

A MATARKA és az egyedi azonosítók, különös tekintettel a szerzőkre

Burmeister Erzsébet, Miskolci Egyetem, Könyvtár, Levéltár Múzeum

erzsi@uni-miskolc.hu

A publikációk egyre növekvő tömege már az elektronikus korszak előtt szükségessé tette a bibliográfiai adatok leírásának szabványosítását, egyedi azonosítók bevezetését, jó példa erre az ISBN, ISSN, stb. számok megjelenése a XX. század második felében.

Az adatbázisok és az internet világában talán még fontosabb a szerzők és az elektronikus dokumentumok egyedi azonosítása a gyors fellelhetőség érdekében (a MATARKA adatbázis esetében csak ennek a két fajta azonosítónak van jelentősége, ezért a többivel nem foglalkozom).

A Miskolci Egyetem könyvtára 2002 óta fejleszti a MATARKA (www.matarka.hu) adatbázist és szolgáltatást. Az adatbázis szerzői neveinek száma 2016-ra meghaladja a 330 ezret, a címekből a teljes szövegekre mutató ugrópontok (linkek száma) a 600 ezer felett. Egyre sürgetőbbé válik, hogy az azonos nevű szerzőket szét tudjuk választani (és nemcsak bibliográfiai leírási eszközök segítségével, vagyis születési és foglalkozási adatok megjelenítésével) és szerzőinkről „több adatot” tudjanak használoink kattintással elérni. Az elektronikus dokumentumok esetében pedig egyre inkább a megbízhatóbb DOI-ból képezhető URL-ekre szeretnénk támaszkodni.

Elektronikus dokumentumok azonosítása: a DOI

A DOI (Digital Object Identifier) a dokumentumot azonosítja, nem a helyét. Az 1997-ben USA-ban megalakult Nemzetközi DOI Alapítvány (IDF = International DOI Foundation) a DOI „gazdája”, benyújtó ügynöksége a DOI ISO 26324:2012 szabványnak, ő a DOI rendszer működtetője, ő fogja össze a DOI ügynökségeket és az alapítvány többi tagját. A DOI ügynökségek felelősek a DOI rendszer integritásáért. A DOI számokat az ügynökségek adják ki a DOI számokért folyamodó intézményekkel, főleg kiadókkal együttműködve, majd az egyediség ellenőrzése után „élesítik”.

DOI használat a MATARKA-ban

A MATARKA adatbázisban a Networkshop előadás időpontjában 563307 cikk mellett található valamilyen teljes szövegre (vagy egész kis hányadban összefoglalóra) vezető link. Ebből 301173 az Elektronikus Periodika Archívumra, 41470 az MTA Real repozitóriumára, 19793 pedig a Hungaricana-ba vezet el. A felsoroltak a *megbízhatóak*. Az összes többi, vagyis 200871 link különböző kiadók gyakran változó felépítésű URL-jeit jelenti. Ha azonban a kiadó elhatározza magát külső nyomásra vagy saját felismerése alapján arra, hogy alkalmazza a DOI-t, akkor a bizonytanság megszűnik és a DOI ügynökség felé megadott, a DOI-hoz összekapcsolódó elérhetőséget biztosítani kell.

2015 szeptemberétől a MATARKA adatbeviteli űrlapján megjelent a DOI mező, és egyidejűleg a tartalomjegyzékben és a találati listákban is a felhasználók számára. 2016 márciusában a DOI-val ellátott cikkek száma 2420 és 27 különböző DOI előtag van, vagyis 27 különböző kiadó szervezettől származnak a DOI-s cikkek.

Közgazdasági szemle

Teljes szöveg: [Elektronikus Periodika Archivum](#) - 1995/1 számtól
Honlap: www.kozgazdasagiszemle.hu
RSS

62. évf. 9. sz. / 2015

Kosárba

Varga Attila - Sebestyén Tamás

Innováció Kelet-Közép-Európában. Az EU keretprogramjaiban való részvétel szerepe az innovációs teljesítményben

Teljes szöveg-DOI: [10.18414/KSZ.2015.9.881](#)

881-908

ESSZÉ

Kornai János

Milyen is hát a tőke a 21. században? Megjegyzések Piketty könyvéhez

Teljes szöveg-DOI: [10.18414/KSZ.2015.9.909](#)

909-942

Medvegyev Péter

Nagyobb kockázat, nagyobb növekedési ütem – kockázat nélkül nincs növekedés! Hozzászólás Thomas Piketty A tőke a 21. században című könyvéhez

943-951

Szerzők: Juhász Judit - Bajmóczy Zoltán - Gébert Judit - Málóvics György

Szegénység, képességek, lehetőségek. Szegénységkezelési szempontok Amartya Sen elméletének tükrében

Teljes szöveg (PDF)

Közgazdasági szemle, 2015. (62. évf.) 5. sz. 544-564. old.

