

2. ADMINISZTRATÍV ADATOK A KÖZGAZDASÁGI KUTATÁSBAN

2.1. A KAPCSOLT ADMINISZTRATÍV PANELADATOK HASZNÁRÓL A KRTK ADATBANK ADMIN-ADATBÁZISAINAK PÉLDÁJÁN

KÖLLŐ JÁNOS

2007. július 25-én lépett hatályba a „2007. évi CI. törvény a döntéselőkészítéshez szükséges adatok hozzáférhetőségének biztosításáról”, új korszakot nyitva a kormányzati döntés-előkészítés és hatásértékelés mellett a hazai empirikus társadalomkutatásban is. A törvény megalkotását a Pénzügyminisztérium Közgazdasági Kutatási Osztálya kezdeményezte, felismerve, hogy az államigazgatási regiszterek összekapcsolásával korábban nem látott gazdasági és méretű longitudinális adatbázisok építhetők, ugyanakkor tapasztalva azt is, hogy a minisztériumok és főhatóságok gyakran vonakodnak átadni az adataikat a kormányzati társintézményeknek. A CI. törvény *kötelezte* a kormányzati szerveket az adatok megosztására, és kialakított egy olyan eljárási rendet, ami – a személyes adatok védelmét szem előtt tartva – lehetővé tette a regiszterek, és az egyes regisztereken belül a különböző évekre vonatkozó adatok összekapcsolását. Az első próbálkozásoktól a törvényig vezető útról lásd az kutatási osztály akkori vezetőjének, a törvény kezdeményezőjének írását (Scharle, 2019)!

Az azóta eltelt évtizedben több államigazgatási szerv is végzett szakmailag igényes elemzést kapcsolt adatokkal: elsőként maga a Pénzügyminisztérium (lásd például *Elek és szerzőtársai*, 2009), később az Oktatási Hivatal, ahol ma már kapcsolt paneladatokra alapozzák a korábban kérdőíves felvételen nyugvó Diplomás Pályakövetési Rendszert (vö. *OH*, 2021) vagy a Magyar Nemzeti Bank (*Borsos–Stancsics*, 2020).

Az MTA által kezdeményezett adatkapcsolások

A törvény végrehajtási rendelete (335/2007. (XII. 13.) kormányrendelet) az adatkérésre jogosultak között sorolta fel a Magyar Tudományos Akadémia elnökét. Élve az így kapott lehetőséggel, a korábban az MTA alá tartozó Közgazdaság- és Regionális Tudományi Kutatóközpont (KRTK) Adatbankja négy ízben kezdeményezett adatkapcsolást.¹

Mint az *2.1.1. táblázatban* látható, az Admin2–4-adatbázisok a korábbi három helyett már öt adatgazda regisztereit kapcsolják össze, és a 2003. januári népesség 50 százalékat követik havi szinten rendre 2011, 2017, illetve 2021 végéig. Az Admin4 ezenfelül kiterjed a regiszterekbe 2003 után belépett személyekre is. Az adatbázisok továbbá a teljes időszakban, éves szinten követik azokat a munkáltatókat is, melyeknél legalább egy megfigyelt személy legalább egyszer dolgozott. A főbb adatcsoportok az Admin3- és az Admin4-adatbázisokban az alábbiak:

¹ Az MTA ugyanakkor nem teljes jogú szereplő: míg az államigazgatási szervek teljes regisztereket kapcsolhatnak, az MTA maximum 50 százalékos véletlen mintát vehet a népességből.

- foglalkozási jogviszony, munkakör, havi fizetett munkanapok, nyugdíj-jogosultságot megalapozó keresmények; a munkáltatóhoz való belépés és a kilépés időpontja,
- a munkáltató jellemzői (mérlegadatok, a munkaerő-összetételre vonatkozó mutatók),
- regisztráció munkanélküliként, járadék, programrészvétel, munkaerő-közvetítés (2009-től),
- nyugdíj, gyermektámogatások, jóléti ellátások,
- közmunkarészvétel (2011-től),
- részvétel az oktatásban, 2009-től teljeskörűen, azt megelőzően visszavezetésekkel; iskolai végzettség a fiataloknál; teszt eredmények és a háttérkérdőív változói az Országos Kompetenciamérésből (2008-tól),
- az egészségi állapotra utaló változók: ellátás járó és fekvőbetegként, gyógyszerkiváltás, táppénz, a betegség típusára utaló kódok (2009-től),
- állandó lakhely, járási szinten.

