

Kézirat. Megjelent a Könyvtári Figyelő 2017. 3. számában

ORCID – egy újabb szerző-azonosító tudományos közleményekhez

Holl András¹ – Bilicsi Erika²

A szerzők azonosítása – vagy éppenséggel eltitkolása – a könyv- és folyóirat-kiadás alapvető kérdése, amely továbbgyűrűzik például a szerzői jogkezelő szervezetek és a könyvtárak felé. A jogdíjak számontartása fontos a jogkezelőnek, az álneveken vagy név nélkül közölt művek szerzőjének azonosítása az irodalomtudománynak, az egyértelmű szerzőazonosítás a tudományos teljesítmény mérése szempontjából a kutatást végző vagy támogató szervezeteknek, a szerző szerinti keresés lehetősége minden olvasónak.

Mivel a könyvtárak hivatottak a dokumentumok megőrzésére, feltárására és szolgáltatására, ehhez olyan nyilvántartást vezetnek, melyben pl. szerző szerinti keresés is lehetséges. Minden bizonnyal a könyvtár volt az első intézmény, ahol felmerült a szerzők azonosításának kérdése, bizonyítja ezt, hogy az ókor első rendezett könyvtáraként számon tartott ninivei könyvtár katalógusa tekinthető a világ első könyvtári katalógusának, mely szerzői és szakrendi részből állt (a könyvtárat Asúrbanipal i.e. 668-626 közötti uralkodása idején alapította). Kezdetben maguk a szerzők sem fordítottak figyelmet arra, hogy egyértelműen fel legyen tüntetve a nevük a közleményeiken, ezen a könyvnyomtatás feltalálása változtatott. Miután a kézi másoláshoz viszonyítva gyorsan, nagy példányszámban jelentek meg a kiadványok, kialakult a maihoz hasonló kinézetű és funkciójú címoldal, amin a szerző, cím, nyomda/kiadó, megjelenés dátuma adatok igen hamar „kötelező” elemekké váltak. Egyre nőtt a szerzők száma, egyre nagyobb feladatot jelentett a nagy gyűjteményekben a szerző szerinti nyilvántartás, ill. a visszakeresés biztosítása. De nem csak a nyilvántartás miatt vált fontossá az azonos névalakkal rendelkező szerzők megkülönböztetése. Maga a tudomány is igényt formált arra, hogy a tudósokról nyilvántartások készüljenek, ennek eredményeként születtek meg az egy-egy tudományterület neves képviselőinek rövid életrajzát és bibliográfiáját tartalmazó lexikonok.

Az informatika forradalmasította a könyvtári nyilvántartást, egyre több és több adat tárolását tette lehetővé, aminek segítségével hihetetlen mennyiségű, különböző szempont szerinti visszakeresés valósítható meg, arról nem is beszélve, hogy az egyes könyvtárak adatbázisai összekapcsolhatóak, így névterek hozhatóak létre (és itt természetesen nem csak a személynév besorolási adatokról van szó).¹

Az egységes nyilvántartást és így az összekapcsolhatóságot a szabványok hivatottak biztosítani. A szerző azonosítás tekintetében mérföldkövet jelentett, mikor 2009-ben bekerült a MARC21 szabványba a 046-os mező, melyben szerzők esetében a születés és a halálozás dátuma vehető fel az authority rekordba, valamint a 37X adatmezők, melyek pl. a személynevekhez kapcsolódó pl. születési hely, foglalkozás, munkahely adatok felvételét biztosítják. Ezek kitöltésével részletes személyi nyilvántartást vezethetnek a könyvtárak, aminek eredményeképpen létrehozható nemzetközi névtér is.

