

Bemutatkozik az ISNI, a nemzetközi szabványos névazonosító

A személyek, mindenekelőtt a szerzők egyedi azonosítása a könyvtári (és egyéb hasonló célú) rendszerek, szolgáltatások létrehozása során kardinális kérdés. A könyvtári világban már ismert és elterjedt VIAF (The Virtual International Authority File = Virtuális Besorolási Állomány) mellett egy nemzetközi szabványos megoldás létezik: a nemzetközi szabványos névazonosító (ISO 27729:2012 Information and documentation – International standard name identifier (ISNI)). Egy országos könyvtári rendszer (illetve szolgáltatás) fejlesztésekor kézenfekvően felmerül egy ilyen nemzetközi azonosító bevezetése, ami magában foglalja egy nemzeti ügynökség felállítását.

Tárgyszavak: tulajdonnév; azonosítás; könyvtárügy; szabvány; katalogizálás

A VIAF-tól az ISNI-ig

Az IFLA berkeiben már 1978-ban felvetődött egy közös, nemzetközi besorolási állomány létrehozásának az ötlete, a megvalósítás azonban végül csak 2003-ban indult el három jelentős intézmény: a Német Nemzeti Könyvtár (*Die Deutsche Bibliothek*), a Kongresszusi Könyvtár (*Library of Congress*) és az OCLC (Online Computer Library Center) együttműködésének köszönhetően.¹ Az OCLC közelmúltban készített felmérése² szerint a kapcsolt adatok (Linked Data) vonatkozásában a VIAF jelenleg a leggyakrabban használt adatforrás, jóllehet korábban elsősorban a katalogizálás segédeszközeként tekintettünk rá. Az új szerepből következnek az új elvárások, amelyek értelmében a VIAF-nak – a kapcsolt adatok szolgáltatójaként – hiteles és megbízható adatokat kell tartalmaznia, valamint – könyvtárosi szemszögből – alkalmassá kell válnia az FRBR-izálás folyamatainak támogatására. A VIAF folyamatos frissülése, függése a könyvtári forrásadatbázisoktól ugyanakkor veszélyezteti az adatállomány stabilitását, és felveti ellenőrző mechanizmusok beiktatását. Megoldásként kínálkozik a nemzetközi szabványos névazonosító, az ISNI használata. Utóbbi épít a VIAF-ra, ugyanakkor dedikált humán erőforrás (egy ún. „Quality Team”) foglalkozik a karbantartásával, a névazonosságokból fakadó hibák kiküszöbölésével. Az ISNI nemzetközi szabvány³, bevezetése kérdéseket vet fel a munkafolyamataink racionalizálása tekintetében. Alább *Vincent Boulet* és *Héloïse Lecomte* előadása⁴ alapján betekintést nyújtunk a *Francia Nemzeti*

Könyvtár (Bibliothèque nationale de France, BnF) projektjébe, felvázolva, miként járult hozzá a könyvtár ISNI szolgáltatásának újragondolása az adatfolyamatok optimalizálásához.

Az ISNI DEM projekt

A BnF általában vett célja, hogy elmozduljon az adatok újrafelhasználását támogató automatizált eljárások irányába, összehangolva az újszerű technológiák alkalmazásából fakadó előnyöket a nemzeti könyvtári feladatokkal és minőségi elvárásokkal. Ebből következik, hogy nem tekintik célnak a humán erőforrás kiváltását, sokkal inkább emberek által ellenőrzött automatizmusok, azaz fél-automatizmusok implementálására törekednek. A jelenleg folyó projektek mindenekelőtt a nemzeti bibliográfia és a kötelempéldány-szolgáltatás adatfolyamainak megújításával kívánják ezt a célt elérni.