A használatra kész adatbázisok létrehozásáról, anonimizálásáról, előzetes teszteléséről, folyamatos javításáról, valamint a személyes adatok védelmének biztosításáról lásd Sebők Anna kötetbeli írását a 2.2. *alfejezetben*.²

2.1.1. táblázat: Kapcsolt államigazgatási panel adatbázisok a KRTK Adatbankban

	Megfigyelések		Időszak (havi adatok)	Elsődleges adatgazdák
	egyének	munkáltatók		
Admin1	4 602 000	–	2002–2009	OEP, ONYF, ÁFSZ
Admin2	4 601 999	886 425	2003–2011	OEP, ONYF, NAV, ÁFSZ, OH
Admin3	5 174 486	1 126 343	2003–2017	OEP, ONYF, NAV, NMH, OH
Admin4 ^a	6 393 158	1 537 329	2003–2021	NEAK, ONYF, NAV, NMH, OH

^a Az Admin4-adatbázis az építés fázisában van, a munkáltatói esetszám még nem ellenőrzött. Részletesen lásd: [Adatbank](#).

OEP: Országos Egészségpénztár, ONYF: Országos Nyugdíjbiztosítási Főigazgatóság.

ÁFSZ: Állami Foglalkoztatási Szolgálat, NAV: Nemzeti Adó- és Vámhivatal.

NMH: Nemzeti Munkaügyi Hivatal, OH: Oktatási Hivatal, NEAK: Nemzeti Egészség-biztosítási Alapkezelő.

Mitől más egy kapcsolt adminisztratív panel, mint a hagyományosan rendelkezésre álló adatforrások?

Nem nehéz belátni, hogy egy ilyen szerkezetű és méretű adatbázis milyen radikális változást jelent a tudományos kutatásban és a szakpolitikai elemzésben a nyolcvanas-kilencvenes évek óta leggyakrabban használt adatforrásokhoz képest.

A KSH *Munkaerő-felmérése* (MEF) rendkívül értékes adatbázis a munkaerő-piaci folyamatok tanulmányozásához, negyedéves hullámai 1992 óta szolgáltatnak információval a családokról, az oktatási részvételről, a foglalkoztatásról, a munkanélküliségről, az inaktivitásról, a transzferekről és sok minden másról. A felvételben szereplők hat negyedéven át követhetők. Ugyanakkor a MEF nem tartalmaz adatokat a bérekről, nagyon keveset mond a munkáltatókról (ágazat, méret, földrajzi elhelyezkedés), a követési időszak pedig másfél év, pontszerű,

² A KRTK adatbankról és az eddig elkészült állományokról lásd adatbank.krtk.mta.hu.

negyedéves megfigyelésekkel, szemben az Adminokkal, ahol a megfigyelések havi (vagy akár napi) szintűek és immár közel húsz éven keresztül követik a minta tagjait, részletes adatokkal szolgálva a mindenkori munkáltatóikról és munkatársaikról is. A MEF mintamérete az Admin-adatbázisokénak alig egy százaléka, alkalmatlan kisebb csoportok vizsgálatára.

Az NMH, majd 2019-től a KSH *Bértarifa-felvétele* (BT) szintén kiemelkedő jelentőségű forrás, amely 1986-ban, 1989-ben és 1992 óta minden évben, lényegében azonos szerkezetben vett fel adatokat, kezdetben a húsz-, 1995-től a tíz-, 2000-től pedig az ötfős vagy nagyobb vállalatokról, a költségvetési intézményekről és a nagyobb nonprofit szervezetekről, valamint az ott dolgozókról. (Az egyéni adatok a húszfősnél kisebb cégek és a költségvetési intézmények esetében teljeskörűen, a nagyobb munkáltatóknál pedig egy nagyjából 10 százalékos minta alapján állnak rendelkezésre). Az egyes hullámok a versenyszférából 150–170 ezer, a költségvetési és nonprofit szektorokból közel 700 ezer egyéni, valamint több mint 20 ezer munkáltatói szintű megfigyelést tartalmaznak, 2019-ig telephelyi szinten. A Bértarifa-felvétel az indulásakor világviszonylatban is az első kapcsolt munkáltatói-munkavállalói (*linked employer-employee data*, *LEED*) adatbázisok közé tartozott, és vállalati mérlegadatokkal kiegészítve addig nem látott lehetőségeket nyitott a munkáltatói magatartás, valamint a kereseti egyenlőtlenségek tanulmányozására. Ugyanakkor a felvétel nem longitudinális, a vállalatok ugyan követhetők, de az egyének nem.