1 <http://orcid.org/0000-0002-6873-3425>

2 <http://orcid.org/0000-0003-2651-2955>

024 7# \$a 0000000123456789	372 ## \$a középiskolai tanár
\$2 viaf	\$s 1959
046 ## \$f 19300101	\$t 1989
\$g 20001231	373 ## \$a Érdi Városi Gimnázium
100 ## \$a Minta Márta	\$s 1959
370 ## \$a Nagytétény	\$t 1989
\$b Budapest	374 ## \$a történelemtanár
371 ## \$a Barackos utca 2.	\$s 1959
\$b Nagytétény	\$t 1989
\$d Hungary	372 ## \$a egyetemi tanár
\$e 1225	\$s 1990
\$s 1930	\$t 2000
\$t 1963	373 ## \$a Budafoki Egyetem
371 ## \$a Szilvafa utca 1.	\$s 1990
\$b Budapest	\$t 2000
\$d Hungary	374 ## \$a tanársegéd
\$e 1117	\$s 1990
\$s 1964	\$t 1993
\$m minta.marta@mail.com	374 ## \$a adjunktus
	\$s 1994
	\$t 2001
	375 ## \$a nő

1. kép

MARC21 authority rekord

Ilyen nemzetközi kezdeményezés a VIAF Konzorcium (VIAF - Virtual International Authority File), ami a United States Library of Congress, a Deutsche Nationalbibliothek és az OCLC 1998-as kezdeményezését követően 2003-ban alakult meg és mára 46 tagszervezet adatait szolgáltatja (ez még több könyvtárat jelent, hiszen pl. nem az olasz nemzeti könyvtár tagja a konzorciumnak, hanem a Central Institute for the Union Catalogue of the Italian Libraries, de együttműködik a VIAF-al pl. az ISNI és a Wikidata is). A VIAF célja, hogy a nemzeti besorolási adatokat egy közös keresőfelületen tegye elérhetővé – így az egységesített szerzői nevek mellett pl. földrajzi nevek, testületi nevek is kereshetők benne. Az adatokat a központi rendszerben összefésülik - ez hivatott biztosítani, hogy minden besorolási adat csak egyszer szerepelhessen –, az így létrejött rekord egyedi azonosítót kap, melynek alapján hivatkozni lehet rá.² Azonban meg kell jegyezzük, hogy az adatokat nem egy központi rendszerben tárolja az OCLC, csak egy felületen szolgáltatja. Azaz ha az egyes könyvtárak javítanak a besorolási adatokon vagy új könyvtár csatlakozik a konzorciumhoz, az OAI protokoll használatával a VIAF-ban elérhető adatok is módosulnak. Ez azt jelenti, hogy bár egyedi azonosítót kapnak a központi rendszerben összefésült rekordok, azok adattartalma nem állandó.

Szintén 2009-ben került be a MARC21 szabványba a 024-es mező, mely a besorolási adatok nemzetközi szervezet által szabványosított azonosítójának felvételét teszi lehetővé, azaz pl. a VIAF azonosító ebben a mezőben feltüntethető. Ezáltal pedig eljuthatunk a szerző VIAF-ban tárolt munkásságához is.

Manapság, az interneten egyre szaporodó adatbázisok világában már nem csak a könyvtárosok problémája a szerzők egyértelmű azonosítása. A citizen science felé hajló kezdeményezés a nevpont.hu.³ 2014 óta működik, alapvetően lexikonok anyagára épít, és a hozzá kapcsolódó Facebook oldal lehetőséget adhat arra, hogy a közösség is hozzájáruljon az adatok kiegészítéséhez.

Az azonosítók történetének következő állomását az ellenőrző funkció bevonása jelentette, mégpedig olyan ügynökségek személyében, melyek egy-egy terület szerzőinek adataiért felelősek, tekintettel arra, hogy a szerzők adatai a birtokukban vannak.

Ilyen azonosító az International Standard Name Identifier (ISNI), mely a kreatív művek – nem kizárólag írott anyagok – szerzőinek azonosítására szolgáló megoldás. Az ISNI azonosító háttérében az ISO 27729:2012 szabvány áll, a VIAF-hoz hasonlóan, szintén az OCLC üzemelteti. Az ISNI – az ISBN vagy az ISSN azonosítók kiadásának mintájára – helyi regisztrációs ügynökségek segítségével működik.⁴ A beküldött adatoknak meg kell felelni egy minimális minőségi követelménynek, ami az bekerülő adatok minősége mellett a duplumok elkerülését hivatott biztosítani.⁵ A rendszerben tárolt adatok nem közvetlenül hozzáférhetőek, a nyilvános keresőben csak egy szűkített adattartalom érhető el, ld. <http://isni.org/isni/0000000109273206>.