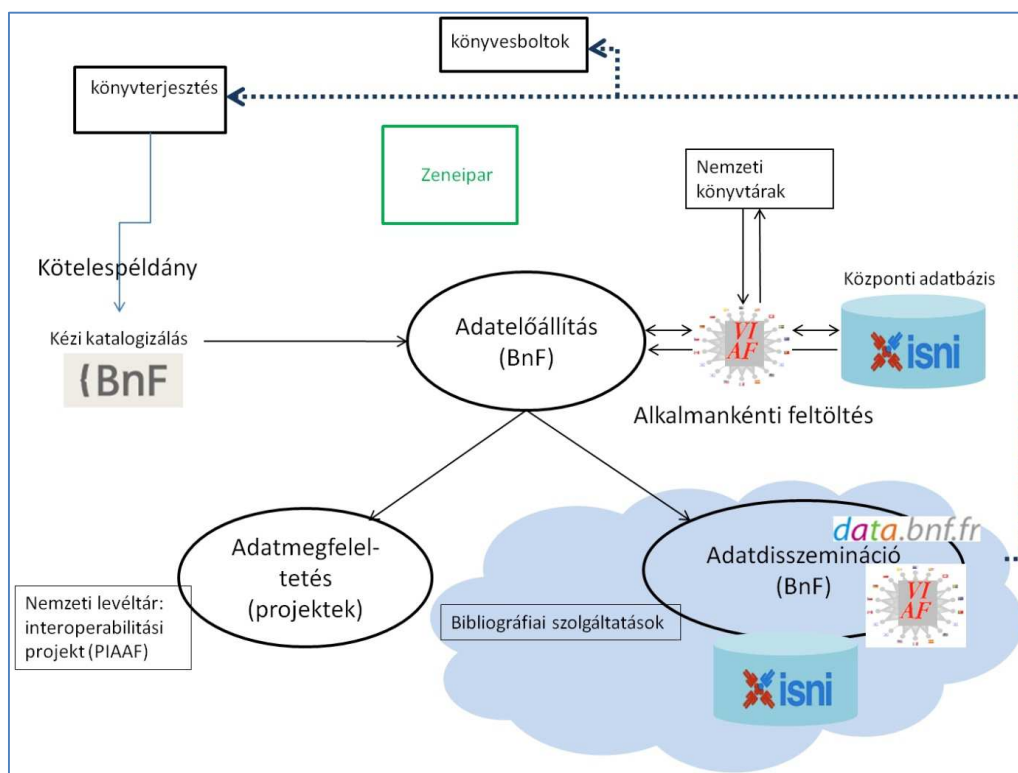
A BnF több ponton is jelen van az ISNI rendszerében: azon felül, hogy részt vesz a nemzetközi ISNI-ügynökség vezetésében, szakértőt delegál az adatbázis minőségéért felelős munkacsoportba, illetve felületet biztosít az ISNI-nyilvántartás, az adatszolgáltatók és végfelhasználók közötti adatkommunikációhoz. A könyvtár emellett ISNI-ügynökségként is működik, és ilyen minőségében törekszik az azonosító nemzeti szintű elterjesztésének kedvező technológiai környezet kialakítására. E téren együttműködik különböző kutatási létesítményekkel, kulturális intézményekkel, levéltárakkal, kiadókkal stb. A terjesztést szolgálja emel-

lett, hogy nemcsak katalógusaiban, szemantikus szolgáltatásában (data.bnf.fr), de a külső adatbázisokkal való kommunikációban is használja az azonosítót, amelynek egyik felhasználási területe a – jelenleg még csak tervezett – nemzeti besorolási állomány (Fichier national d'entités).

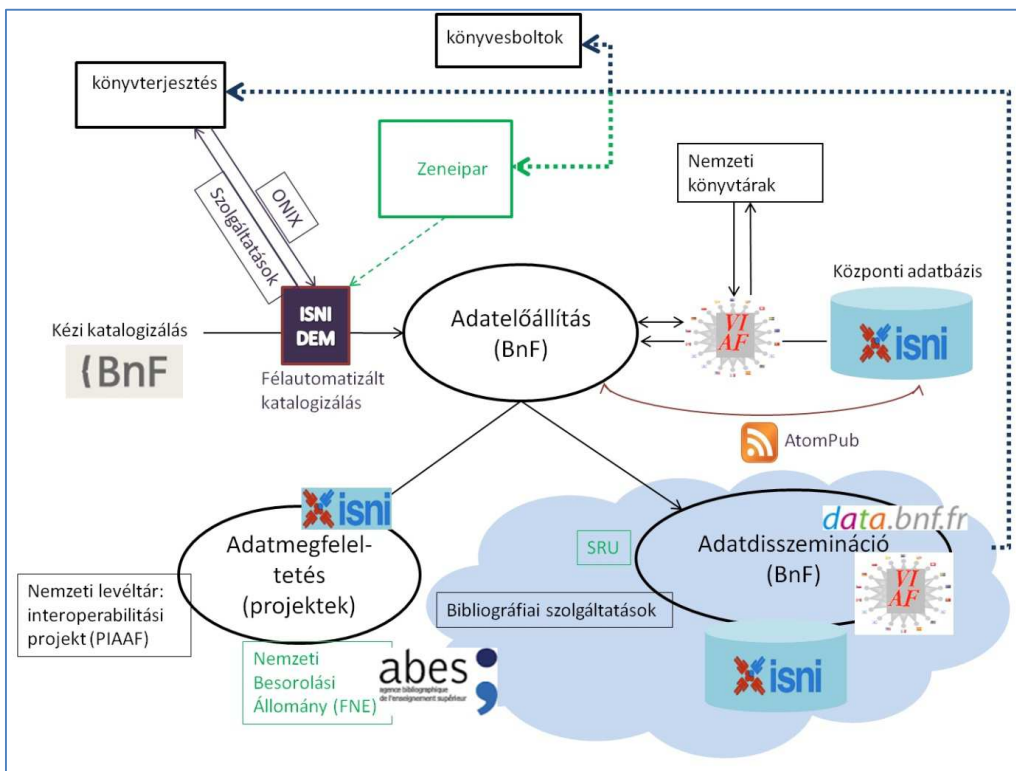
Az alábbi két ábra rávilágít a félautomatizálás irányába való elmozdulás és az ISNI-szolgáltatás kiterjesztése közötti összefüggésre (1., 2. ábra).

