

A KSH Könyvtár társadalomtudományi hivatkozásgyűjtő adatbázisa, a Mathias

A tudományos teljesítmény méréséhez ma már a társadalomtudományok területén is elengedhetetlenül hozzátartozik a hivatkozásgyűjtés. Nemzetközi szinten (különösen a természettudományokban) a hivatkozások kereshetőségét elektronikus adatbázisok segítségével több-kevesebb sikerrel megoldották. Magyar nyelvterületen egységes, több évtizedre visszamenően digitalizálva rendelkezésre álló folyóirat-állomány hiányában hasonló adatbázisok nem segítik a kutatók munkáját és értékelését. Egységes koncepció hiányában a további lemaradás elkerülése érdekében teret kell engedni egy-egy tudományterületen önálló adatbázisok kiépítésének – természetesen az együttműködés lehetőségét sem kizárva más intézményekkel, kezdeményezésekkel. Cikkünkben a KSH Könyvtár hivatkozásgyűjtő adatbázisának létrehozásáról, a fejlesztés során felmerült problémákról számolunk be.

Nemzetközi vonatkozásban először kötetkatalógusban tették közzé a hivatkozásokat, napjainkban pedig elektronikus adatbázisokban érhetők el. Ezek közül a legismertebb a *Web of Science (WoS)* és a *Scopus*, amely egyes szakterületeken magyar vonatkozásban is kiválóan alkalmazható. Ilyen diszciplína például az orvostudomány, amelynek művelőitől elvárás, hogy eredményeiket nemzetközi folyóiratokban is publikálják. Más, inkább magyar nyelvű publikációkat előállító tudományterületek munkatársai azonban azzal a problémával szembesülnek, hogy jóformán csak „lapozgatós” módszerrel tudják a rájuk vonatkozó hivatkozásokat visszakeresni. A *Magyar Tudományos Művek Tára (MTMT)* alkalmas ugyan a hivatkozások kezelésére, azonban a kutatóknak vagy adminisztrátoraiknak nagyrészt maguknak kell azokat felkutatni. Ez a probléma áll fenn a társadalomtudományok, közelebbről a statisztika területén is. A *Központi Statisztikai Hivatal Könyvtár* ezért kezdte el kiépíteni 2012-ben a *Magyar társadalomtudományi hivatkozások adatbázisát*, a *Mathias*t, amely a magyar nyelvű társadalomtudományi folyóiratok hivatkozásait rögzíti és teszi visszakereshetővé. Az adatbázis célja, hogy a társadalomtudományi hivatkozásokat teljességre törekvően, módszeresen gyűjtve és feldolgozva kereshetővé tegye.

A munka első fázisában a *Központi Statisztikai Hivatal* élő, illetve nemrég megszűnt folyóiratainak feldolgozására került sor. Ezeken belül is a XX. század második felében megjelent magyar vonatkozású művekre történő hivatkozásokat rögzíti a könyvtár. Magyar vonatkozás alatt azt értjük, hogy a hivatkozott szerzők közül legalább egy magyar kötődésű. (Így került például a figyelt szerzők közé a Kiss László néven született Leslie Kish (1910–2000), aki tizenhat éves korától haláláig Amerikában élt.) Hivatkozásnak tekintünk továbbá minden műre történő utalást, jelenjen az meg lábjegyzetben, képaláírásban vagy bibliográfiában. Nem rögzíti a könyvtár az 1950 előtt megjelent művekre vonatkozó hivatkozásokat. Mindezek a megkötések az adatbázis gyakorlati alkalmazása miatt történtek. A *Mathias* azonban nemcsak a ma aktív kutatók számára kínál lehetőséget, hanem tudománytörténeti szempontból is új forrást kínál. „Melléktermékként” a XX. század első felében vagy korábban jelentős

teljesítményt nyújtó tizenegy szakember (*Buday László* [1873–1925], *Jekelfalussy József* [1849–1901], *Keleti Károly* [1833–1892], *Kovács Alajos* [1877–1963], *Kőrösy József* [1844–1906], *Laky Dezső* [1887–1962], *Ráth Zoltán* [1863–1902], *Szél Tivadar* [1893–1964], *Thirring Gusztáv* [1861–1941], *Thirring Lajos* [1899–1982], *Vargha Gyula* [1853–1929]) tudományos utóélete is nyomon követhető segítségével. (Igyekeztünk azokat a tudósokat kiválasztani, akiknek a munkásságával kapcsolatban jelenleg is folynak a kutatások. A magunk döntően statisztikatörténeti érdeklődését a *Mathias* tervezésekor igyekeztünk háttérbe szorítani.)