Számos adatbázis üzemeltető a saját rendszerén, termékein belül használ saját egyedi azonosítót a szerzők azonosítására, mint pl. a Scopus Author ID, amit az Elsevier használ adatbázisában, de ilyet alkalmaz a Microsoft Academic Search is. Ezek előnye, hogy segítségükkel többé-kevésbé megkülönböztethetőek az azonos névalakú szerzők, így egy-egy kutató munkássága megjeleníthető az adatbázisokban, ehhez kapcsolódóan pedig tudományometriai adatok, statisztikák kérdezhetőek le az adott adatbázisban tárolt közleményekre vonatkozóan. Azonban a szerzők azonosítása nem megbízható. Kifejezetten problémás a magyar nevek adminisztrációja a keleti névsorrend és az ékezetek miatt, a hibák javítása pedig gyakorlatilag kivitelezhetetlen, hiszen olyan nagy mennyiségű adattal dolgoznak ezeknek az adatbázisoknak az üzemeltetői, hogy az egyedi javítási kérések teljesítése csak ritkán valósul meg – bár a hibabejelentésre kialakított felülettel is rendelkeznek. Sokkal pontosabb eredményt értek el azok az adatbázisok, melyek engedik a szerzők közreműködését vagy arra alapozzák működésüket, mint pl. a Google Scholar author ID, az amerikai Clarivate Analytics által használt ResearcherID a Web of Science adatbázisban és pl. az EndNote-ban vagy a hazai adatbázisok közül pl. a Magyar Tudományos Művek Tára (MTMT).

A tudományos művek szerzőinek azonosítása sok szempontból az említett törekvésekkel átfedő követelményeket támaszt, mégis elég speciális terület, ahol érdemes volt egy külön kezdeményezést létrehozni – ez az ORCID (Open Researcher and Contributor ID).

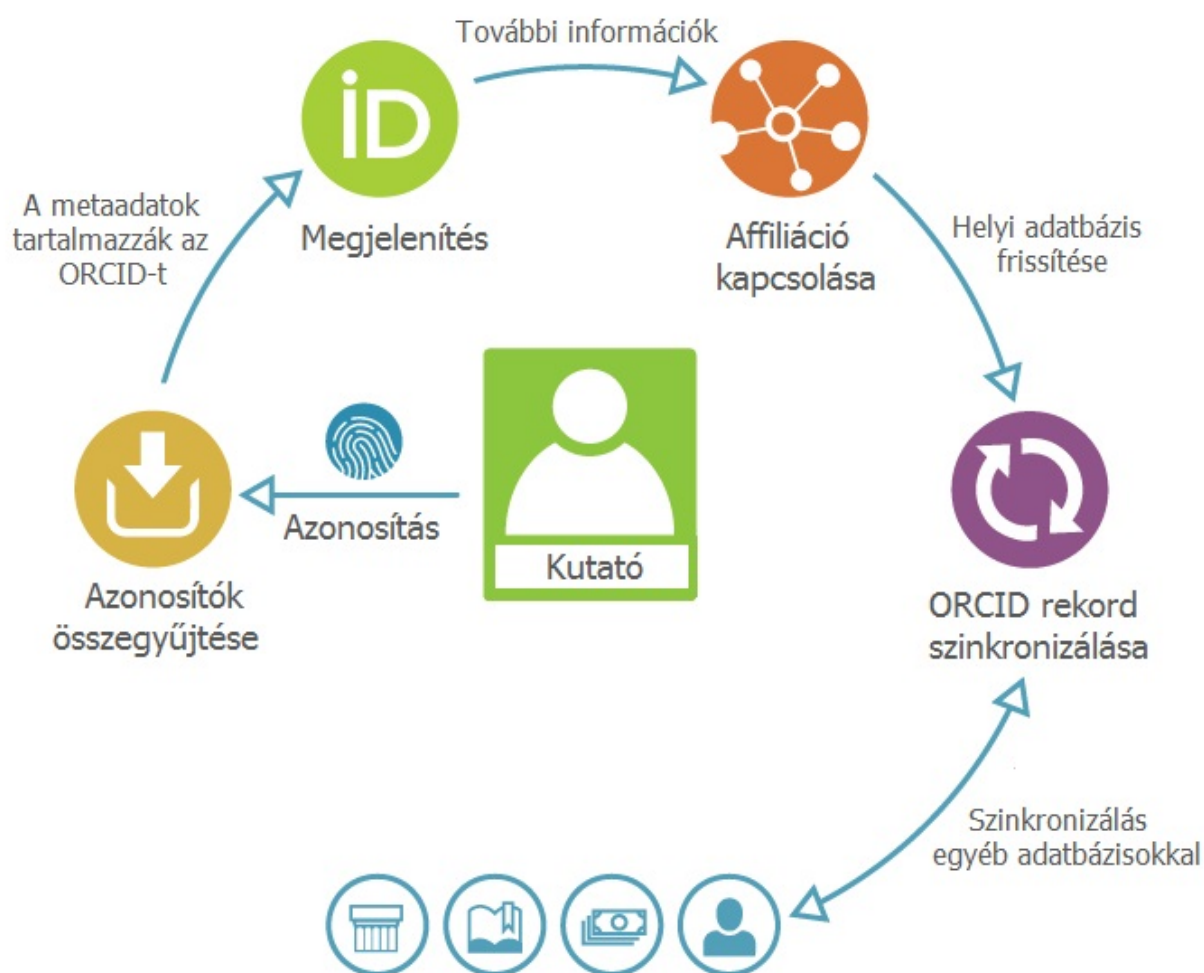
Az ORCID 2009-ben alapított szervezet, kiadók, kutatást végző intézmények hozták létre, de együttműködik pl. az ISNI International Authority szervezettel is, azonosítója az ISO szabványnak megfelel. Nem csupán azonosítók kiosztásával foglalkozik, az azonosítóval rendelkező kutatók tudományos profilját is megjeleníti, továbbá infrastruktúrát biztosít a tudományos művek szerzőségi információinak hiteles és pontos áramlásához a tudományos kommunikációban érdekelt felek között.⁶ Non-profit alapon működik, a tagdíjakból fedezi a költségeit. A tagok között megtalálhatóak kiadók a legnagyobbaktól a feltörekvőig (mint az Elsevier, Springer Nature, Wiley, Taylor & Francis, az eLife és a PLoS), tudományos társaságok (mint az American Association for the Advancement of Science, az American Astronomical Society vagy a Royal Society of Chemistry), egyetemek (köztük a Harvard, az MIT, valamint az University of Oxford) és nemzetközi kutatási szervezetek (mint a CERN). 2017 nyarán a tagintézmények száma megközelítette a 700-at, a regisztrált kutatóké pedig elérte a 3,5 milliót.