Az első ábra a projekt előtti állapotokat mutatja. Jól látható, hogy a könyvtáron kívüli adatfolyamok becsatornázása az ISNI-számigénylés megvalósításával történik. Fontos szerepet játszik ezen a ponton a könyvterjesztésben alkalmazott ONIX szabvány.⁵ A félautomatizált katalogizálás azt jelenti, hogy a bibliográfiai leírásokat készítő munkatársak építhetnek a könyvpiari értéklánc egyéb komponenseiből (pl. kiadók, könyvterjesztők) érkező rekordokra. A munkafolyamat változását az alábbi két ábra illusztrálja (3., 4. ábra).

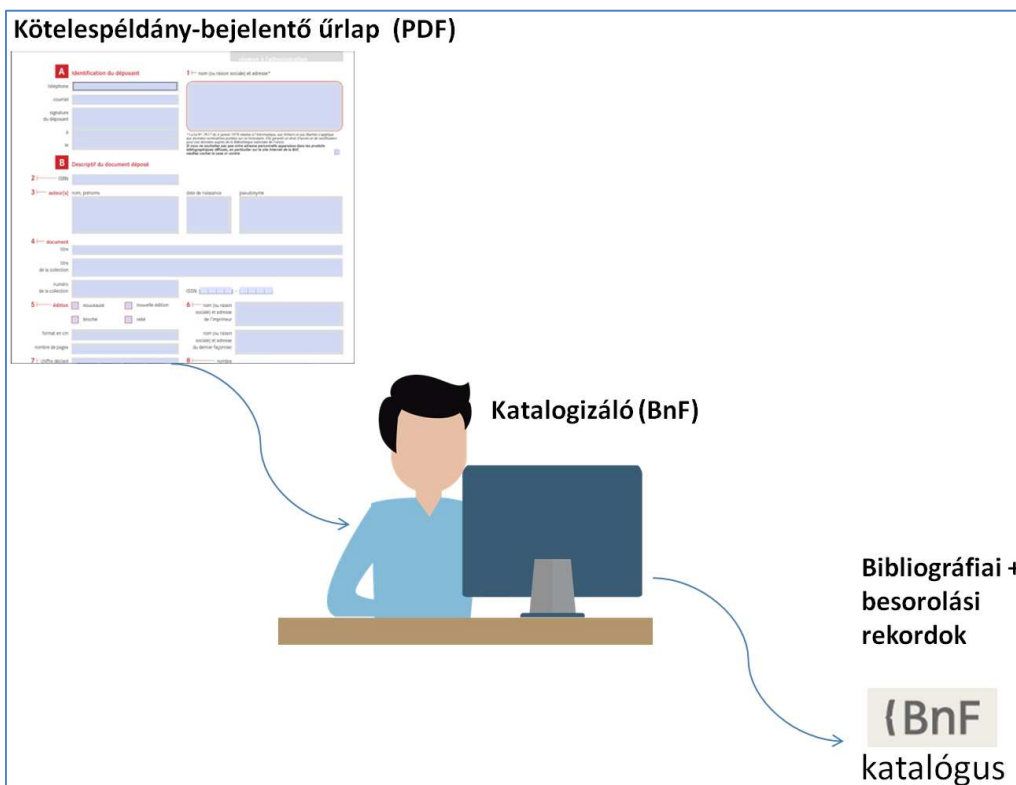
Aktuálisan a kötelezpéldányként beérkezett mintegy – 65 ezer szerzőtől származó – 75 ezer könyv 90 százalékának a feldolgozása az első ábrán vázolt munkafolyamat szerint zajlik. A kiadónak bejelentést kell tennie egy elektronikus űrlap kitöltésével a beküldött kiadványról. A kinyomtatott és a kiadvánnyal együtt kézbesített űrlap segíti a dokumentum azonosítását. A piac 10 százalékát lefedő két kiadói csoport viszont már közvetlenül küldi be a kiadványok adatait ONIX-formátumban az ún. „kötelezpéldány extranet”⁶ applikáción keresztül. Az ONIX-rekordokból szoftveres úton kinyert közreműködői adatokat a BnF munkatársai besorolási rekordokká alakítják, majd az Atom Publishing Protocol (AtomPub)⁷ segítségével feltöltik az ISNI adatbázisába. A kiutalt új névazonosítók 8–10 napon belül kerülnek vissza a kiadókhöz. A BnF és az ISNI adatbázisai közötti adatcsere felgyorsulása a következő két ábrával szemléltethető (5., 6. ábra).



