az adatok elektronikus formában vannak, még nem adták ki másnak, meg kell fizetnünk a kiadás díját. Fontos tudni, hogy teljesen reális helyzet, hogy az adatok feldolgozásához komoly munkára legyen szükség - sok államigazgatási adatbázist olyan szerkezetben tárolnak, ami nem barátja az elemzésnek. Ha igen, alapesetben a feldolgozás összes költségét ki kell fizessük. Ennek díja lehet egészen kicsit (pl. adatok lekérdezése), vagy egészen nagy (papír-alapú adatok digitalizálása). Ha a kért ár túl magasnak tűnik, kérjünk indoklást! A feldolgozás ára és az ár indoklása is közérdekű adat, így azt magyarázni kell.

2. Ha személyes adatokról van szó, akkor a törvényben meghatározott személyek kérhetik azok személyes jellegétől való megfosztását. Ezt a kérelmet nem lehet megtagadni, ha a megfoszthatóság bizonyítható. Mivel a személyes jellegtől való megfosztás maga is egy számítástechnikai művelet, kivitelezése időt és pénzt igényel. Bizonyos esetekben a művelet "sikeressége" nem egyértelmű. Ha a személyes jelleget az igényelt adatbázisban szereplő adatok (is) adják, nem csak egy azonosító, szükség lehet arra, hogy a személyes jelleg elvesztését ezeken keresztül is ellenőrizzék. Ha mindezen túl vagyunk, visszaléphetünk az 1. pontra, és kezdhetjük úgy a folyamatot, mintha eleve a személyes jellegtől megfosztott adataink lettek volna.

3. Ha több adatbázis összekapcsolására van szükségünk, elsőként az összekapcsolást kell kezdeményezni. Erre hasonló szabályok vonatkoznak, mint a személyes jelleg megszüntetésének kérelmezésére.