Az ORCID szolgáltatásai nyílt forráskódú szoftvereken alapulnak, elterjedten használt, szabad szoftver komponensekből építkeznek (PostgreSQL adatbázis-kezelő, Solr keresőmotor). A szabványos alkalmazásprogramozási felület (API) RESTful webszolgáltatásokat használ XML és JSON kimenettel.

Mint már említettük, az ORCID azonosítók egyben ISNI azonosítók is, az ISNI azonosító-tartományából egy elkülönített részt használnak fel. Egy azonosító – mint az ISNI azonosítók általában – 16 számjegyből áll, négy számjegyenként kötőjellel elválasztva jelenik meg (az utolsó

pozícióban nem csak számjegy, hanem egy „X” betű is szerepelhet). Az ORCID azonosító „actionable” - azaz URL-é alakítható, ami a kutató adatlapjára mutat.⁷

Több, mint 3000 folyóirat gyűjti be szerzőitől az ORCID azonosítókat jelenleg, de egyes kiadók az azonosító használatát meg is követelik. Néhány kiadó nyílt levélben⁸ felhívást adott közre az ORCID bevezetéséről – a levélhez folyamatosan csatlakoznak új aláírók. A levél végén található táblázat tartalmazza az aláírók nevét, és az időpontot, amitől kezdve kötelező az ORCID megadása. A nyílt levél szerint nem csupán elő kell írni az ORCID használatát, de a kiadónak az ORCID API-t használva kell ezeket begyűjtenie; használni kell a CrossRef „auto-update” funkcióját (a cikk DOI-jának regisztrálásakor a CrossRef automatikusan frissíti az új közleménnyel a szerző publikációs listáját az ORCID adatbázisban); és végül az ORCID azonosítókat meg kell jeleníteni a cikkekben – a metaadatokban és a szövegben is (legalábbis az elektronikus változatban).