A *vállalati mérlegek* és az azokból készült összeállítások (Amadeus, Opten, Orbis) jelentik a közgazdasági kutatás egy további fontos adatforrását, melyekben a vállalatok hosszabb időn át követhetők, még ha a szétválások és egyesülések kiszűrése tökéletesen meg nem oldható feladatot jelent is. Ugyanakkor nem látunk a vállalati szintnél mélyebbre, csak a pénzügyi változókat, a létszámot és az átlagbért ismerjük, ami erősen korlátozza a kutatható kérdések körét.

Az *oktatási, egészségügyi és demográfiai adatbázisok* részletes képet nyújtanak az érintett területekről, de az elemző azokból nem lát ki: tudható, például, hogy valaki milyen iskolába járt, milyen tesztteredményeket ért el, vagy milyen jellegű betegséggel küzd, de az nem, hogy ezeket a létállapotokat milyen folyamatok előzték meg vagy követték.

A *munkanélküli-nyilvántartás* szintén jó minőségű és részletes adatokkal szolgál a munkaügyi kirendeltségek ügyköréről, de nem ad képet a munkanélküliséghez vezető, illetve az onnét kivezető útról.

Az Admin-adatbázisok majdnem mindent „tudnak”, amit a fenti típusok nyújtanak, és sok mindent, amiről az ezeket használó kutatásokban álmodni sem lehet. A „majdnem” kitételt két körülmény indokolja: a fiatalokat és a munkanélkülieket leszámítva nem ismert az iskolai végzettség (szemben a MEF-fel és a BT-vel), továbbá, nem tudjuk, kik élnek egy háztartásban (szemben a MEF-fel). Az előbbi durván megbecsülhető a betöltött munkakörök alapján, az utóbbi viszont nem pótolható.

Tudományos hozam

A kapcsolt államigazgatási regiszterek összekapcsolásával létrejött longitudinális adatbázisok jelentősen hozzájárultak annak a „hitelességi forradalomnak” (*credibility revolution*) a kibontakozásához, ami korábban nehezen teljesíthetőnek számító, gyakran teljesíthetetlen követelményeket állított az oksági kapcsolatok verifikálására törekvő empirikus kutatás elé (*Angrist–Pischke*, 2010). Az új standardoknak ilyen adatok birtokában könnyebb megfelelni: mód nyílik a szintek helyett (mellett) a megfigyelési egységeken belüli változások tanulmányozására, valamint a meg nem figyelt heterogenitás hatásának kiszűrésére akár többféle fix hatás egyidejű alkalmazásával. A nagy- és változóban gazdag mintákból könnyebb megfelelő kontrollcsoportot választani.

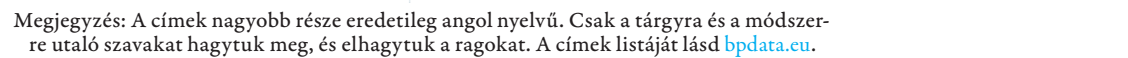
Az államigazgatási adatok nemcsak mélyebb elemzést tesznek lehetővé, hanem nagyságrendileg bővítették is a kutatható kérdések körét. Egyidejűleg figyelnünk meg olyan, korábban csak külön-külön tanulmányozható jellemzőket, mint például a foglalkoztatás, a keresetek, a munkahely- és foglalkozásváltás, a munkanélküliség, a jóléti juttatások, az egészségi állapot, az oktatási részvétel és az iskolai teljesítmény. Olyan szinten látunk a gyár kerítésén vagy az intézmény kapuján belülre, ami elképzelhetetlen volt a korábban elérhető adatbázisokban. Megvizsgálhatók a munkavállalási korú inaktívak, akik semmilyen államigazgatási regiszterben nem szerepelnek, csakis az összekapcsolt adatokban azonosíthatók, a tanulók, a foglalkoztatottak, a munkanélküliek (továbbá, ha szükséges, a nyugdíjasok és a gyermektámogatásban részesülők) kizárásával.