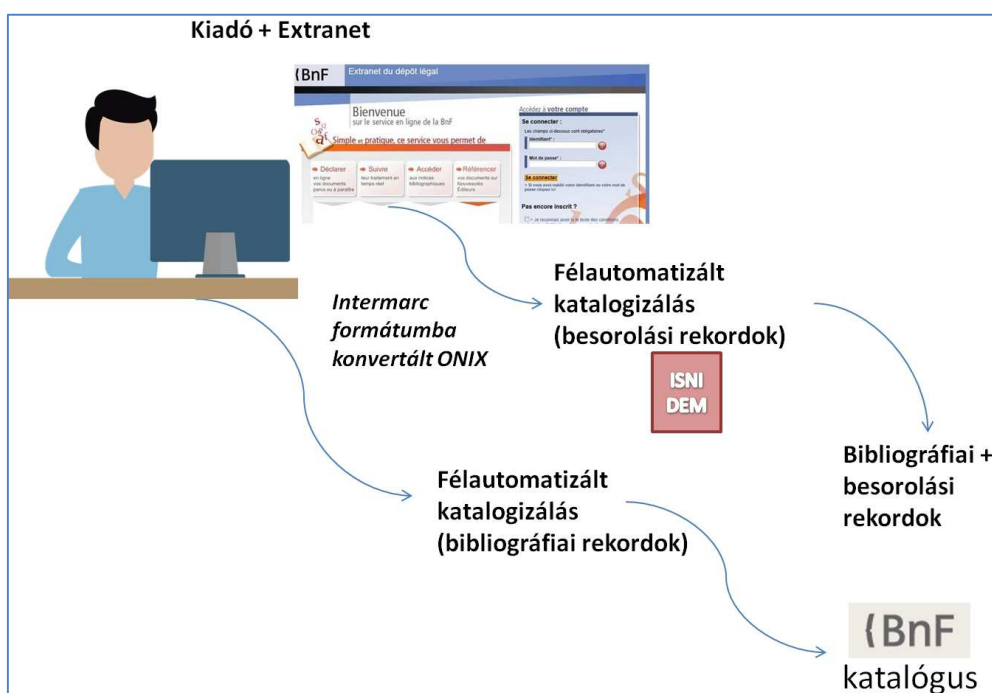
1. ábra Adatfolyamok az optimalizálás előtt



2. ábra Adatfolyamok az optimalizálás után



3. ábra A besorolási és bibliográfiai rekordok létrehozásának hagyományos útja



4. ábra A besorolási és bibliográfiai rekordok létrehozásának félautomatizált útja

Az ISNI-igénylések előállítása tehát a kiadóktól származó ONIX- adatok alapján történik egy új fejlesztésű szoftver, az ISNIDEM segítségével. Az alkalmazás mindenekelőtt ellenőrzi, hogy létezik-e már a szerző a katalógusban. Ha a válasz igen, az igénylést összekapcsolják a vonatkozó besorolási rekorddal (egyelőre manuálisan, de a tervek szerint a jövőben automatizált módon). Ha a szerző még nem szerepel a BnF katalógusában, létrehoznak egy új besorolási rekordot az ISNIDEM rendszerben. Egyelőre ez is manuálisan történik, de ugyancsak tervezik automatizálni a folyamatot. Az igényléseket elküldik az ISNI központi adatbázisába, a kiadókat pedig rögtön értesítik, ahogy az azonosítókat megkapták. A folyamat, amely a kiadóval folytatott párbeszédé teljesebbé válik, legfeljebb egy hetet vesz igénybe (7. ábra).

Egészen konkrétan tehát az történik, hogy az extraneten keresztül lezajló kötelezpéldány-bejelentés pillanatában az ISNIDEM minden közreműködőhöz automatikusan hozzárendel egy számigénylő űrlapot. Az igénylések státusza és prioritása a rendszerben nyomon követhető. A 8. ábrán láthatjuk az automatizált módon kitöltött űrlapot, amely tartalmazza a közreműködőre vonatkozó minden kinyerhető információt, beleértve a bibliográfiai adatokat. Ha a szerző már szerepel a BnF adatbázisában, a rendszer hozzárendeli a megfelelő beso-