Teljes szöveg: Csak a felsőoktatási intézmények számára: [Arcanum Digitális Tudománytár \(1877-2009\)](#)

[Elektronikus Periodika Archivum](#) 1995/1 számtól

Szerzők: Gébert Judit

Mit is kell fenntartani? Fenntarthatóság a képességszemlélet perspektívájából

Teljes szöveg-DOI: [10.18414/KSZ.2015.9.972](#)

Közgazdasági szemle, 2015. (62. évf.) 9. sz. 972-989. old.

Teljes szöveg: Csak a felsőoktatási intézmények számára: [Arcanum Digitális Tudománytár \(1877-2009\)](#)

[Elektronikus Periodika Archivum](#) 1995/1 számtól

Szerzők: Bajmóczy Zoltán - Gébert Judit

Arguments for deliberative participation in local economic development

[Acta oeconomica](#), 2014. (Vol. 64.) no. 3. 313-334. old.

1. ábra: DOI-t tartalmazó tartalomjegyzék és találati lista

Szerzői azonosítók

Szerzői azonosító sokféle van, többé-kevésbé ismertek és használatosak Magyarországon. A MATARKA adatbázisban 2016-ban vezettünk be néhányat. Adódott a *VIAF*, mint a legnagyobb szerzőket azonosító adatbázis, az *ORCID* és az *MTMT* azonosító, mert ezek a felsőoktatási intézmények számára fontosak, valamint a *nefter.hu* szolgáltatás URL-jei. A jövőben várható magyar szerzői névtér azonosító megjelenése, erre azonban még egy ideig várni kell.

Egyedi azonosító megnevezése	Egyedi azonosító felépítése
Scopus author identifier	7004182651
Google Scholar ID	nem jelenik meg
ISNI	0000 0000 7925 6201
MS Academic Search	nem jelenik meg
MTMT	nem jelenik meg, csak a szerző látja, 10019226
ORCID	0000-0003-4419-142X
Researcher-ID (Web of Science)	A-7208-2011
VIAF	12991180

A szerzői azonosítók közül röviden csak a külföldieket mutatom be.

VIAF = The Virtual International Authority File

Létrehozását már 1998-ban javasolták, végül is 2003-ban jött létre VIAF az OCLC, Library of Congress, Deutsche Nationalbibliothek és Bibliothèque nationale de France közös projektjeként, 2012-ben lett az OCLC nyilvánosan elérhető szolgáltatása (www.viaf.org). „A virtuális gyűjtemény célja, hogy az egyes nemzeti besorolási adatállományokat egyetlen, egységes besorolási adatállományként szolgáltatassa.”[3] Partnerintézményei főleg nemzeti könyvtárak, az OSZK is, 2013-ban 32 tag volt. Több név authority fájljt egyesít egyetlen névszolgáltatásba. Szerzői neveken kívül egyéb authority állományokat is épít: testületi neveket, földrajzi neveket, stb.

ORCID

Az Open Researcher and Contributor ID rövidítése az ORCID (<http://orcid.org/>).

Az ORCID közösség tagjai egyéni kutatók, egyetemek, kutató intézmények, kutatásokat finanszírozó intézmények, kiadók, tudományos társaságok, repozitóriumok. Az intézmény tagok előfizetési díjat fizetnek (a személyek nem), s ezért különböző szolgáltatásokat kapnak. Májig több mint 1,7 millió azonosítót adtak ki.

Két alapvető funkciót kínál:

Regisztráció egyedi szerzői azonosítóhoz és a kutatói tevékenységek (pl. publikációk) nyilvántartásához

Application Program Interface-k (API), melyek a rendszerek közötti kommunikációt teszik lehetővé.

A szerzők annyi adatot töltenek fel magukról, amennyit jónak látnak. Lehetőségük van bemutatni saját publikációik listáját azok elérhetőségével együtt, bemutathatják életútjukat, oktatási és kutatói tevékenységüket, stb

Szerzői azonosítók bevezetése a MATARKA-ban

Nagyon sok szerző van már a MATARKA-ban (300000 felett), ezek 60 % magyar. Az azonosítók bevezetésének több haszna is van:

- szerzőnevek egységesítése, utalozásuk, összekapcsolásuk
- hibák, elírások javítása
- kapcsolatteremtés más adatbázisokkal (VIAF, ORCID, MTMT, nevpont.hu), hisz ugrópontok vezetnek hozzájuk
- azonos nevű szerzők szétválasztása

Nehéz és időigényes az azonos nevű szerző szétválasztása, azonosítókkal ellátása, mert sok forrást át kell vizsgálni a hibák elkerülése érdekében. Nem engedhető meg, hogy a mérnök Kiss József cikke a fizikus Kiss Józsefhez kerüljön (légből kapott példa!). Az azonosítók felvétele természetesen technikailag egyszerű: külön segédprogram segít ebben.