Ami talán a legfontosabb: az adatbázisok rendkívüli mérete lehetővé teszi a legkülönbözőbb kritikus események és kisebb csoportok megfigyelését: embereket, akik egy időben hagyták el az iskolát, váltak munkanélkülivé vagy helyezkedtek el, akik egyszerre lettek betegek; vállalatokat, melyek egyszerre bővítették vagy csökkentették a kibocsátásukat vagy a létszámukat és a többi. Az időablak széleit leszámítva, minden pillanatban ismerjük a megfigyelt emberek és cégek előtörténetét és a sorsuk későbbi alakulását, az esetszámok pedig még olyan szűk csoportok esetében is tekintélyesek, mint például a sorkatonáké, a korai nyugdíjba vonult rendőröké vagy a börtönviseltéké.

Kutatások az Admin1–3 adatbázisokkal

Az 2.1.1. ábra szófelhője azokon az Admin-adatokon alapuló kutatásokon nyugszik, melyekből már megszülettek a publikációk vagy legalább az azokat megelőző műhelytanulmányok. A felhő a publikációk címében szereplő szavakon alapul, a címszavak mérete arányos az előfordulási gyakoriságukkal. A 48 publikáció részletes listáját lásd a bpdata.eu tárhelyen, ahol az egyes tanulmányok tartalmi összefoglalói is elérhetők (ha vannak) a bibliográfiai tételhez kapcsolt linke kattintva. Az 2.1.1. ábrában közölt, csak a kutatások tárgyára utaló lista is jól érzékelteti, hogy milyen sokféle elemzési lehetőséget teremtettek a kapcsolt panelek.

menedzserek aktivitás

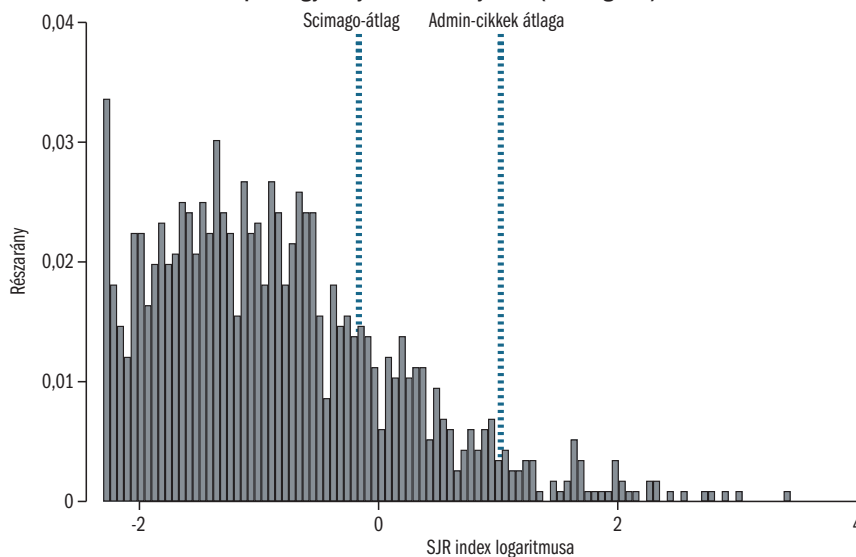


A világvizonylatban is kiemelkedőnek mondható Admin3-adatbázis csupán 2021-től érhető el (a tesztelése és javítása még ezen sorok írásakor is tart), ezért az azon folyó kutatások közül még nagyon kevés érthette meg a folyóiratban való megjelenést. A Scimagóban jegyzett folyóiratokban eddig 21, az Admin-panelesen alapuló cikk jelent meg, az átlagosnál sokkal jobban jegyzett helyeken. Ennek alátámasztására két statisztika is említhető. Az Admin-cikkek IS (*article influence score*) indikátora 3,7, ami az átlagosnál (IS = 1,0) sokkal magasabb érték, arra utal, hogy az őket befogadó folyóiratok idézettséggel mért tudományos hatása az átlagosnál sokkal erősebb.³ A másik választott mutató az SJR (Scimago Journal Rank), melynek átlagos értéke a Scimagóban mért 1160 közgazdasági, ökonometria és pénzügyi folyóirat körében 0,85, 1,85-