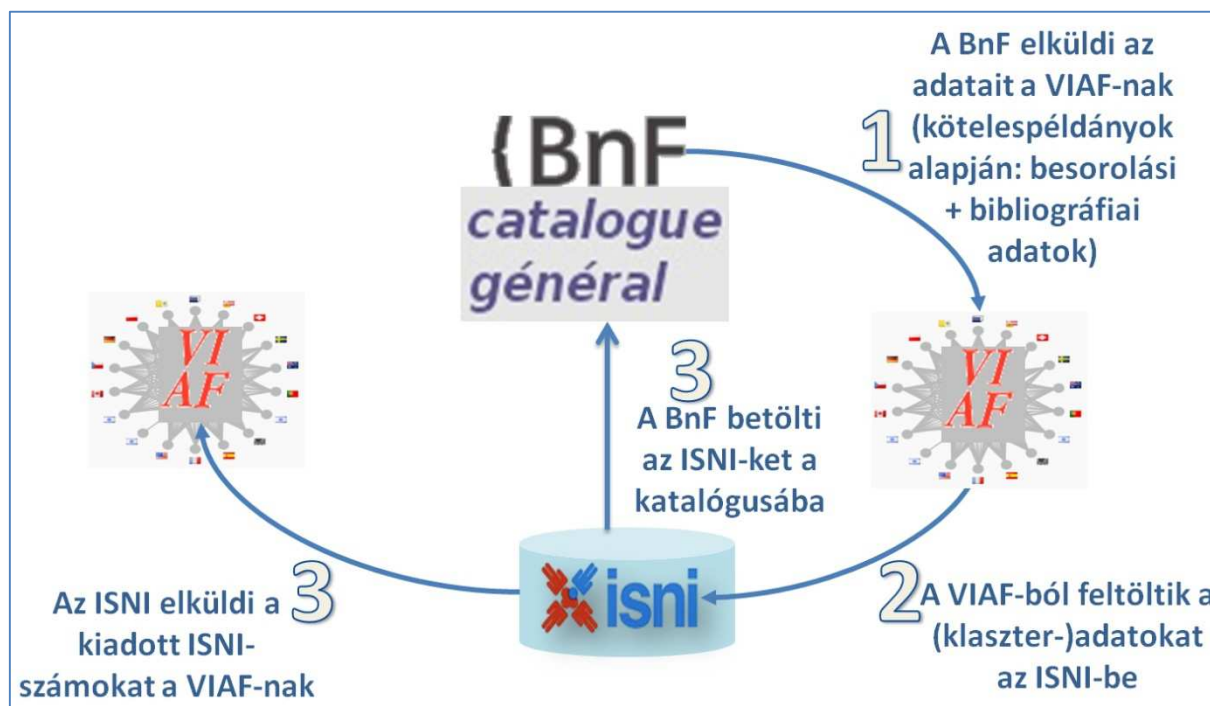
rolási rekord azonosítóját. Azért is fontos a folyamatok minél nagyobb fokú automatizálása, mert a ISNI-ügynökség mindössze két munkatársat foglalkoztat, közülük az egyiket részmunkaidőben.

2016 végén a BnF katalógusa 1,7 millió személynév és 380 ezer testületnév-besorolási rekordot tartalmazott. Összesen 1,4 millió ISNI-t vezettek vissza a rekordokba, ezek eloszlása a személyi nevek és testületi nevek rekordjai között 80:20%. A cél a 100 százalékos lefedettség elérése a személyneveket illetően.

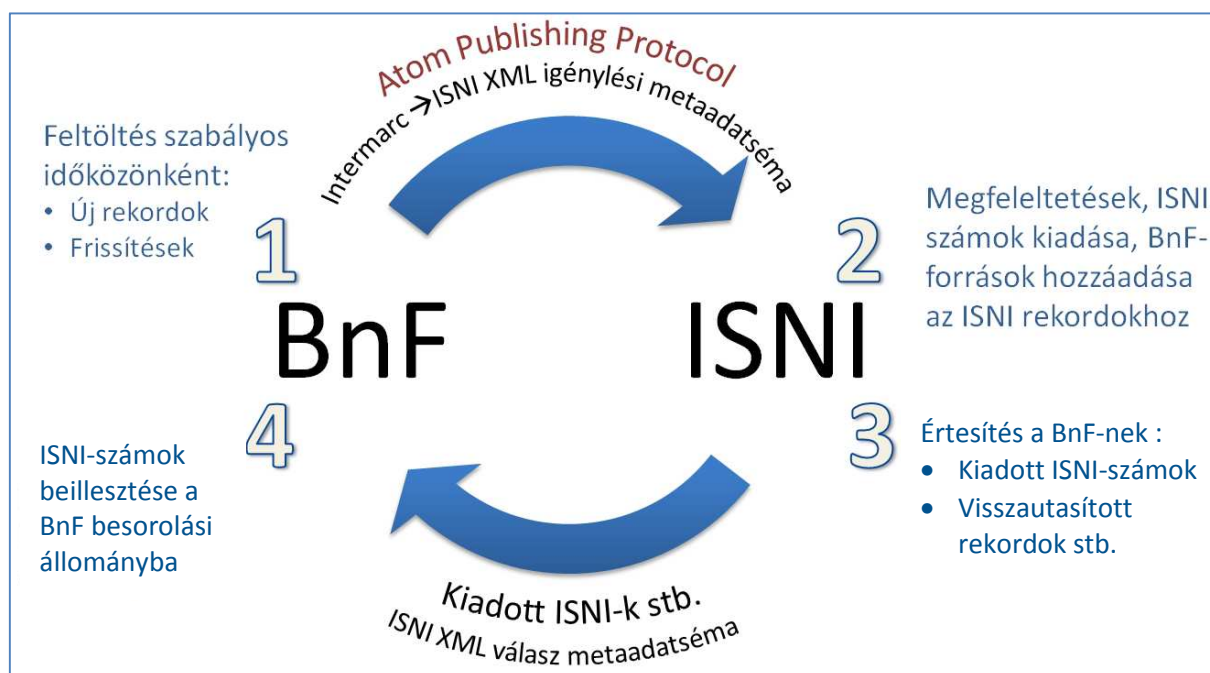
Ahogy említettük, a BnF-nél kiemelt cél az adatok újrahasonosításának⁸ előmozdítása, ami egyben hatékonyan segíti elő az ISNI elterjedését. A katalógusadatok nemcsak bibliográfiai és szemantikus szolgáltatásokban (data.bnf.fr), illetve a VIAF-ban hasznosulnak, hanem különböző kiadói és könyvterjesztői adatbázisok is építenek rájuk (pl. Dilicom).