A könyvtár munkatársai 2013 márciusáig a fenti szempontok alapján feldolgozták a *Statisztikai Szemle* 1990 és 2012, a *Demográfia* 2000 és 2012, a *Területi Statisztika* 1998 és 2012, illetve a *Gazdaság és Statisztika* 1992 és 2006 közötti évfolyamait. A *Statisztikai Szemle* 1923 óta – a háborús időszakot leszámítva – havi rendszerességgel megjelenő lektorált folyóirat, a magyar statisztikatudomány legfontosabb fóruma. Az itt megjelent tanulmányok lefedik a társadalomtudományok, azon belül is kiemelten a statisztika, a közgazdaság-tudomány és a demográfia egyes ágait, illetve betekintést nyújtanak az egyéb tudásterületek statisztikai alkalmazásaiba. A *Területi Statisztika* nemcsak a szűkebb értelemben vett statisztika, hanem a regionális tudomány valamennyi területéről közöl cikkeket. Az 1957-ben alapított *Demográfia* a *KSH Népeségtudományi Kutatóintézetének* tudományos folyóirata. A *Gazdaság és Statisztika* 2006-os megszűnéséig közel hat évtizeden át az ipar- és gazdaságstatisztikusok elsőszámú publikációs fóruma volt.

A feldolgozandó dokumentumok körének kijelölése után az adatbázis szerkezetének megtervezése jelentette a következő feladatot. Már a kezdet kezdetén eldöntöttük, hogy a lehető legegyszerűbb megoldást választjuk, és mind az adatfeltöltéshez, mind a lekérdezéshez táblázatkezelő programot fogunk használni. Tettük ezt többek között azért, mert a *Mathias* építésének első évében nem számoltunk néhány tízezernél több rekorddal. Ekkora mennyiség kényelmesen kezelhető akár *Excel*el is. Az egyszerűen konvertálható formátum mellett szólt az is, hogy ez később jócskán megkönnyítheti az együttműködést más adatbázisgazdákkal.

Egy-egy rekord értelemszerűen két részre oszlik: egyfelől a hivatkozó, másfelől a hivatkozott mű leírására. A hivatkozó művek sokszínűsége viszonylag gyorsan felülírta az eredeti terveket: jelenleg összesen huszonhárom mező áll rendelkezésre egy-egy tétel rögzítéséhez. Bár sokáig vívódtunk, külön mezőkbe soroljuk-e a társszerzőket, végül – meglehetősen vitatható módon – csupán írásjelekkel választottuk el őket egymástól.

A szerzőség megállapításakor nehézséget okoznak a KSH 1989 előtt megjelent összeállításai. Ezeknél szerzőségi közlésként többnyire csupán a Központi Statisztikai Hivatal, esetleg egy-egy főosztály megnevezése szerepel. Az előszókból és további rejtett adatokból azonban olykor kideríthető az összeállító, szerkesztő személye. Ebből a megfontolásból a KSH kiadványaira történő hivatkozások minden esetben szerepelnek a *Mathias*ban. A későbbiekben szükség lesz ezeknek a rekordoknak a folyamatos ellenőrzésére, javítására, bővítésére, a katalógussal való összevetésére.

Természetesen számos egyéb problémával is találkoztunk az adatrögzítés során. Több folyóiratnál is megfigyelhető gyakorlat, hogy akár egyazon számon, de egyetlen cikkben belül is két-háromféleképpen hivatkozik. Egyes szerzők az adattárakat eleve nem tekintik hivatkozásra érdemesnek. Jó esetben legalább egy táblázat vagy egy ábra alatt feltüntetik az összeállítás címét, de pontos bibliográfiai leírást nem mellékelnek hozzá, és az irodalomjegyzékbe sem veszik fel. A kiadvány címe helyett több esetben a sorozatcím és a kötetszám szerepel. Néhányszor a hivatkozott periodikum nem szokványos megjelenési adataival gyűlt meg a bajunk: a *Medvetánc* 1986. 4. – 1987. 1. összevont számát például a legnagyobb jóindulattal sem tudtuk kielégítő módon megjeleníteni az adatbázisban.

Terveink szerint a következő hónapokban az említett folyóiratok kurrens és régebbi számainak feldolgoása mellett megkezdjük a statisztikai-demográfiai tárgyú kutatási jelentések, monográfiák, Festschriftek, tanulmány- és konferenciakötetek hivatkozásanyagának rögzítését is.