ös szórással. Az Admin-cikkek esetében az SJR érték 2,78, az átlagosnál több mint egy szórás egységgel magasabb.⁴

Az Admin-cikkek mezőnybeli elhelyezkedését jól mutatja a 2.1.2. ábra hisztogramja, melynek vízszintes tengelyén az SJR-értékek láthatók (természetes alapú logaritmusban, hogy a szélső esetek ne nyomják agyon a hisztogramot), míg az oszlopok magassága az adott SJR-értékhez tartozó folyóiratok mezőnybeli részarányát mutatják. A két szaggatott oszlop közül a bal oldali a Scimago-főátlagot, a jobb oldali az Admin-cikkek átlagát jelzik. Látható, hogy az SJR-értékek eloszlása erősen aszimmetrikus, ennek köszönhető, hogy a Scimago-ban jegyzett folyóiratoknak mindössze öt százaléka ért el az Admin-publikációk átlaga fölötti SJR-indexet.

2.1.2. ábra: Az Admin-adatokon alapuló publikációkat közlő folyóiratok átlagos SJR-értéke a Scimago-ban jegyzett, 1160 közgazdasági, ökonometriai és pénzügyi folyóirat mezőnyében (hisztogram)



Adatforrás: [Scimago](#).

A tudománymetriai mutatók a jövőben minden bizonnyal javulni fognak az Admin3 és Admin4 gazdagságának köszönhetően, már csak azért is, mert a különleges adatháttér számos külföldi kutatót vonz. Az említett publikációk és prepublikációk szerzői listájában 34 külföldi és 25 külföldön dolgozó magyar kutató neve fordul elő. Utóbbi arra utal, hogy a gazdag adatháttér segít fenntartani a külföldön dolgozó magyar kutatók anyaországához fűződő kapcsolatát.

Költségek

Egy kapcsolt adminisztratív panel sokba kerül: az adatgazdáknak és az összekapcsolást végző koordinátornak fizetett 15–35 millió forintot felül lega-

⁴ Az SJR mutató a folyóiratok tudományos hatását méri, amely mind a folyóirat által kapott hivatkozások számát, mind pedig azoknak a folyóiratoknak a fontosságát vagy presztízsét jelenti, amelyekből az idézetek származnak. Logikáját tekintve a sakkban használt Élő-pontszámhoz hasonlít.

lább ennyit emészt fel a kutatásra kész adatbázis több évig tartó felépítése, és akkor még nem számoltunk a rendkívüli méretek kezeléséhez szükséges gépi kapacitás és az erre alkalmas szoftverek költségeivel. Ám ha belegondolunk, mennyibe kerülne a listánkban szereplő kutatások adatainak előállítására kérdőíves módszerekkel, akár csak tizedekkora mintákon, akkor az adminisztratív panelt kifejezetten olcsónak mondhatjuk. Az Admin3-adatbázist e sorok írásakor 64 kutatási projekt használja. Ha százmillióس bekerülési költséggel számolunk, az projektenként másfél millióس adatköltséget jelent – néhány fő-kuszcsoportos beszélgetés árát.

Gyakorlati tudnivalók

Az Admin-adatbázisok alapjául szolgáló, spell szerkezetű forrásfájlok köz-adatnak minősülnek, tudományos célra bárki számára elérhetők.⁵ A KRTK Adatbank által kutatásra előkészített verziók védett szerveren, dedikált távoli eléréssel elemezhetők, a részletesebb bontású egészségügyi adatok pedig szigorúan ellenőrzött adatszobában kutathatók. Az adatbázisokat az intézményi felhasználók térítés ellenében érhetik el, melynek mértéke a szerverszolgáltatás és az esetlegesen szükséges többletkapacitás költségének függvénye. Az adatokat az akkreditált kutatóhelyeken dolgozó egyéni kutatók és szakdolgozók a rendelkezésre álló gépi kapacitás korlátain belül – sorállásos alapon – ingyenesen használhatják, kizárólag tudományos célra. A KRTK kutatói és az általuk vezetett projektek elsőbbséget élveznek.