A 2017-es londoni ISNI Summit során felmerült kérdések

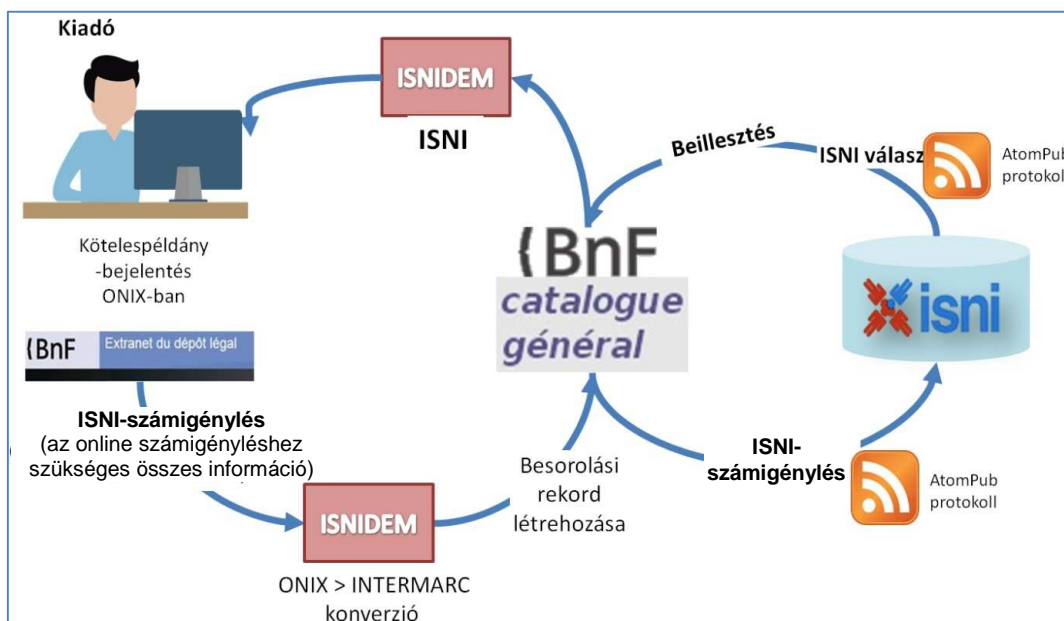
2017 februárjában az ISNI Nemzetközi Ügynökség vezetősége és az OCLC illetékes vezető szakem-



5. ábra BnF-ISNI adatszere (2013–2016)



6. ábra BnF-ISNI közvetlen (automatizált) adatkapcsolat (2016 végétől)



7. ábra Az ISNI-számigénylés kezelése az ISNIDEM rendszer beiktatásával

Liste des demandes Fiche de demande d'ISNI

Action

- Renseigner le NNA pour cette demande
- Annuler la demande
- Passer la demande au statut 'A statuer'
- Créer la notice autorité
- Ajouter un commentaire


Données de gestion de la demande d'ISNI

ID Demande : 813 Statut : ●●●
 Expéditeur : Interforum Priorité : P1 Modifier
 Id expéditeur (DL Extranet) : 2 Date demande : 12/01/2017
 NNA : Date réponse :
 ISNI :

Données auteur (Source : ONIX du demandeur)

ID Local : 128016 Type : PEP
 Nom, prénom : DE BODARD, ALJETTE N° de contributeur ONIX : 1
 Pseudonyme : Fonction : Auteur du texte
 Date de naissance : Langue Source :
 Date de mort : Langue cible :
 Nationalité : Site web :
 EAN du même auteur :

Données bibliographiques (Source : Extranet du dépôt légal)

 Titre principal : La chute de la Maison aux Flèches d'Argent N° de déclaration : 10000000312278
 Sous Titre : Statut de la déclaration : Déclaration validée
 Langue : fre NNB :
 Titre original : Editeur : Fleuve éditions
 Langue Originale : eng Id éditeur (DL Extranet) : 14132
 Date de publication : 12/01/2017 Voir Onix
 EAN : 9782265116337