2. kép
Az ORCID működése

A kutatók számára ingyenes az ORCID azonosító létrehozása és a profil fenntartása, ha ezt egyénileg teszik meg. Az egyetemek, kutatóintézetek beléphetnek az ORCID-ba, és elláthatják kutatóikat, oktatóikat azonosítókkal – de megtehetik azt is, hogy kötelezik őket arra, hogy egyénileg

regisztráljanak. Ez utóbbi megoldás az intézménynek nem kerül pénzbe, a tagság viszont jobb hozzáférést biztosít az ORCID adatbázisához, valamint lehetővé teszi azonosítók készítését az API-n keresztül. A legolcsóbb intézményi tagsági kategória éves díja 5000 USD. A hazai egyetemek közül tudomásunk szerint az ELTE⁹ és a Miskolci Egyetem rendelkezett az ORCID azonosítók használatáról.

Egy ORCID azonosító létrehozása mindössze néhány percbe kerül. Amennyiben a kutató nem biztos benne, nincs-e már azonosítója, érdemes saját nevére rákeresnie – ékezetes és ékezet nélküli változatban is próbálkozva. Ha nincs találat, a regisztrációs űrlapon nevét, e-mail címét kell megadnia (és egy jelszót választania). Itt kiválaszthatja, hogy általánosságban az adatait mindenkivel meg kívánja-e osztani, esetleg csupán egyesekkel, vagy senkivel sem. Ezt a választást minden egyes adat esetében külön-külön felül lehet majd bírálni. Ugyancsak meg kell adni, milyen gyakran kívánunk leveleket kapni az ORCID-től (a „soha” is lehetséges választás). A sikeres regisztráció után általános adatokat (foglalkozás, stb.), valamint tanulmányokat és munkahelyeket, elnyert pályázatokat és publikációs listát lehet létrehozni. Az utóbbinál a kézi felvitel mellett importálni is lehet különböző adatbázisokból, például a CrossRef-ből, a Scopus-ból vagy a kutató esetleges korábbi ResearcherID profiljából. Ez egyúttal azt is jelenti, hogy az ORCID gyűjti a szerző egyéb, bibliográfiai adatbázisokban használatos azonosítóit. Ezt teszi számos bibliográfiai adatbázis is, mivel az egyes szerző azonosításra létrehozott adatbázisok között egyre nagyobb az átfedés, de még nem váltható ki egyik a másikkal.¹⁰

Az ORCID infrastruktúrája lehetővé teszi a szerzőségi adatok automatikus áramlását, hitelesen, a forrás feltüntetésével. A folyóirat a kéziratbeküldésnél az ORCID-on keresztül történő azonosítást kér: a szerző előtt felbukkan egy ablak, ahol az ORCID-os jelszavával engedélyezheti az azonosító átadását a kiadónak. Így biztosított, hogy a beküldő szerző azonosítója helyesen kerül a rendszerbe, és valóban ő végezte a feltöltést. A cikk megjelenésekor a szerzők azonosítóit átkerülnek a CrossRef adatbázisába (a legtöbb folyóirat CrossRef DOI-kat használ). A CrossRef viszont az auto-update szolgáltatásával a cikk megjelenésekor frissítheti a szerző ORCID-os publikációs listáját (a szerzőnek lehetősége van ezt esetenként vagy általánosságban engedélyeznie). A szerzői publikációs listát viszont az ORCID API használatával további bibliográfiai adatbázisokkal lesz majd lehetséges szinkronizálni - például az MTMT-vel.

Az ORCID azonosítók használata már beépült egyes szabad szoftverekbe, mint az EPrints és a DSpace repozitóriumi szoftverek, és az Open Journal Systems folyóirat szerkesztőségi és publikáló szoftver. Felhő-alapú kollaboratív kézíratszerkesztőt ajánl szerzőinek az EDP Sciences: a *Writing studio*-t. Ebbe a rendszerbe is be lehet lépni az ORCID API autentikációs szolgáltatásán keresztül. Megjelentek az ORCID használatával kapcsolatos feltételek egyes támogatások ajánlásai között is. Az Európai Bizottság OpenAIRE FP7 Post-Grant Open Access Pilot alternatív támogatási mechanizmusára olyan Open Access folyóiratok pályázhattak, amelyek nem számolnak fel közlési díjat. A pályázatok elbírálásánál előnyben részesített vállalatok között szerepelt az ORCID bevezetése is.