A találati listákban, ill. tartalomjegyzékekben kis képek jelennek meg a szerzők neve mellett, amelyekre kattintva át lehet lépni a szolgáltató (VIAF, ORCID, MTMT, nevpont) megfelelő oldalára.

Amit meg kellett oldani: a tényleges ellenőrzés, vagyis, hogy egy adott napon bekerült szerző szét van-e már választva, csak utólag történhet. A MATARKA segédprogramjai között van egy napi ellenőrzéseket végző, melynek most már 4 feladata van:

1. bekerült új szerzőnevek és szavak ellenőrzése
2. URL-ek formai ellenőrzése
3. DOI-k ellenőrzése
4. szétválasztott szerzőkből került-e újabb be

Az ellenőrzés által kidobott hibák javítása az ellenőrzést végző könyvtáros feladata és ő dönti azt is el, hogy az újonnan bekerült azonos nevű szerzős cikk melyik szerzőhöz tartozik. Lásd 2. ábra.

EGYEDI SZERZŐ ELLENŐRZÉS

Újból bekerül olyan szerző, aki már egyedi azonosító és több van belőle
Mai dátum -- 2016-03-03

Hány szerző: 232

Kattints be azt a szerzőt, akihez az új cikket hozzá kell rendelni, a pirossal jelzettnél van jelenleg.
Ha a piros maradjon, akkor is kattints be.

14264-- (2016-03-03) -- Apor Péter, orvos -- Cikkszám:2359926: Kardiometabolikus betegségek megelőzése és gyógyít (Füzetszám:137249)

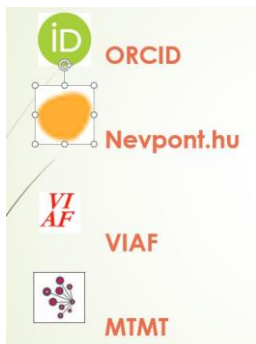
- ☒ --Apor Péter (orvos)
- ☐ --Apor Péter (történész)
- ☐ --Apor Péter (mérnök)

Bevitel

2. ábra: Részlet a napi ellenőrzést végző segédprogram futtatási listájából.

Az adatbázisban - nem úgy, mint a DOI-nál - a szerzői azonosítók nem számokként jelennek meg, hanem kis markerek (ikonok, logók) állnak a szerzők neve mellett a találati listákban, amelyek elvezetnek a megfelelő szolgáltatáshoz.

Alább a bal oldali képen látható a szolgáltatások és a hozzájuk tartozó markerek listája, a jobb oldali képen pedig egy találati lista, ahol láthatók a markerek.



Kaptay György miskolci professzor neve mellett VIAF, MTMT és ORCID logó is van. Ez aránylag ritka eset, de előfordul. A legjellemzőbb az, hogy egy vagy két markerre lehet kattintani és ennek révén több információt megszerezni a szerzőről.



3. ábra: Az egyedi azonosítókhoz tartozó markerek

4. ábra: Markerek a szerzők neve mellett

A szerzők szétválasztása megbízható módon nem mindig lehetséges. Ilyenkor megoldás lehet az, hogy a biztosan elkülöníthető szerzőket leválasztjuk és az összes többi cikket foglalkozás megadás nélkül adjuk hozzá a szerzőhöz.

Erre mutat szemléletes példát az alábbi ábra:

Csorba József szerzőnév hatszor foglalkozással szerepel, egyszer foglalkozás nélkül.

Találatok	
szerzőnevekben	
8 találat	
Találatok listája	
Csorba József (Információs tanácsadó)	33 cikk
Csorba József (méhész)	2 cikk
Csorba József (addiktológus)	7 cikk
Csorba József (biotechnológus)	5 cikk
Csorba József (orvos)	2 cikk
Csorba József (mérnök)	2 cikk
Csorba József	5 cikk
Csorba Józsefné	6 cikk

5. ábra: Szerzőnevek szétválasztás után

Irodalom

- (1) A szolgáltatások honlapjai: doi.org, viaf.org, mtmt.hu, nevpont.hu
- (2) Nemzeti névterek. <http://syi.hu/ns/general.html> [utolsó letöltés: 2015.12.10.]