Données catalogue

- Recherche dans notices complètes SANS date Nb resultat : 1

Nom	Prénom	Date	Qualificatif	ISNI
Bodard	Alette de	1982-...		0000 0000 7806 8831

+ Recherche dans notices élémentaires Nb resultat : 0

8. ábra Automatizált módon kitöltött űrlap

berei tanácskozást tartottak Londonban, ahol a nemzeti ügynökségeként funkcionáló könyvtárak is képviseltették magukat. Az egyik legfontosabb körüljárt probléma a VIAF és az ISNI összefüggése, a fenntarthatóságot biztosító üzleti modell megtalálása volt. Tudvalevő, hogy a VIAF elsődleges felhasználási területe a könyvtárak számára az

adataik láthatóságának növelése, míg az ISNI a kiadókkal, könyvterjesztőkkel és jogkezelő szervezetekkel való együttműködés szempontjából jelent elsősorban perspektívát. A két azonosító a közöttük lévő átfedések ellenére mind műszaki értelemben, mind a finanszírozást illetően különbözik. Az eltéréseket az 1. táblázat⁹ foglalja össze.

1. táblázat

A VIAF és az ISNI közötti eltérések

	VIAF	ISNI
Fenntartás/Működtetés	az OCLC szolgáltatása, amelynek fejlesztését, működését egy a könyvtárak képviselőiből álló tanácsadó testület (VIAF Council) koordinálja	nemzetközi szabvány, fenntartása a Nemzetközi ISNI-Ügynökség hatáskörébe tartozik, technikai háttéréről az OCLC gondoskodik
Finanszírozás/Üzleti modell	ingyenes szolgáltatás	tagsági díj a nemzeti ügynökségek részéről, valamint az egyes azonosítók kiosztása után járó díjak
Lefedettség	személyek, testületek, művek, kifejezési formák és földrajzi nevek	személyek és testületek
Adattartalom	csak nyilvános adatok	személyes (titkosított) adatok is
Az azonosító szerepe	besorolásirekord-csoportok létrehozása (klaszterezés) a nagyobb nemzeti könyvtárak, tudományos könyvtárak és egyéb közgyűjtemények besorolási rekordállományára építve, az adatok közzététele és terjesztése	szabványos nemzetközi (ISO) azonosítók létrehozása építve a VIAF-ból származó releváns adatokra; adatkapcsolatokat generál más (nem-VIAF) forrásokkal – platformként szolgál a cross-domain adatkapcsolódásokhoz; kizárólag a kiutalt ISNI-azonosítókat teszi közzé és terjeszti (tehát azokat, amelyek esetében az egyes rekordokról bizonyossággal elmondható, hogy egymástól különböznek, illetve hiteles információk alapján deduplikálták őket)
Technikai megoldások	aratja az adatokat, csoportosítási (klaszterezési) eljárást alkalmaz, a rekordok megjelenítése és letöltése többféle formátumban lehetséges	kötegelt feltöltést alkalmaz, illetve csoportosítási eljárást, végzi a rekordcsoportok (klaszterek) karbantartását; az online számigénylés API-n keresztül megvalósítható; lehetséges végfelhasználói közreműködés (crowdsourcing) (kontrollált módon), valamint az online szolgáltatások testreszabása a tagok és nyilvántartó ügynökségek számára; szakemberekből álló csapat (Quality Team) biztosítja az adatbázis minőségének ellenőrzését és a hibák javítását; a rekordok megjelenítése többféle formátumban (XML, RDF) lehetséges
Adatforrás	könyvtárak besorolási állománya, a kulturális örökség terület egyéb adatszolgáltatói (múzeumok, levéltárak)	különböző adatszolgáltatók (könyvtárak, jogkezelő szervezetek, kutatói adatbázisok, a zeneipar és könyvpiar szereplői); az ISNI-ügynökségek által feltöltött vagy az online számkiutalás során szolgáltatott adatok
Adatmodell	a VIAF semlegesnek tekinthető az egyes szabványok és formátumok tekintetében	az ISNI ISO szabványként egyedi adatmodell kialakítását igényli, ugyanakkor adatkezelési politikája alkalmassá teszi az eltérő forrásokból származó adatok befogadására