2017-ben több ORCID témájú rendezvény volt Magyarországon. Május 25-én az MTA Könyvtár és Információs Központban tartottak előadást a szervezet vezetői.³ A THOR (Technical and Human Infrastructure for Open Research) szintén az Európai Bizottság támogatásával, és az ORCID részvételével létrehozott projekt. Célkitűzése, hogy a kutatási folyamat során egyedi azonosítók használatával akadálymentesítsék a tudományos kommunikációt. A THOR projekt 2017 szeptember végén kétnapos rendezvényt szervezett Magyarországon.¹¹

Számos egyetemi könyvtár illesztette be a hallgatóknak nyújtott szolgáltatásai közé ORCID azonosítók készítését. A Texas A&M University ezen túl integrálta az azonosítók használatát a

kampusz különböző információs rendszereibe (Clement, 2014). Az első ORCID-dal megjelölt publikáció a hallgatók disszertációja lehet. Mint láthatjuk, az ORCID megoldást jelent az ezután megjelenő tudományos művek szerzőinek azonosítására. Az elhunyt szerzők azonosítása, úgy tűnik, továbbra is gondos nyomozómunkát igényel majd.

1 Bővebben ld. UNGVÁRY Rudolf: A besorolási adatsere-formátum bővülése. A legutóbbi két évtized fejleménye. = Tudományos és Műszaki Tájékoztatás. 58. évf. 2011. 9. sz. 371-386.
p.http://epa.oszk.hu/03000/03071/00047/pdf/EPA03071_tmt_2011_09_371-386.pdf

2 VIAF permalink pl. <http://viaf.org/viaf/95251869>. Azonban a résztvevő könyvtárak rekordjainak minősége, részletessége igen változatos, így a megfeleltetés nem maradéktalanul sikeres, pl. Arany János legteljesebb adatait tartalmazó VIAF rekord a 95251869 számú, de megtaláljuk a 285393883, 89023145, 246767100 azonosítókon is. Az adattisztításban a VIAF munkatársai együttműködnek az ISNI munkatársaival, az ISNI azonosítót a VIAF rekordok tartalmazzák, ahogy az ISNI rekordok is a VIAF azonosítókat.

3 <http://www.nevpont.hu>

4 <http://www.isni.org/content/isni-registration-agencies>

5 ISNI Data Quality Policy: <http://www.isni.org/content/data-quality-policy>

6 CLEMENT, Gail: Establishing Your Scholarly/Professional Identity with ORCID: A Care and Feeding Manual for Graduate Students. Instructional booklet for Texas A&M Graduate Students. 2014. <http://hdl.handle.net/1969.1/151483>

7 A szerző ORCID azonosítójából képzett URL: <http://orcid.org/0000-0002-6873-3425>

8 Requiring ORCID in Publication Workflows: Open Letter: <https://orcid.org/content/requiring-orcid-publication-workflows-open-letter>

9 5/2017. (V. 2.) számú rektori utasítás 4. paragrafusa: https://www.elte.hu/file/ru_2017_05.pdf

10 A MATARKA 2016-ban vezette be a VIAF, az ORCID, az MTMT és a nevpont.hu azonosítókat, de a szerzők azonosítása a mai napig is tart, mivel a szerzők száma a 320 ezret is meghaladja. Bővebben ld. BURMEISTER Erzsébet: Szerzők nevének egységesítése, szerzők szétválasztása, egyedi azonosítók. = Tudományos és Műszaki Tájékoztatás. 63. évf. 2016. 7-8. sz. 244-250. p. <https://tmt.omikk.bme.hu/tmt/article/viewFile/77/78>Az MTMT-ben jelenleg megadható nemzetközi szerző azonosító az ORCID, a GoogleScholar ID, a Researcher ID, a Scopus Author ID, de ezeken kívül számos hazai azonosító is megadható, mint pl. az ODT ID vagy a Neptun azonosító.

11 Az előadások elérhetők a rendezvény weblapjáról: <https://orcid.org/content/2017-hungary-workshop>