Az adatokat Stata formátumban tároljuk, de elérhetők az R és Python programok számára is. Felhasználói támogatást a Stata használóknak nyújtunk. Az adatokat már csak a méretek miatt (az Admin3-adatbázis a kiegészítő modulokkal együtt 120 gigabájt) és adatvédelmi okokból sem adjuk ki.

Komoly helyen nem publikálható olyan kutatás, amelynek az adatai nem hozzáférhetők. Ennek biztosítására a becslési mintákat – a minimálisan szükséges adattartalommal – elérhetővé tesszük a lektorok, illetve az eredményeket replikálni kívánó kutatók számára, kérés esetén az anonimitás megőrzésével.

Az adminisztratív paneladatbázisok jövője

2021 nyarán, a nemzeti adatvagyonról szóló 2021. évi XCI. törvény elfogadásával hatályon kívül helyezték a 2007.CI. törvényt. Az államigazgatási adatok kezelése az ekkor alapított Nemzeti Adatvagyon Ügynökség (<https://www.navu.hu/>) kezébe került. Sem a törvény, sem a kapcsolódó végrehajtási rendelet⁶ nem tér ki a tudományos célú adathozzáférés kérdésére, a rendelet nem határozza meg tételesen az adatkapcsolás kezdeményezésére jogosult intézmények (személyek) körét, és nem írja le részletesen magát az eljárást.

Nagyon reméljük, hogy nem fog elveszni az az előny, amelyre a magyar társadalomkutatás a kapcsolódó adminisztratív panelek révén tett szert: a világviszonylatban is kiemelkedő minőségű adatok továbbra is vonzani fogják a hazai

⁵ A forrásfájlok megfigyelési egysége egy eleje- és végdátumból, valamint a kettő közötti történetre utaló változóból álló „spell”. Ebből az adathalmazból állít elő a KRTK Adatbank egy egyén-hónap megfigyeléseket tartalmazó, fix formátumú adatmátrixot, melynek egymilliónál több sora és több száz oszlopa van.

⁶ Lásd a nemzeti adatvagyon hasznosításával összefüggő egyes részletszabályokról szóló 607/2021. (XI. 5.) kormányrendelet.

és nemzetközi kutatói közösséget, segítenek a nemzetközi pályázati források elnyerésében, és komoly tudományos eredményeket alapoznak meg. Ugyancsak reméljük, hogy nem kerül parlagra az a módszertani tudás, amelyet az adatbázisokat felépítő szakemberek és az azokat használó kutatók halmoztak fel az elmúlt tíz évben, valamint az a sok száz milliós számítási kapacitás, amelyet kifejezetten a nagy méretű paneladatbázisok kezelésére hozott létre a KRTK, az MTA, majd az ELKH támogatásával. Bízunk abban is, hogy ezt a tudást a mainál nagyobb mértékben fogja hasznosítani az államigazgatás is, az adatkezelésben, a tényeken alapuló döntés-előkészítésben és szakszerű hatásvizsgálatokban.

Hivatkozások

- ANGRIST, J. D.–PISCHKE, J.-S. (2010): [The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics](#). Journal of Economic Perspectives, Vol. 24. No. 2. 3–30. o.
- BORSOS ANDRÁS–STANCSICS, M. (2020): Unfolding the Hidden Structure of the Hungarian Multi-Layer Firm Network. Occasional Papers, No. 139. Magyar Nemzeti Bank, Budapest.
- ELEK PÉTER–SCHARLE ÁGOTA–SZABÓ BÁLINT–SZABÓ PÉTER ANDRÁS (2009): A bérekhez kapcsolódó adóeltitkolás Magyarországon. Közpénzügyi Füzetek, No. 23. ELTE TáTK, április.
- OH (2021): [Diplomás pályakövetési rendszer 2020 – adminisztratív adatbázisok egyesítése](#). Gyorsjelentés. Oktatási Hivatal, Budapest.
- SCHARLE ÁGOTA (2019). [Hungary: a case study on improving access to administrative data in a low-trust environment](#). Megjelent: *Crato, N–Paruolo, P.* (szerk.): Data-Driven Policy Impact Evaluation: How Access to Microdata is Transforming Policy Design. Springer, Cham, 119–130. o.

K2.1. Az adminisztratív adatbázisok használata a Budapesti Corvinus Egyetemen

TELEGDY ÁLMOS, LENGYEL BALÁZS & KESZEY TAMARA

Az adatok intenzív használatára épülő kutatások megváltoztatják az üzleti, közgazdaság- és társadalomtudományokat, illetve azok oktatását. Az adatok gyűjtésének, tárolásának, feldolgozásának, elemzésének és szolgáltatássá vagy termelésé alakításának módjára vonatkozó új készségek az üzleti oktatásba is bekerültek. Ezek a képzetek azonban ritkák a közép-kelet-európai régióban. Néhány példán kívül (CEU Bécs, Varsói Közgazdasági Egyetem, IBS Budapest) a környező országokban hiányzik az ilyen készségek kritikus tömege. Ennek az igénynek a kielégítésére a Budapesti Corvinus Egyetem Megújulási Programjában stratégiai célként tekint, és hozott előremutató intézkedéseket, mint például az üzleti adattudomány képzés elindítása.

A minőségi adatokhoz való hozzáférés növelése érdekében a Budapesti Corvinus Egyetem és a Közgazdaság- és Regionális Tudományi Kutatóközpont (KRTK) Adatbank között megállapodás jött létre, amely feljogosítja az egyetem összes dolgozóját és hallgatóját, hogy az Adatbank az adminisztratív adatbázisait használhassa.

A hazai adminisztratív közadatokból készült, a kapcsolt munkáltatói és munkavállalói adatokat tartalmazó KRTK-adatbázis használata nagyon fontos az egyetem számára, és két ponton is jól illeszkedik a hosszú távú stratégiai céljaihoz. Először is, az egyetem nagy hangsúlyt kíván helyezni a nemzetközi mércével is kimagasló kutatásra, ami elképzelhetetlen jó minőségű, a nemzetközi közgazdasági kutatá-

sokban is használt, mégis egyedi adatbázisok nélkül. Másodszor, az egyetem nagyon fontos feladata, a minőségi, modern közgazdasági és üzleti oktatás sem elképzelhető olyan adatbázisok nélkül, amelyeken az adatelemzés technikai eljárási módjai, és a hallgatók szembesülhetnek az elemzés nehézségeivel is. Ezekre kiválóan alkalmasak a KRTK adminisztratív adatbázisai, hiszen a hazai közgazdászok sokat használja ezeket, így jelentős tudás halmozódott fel velük kapcsolatban, amiből a hallgatók profitálhatnak. Ugyanakkor az adatok csak védett környezetben kutathatók, így a hallgatók megtanulhatják azt is, hogyan kell érzékeny adatokkal dolgozni.

A KRTK adminisztratív adatbázisait eddig a BCE közgazdasági elemző mesterszakos hallgatói használták elsősorban, de valószínűsíthető, hogy a közeljövőben sokkal több szakon fogják őket alkalmazni az oktatáshoz. A közeljövőben a doktori iskolákban tervezzük további használatukat.

Összegezve elmondható, hogy az adminisztratív adatok használata a Budapesti Corvinus Egyetemen előre láthatóan hozzájárul majd az adatok elemzéséhez értő új generációk kineveléséhez. Ez javítja majd a hazai vállalatok döntéshozatali képességeit, és az adatelemzésből keletkező tudás döntéshozásba történő becsatornázásával segíti a kormányzat és a nemzeti bank szakpolitikájának kialakítását. Ezenkívül az adatok segítik a jó minőségű szakdolgozatok és doktori disszertációk írását, valamint az egyetemi kutató közösséget a rangos nemzetközi publikációk közlésében.