Adatperzisztencia	a klaszterek tartalma a feltöltésekkel gyakran változik	a cél a lehetőségekhez mérten stabil, azaz perzisztens és megbízható azonosítók létrehozása; a rendszer olyan algoritmusokkal rendelkezik, amely csak azoknak a klasztereknek utal ki ISNI-számot, amelyek megbízhatósága elér egy meghatározott szintet
Frissítések	a VIAF havonta frissül	folyamatos frissülés az ügynökségek igényei alapján, illetve az online számkiutalások tükrében, beleértve a Minőségcsoport rekordszerkesztési műveleteit, illetve a végfelhasználói igénylések kezelését
Adatdisszemináció és licenc	a VIAF nyílt kapcsolt adatként terjeszhető ODC-By licenc ¹⁰ alatt	az ISNI egy hivatkozható nyilvántartás, amely egy API-n keresztül nyilvánosan lekérdezhető; az adatok nyílt licenc alatt felhasználhatók, jelenleg tervezik nyílt kapcsolt adattá fejlesztését
Célközönség	könyvtárak, a kulturális örökség intézményei, oktatás és kutatás	globális és több szektort érint (könyvtárak, levéltárak, kutatás, zene- és filmipar, könyv- és folyóirat-kiadás, jogkezelők stb.)
Példa a felhasználásra	a katalógizálás forrása, a könyvtári entitások összekapcsolása repozitóriumokkal (HUB-funkció), mint például a Wikipedia és a Wikidata	linkek automatizált létrehozása a besorolási rekordokhoz a kiadói és könyvtári adatok összekapcsolásával

A VIAF közgyűteményi (elsősorban könyvtári) rekordok alapján és közgyűtemények (elsősorban könyvtárak) által épül, valamint a nevek csoportosítása során nem alkalmaz ellenőrzési mechanizmusokat. Ezzel szemben az ISNI felhasználási köre kiterjedtebb, a teljes kreatív ágazatot magában foglalja, a zeneipart is beleértve, így a benne rejlő adatkapcsolási (linked data) potenciál is jelentősebb. Az ISNI igénylésekor egy szakemberekből álló csapat (ISNI Quality Team) ellenőrzi, hogy az egy azonosítóhoz rendelendő névváltozatok valóban egy személyt takarnak-e, valamint javítják az ISNI referencia-adatbázisban (ISNI Reference Database) előbukkanó hibákat. Az ISNI nyilvántartása tehát megbízhatóbb, hitelesebb adatokra épül. Továbbá az ISNI egy nemzetközi (ISO) szabvány, a VIAF az OCLC tulajdona, jóllehet a kétféle azonosító között van kapcsolat, és az OCLC műszaki partnerként van jelen az ISNI-adatbázis működtetésében.

Az ISNI különös adottsága, ami ugyancsak megkülönbözteti a könyvtári besorolási rekordok nemzetközi szintű azonosítására használt VIAF-tól, hogy különbséget tesz személy és személyiség között, amit a személyes adatok védelmének (privacy) kezelésével valósít meg. Ez közelebbről azt jelenti, hogy lehetőség van arra, hogy egy személy, ha érdeke úgy kívánja, több személyiséget (álnevet) használjon anélkül, hogy a személyiségek, illetve a

hozzájuk kapcsolódó nevek közötti kapcsolat nyilvános volna. Az ISNI névterét használó könyvtári rendszerekbe tehát a kapcsolati információ alpból nem kerül át. A nyilvánosság ugyanakkor – természetesen – a későbbiekben módosítható.

A tanácskozáson megfogalmazódott igény a szolgáltatások fenntartásához szükséges új üzleti modell bevezetésére ugyanakkor több problémát felvet, így például azt, hogy a könyvtárak teljesen eltérő üzleti modellek alapján működnek, amelyeket nehéz összehangolni. Felvetődött ugyancsak a munkafolyamatok standardizálásának kérdése, ezzel kapcsolatban a résztvevők oldaláról megfogalmazódott, hogy az ISNI-k kezelése könyvtáranként másként valósul meg, erősen kérdéses tehát egy mindenkinek megfelelő, standard munkafolyamat kialakítása.

Összefoglalás

A kapcsolt adatok világában a web egyre inkább egy globális adatbázishoz hasonlít, amelyben az adatreprezentáció új eszközeinek köszönhetően egyre több információ válik gépileg értelmezhetővé. Ebben a környezetben az egyedi azonosításnak különös szerep jut, nemcsak a releváns tudományos információk megtalálása, de a tartalmak jogtiszta szolgáltatása miatt is. Korábban afféle disztópiaként tekintettünk arra a világra, amelyben

az individuumok neveit számokra cserélik, az informatika azonban megtanított minket arra, hogy a (besorolási) adatok egységesítésének nincsen jobb módja, mintha a számok nyelvére ültetjük át azokat, ezzel garantálva egyediségüket. A könyvtáros normalizálásra való törekvése a személynevek terén akkor járható tehát csúcsra, ha az entitásokhoz egy (vagy több) egyedi numerikus jelet rendelhet „kitüntetett névalakként”. Az ISNI-t épp erre találták ki.

Jegyzetek

1. BENNETT, R.–HENGEL-DITTRICH, C.– O’NEILL, E.–TILLET, E.: VIAF (Virtual International Authority File): Linking Die Deutsche Bibliothek and Library of Congress Name Authority Files (2006) – <http://archive.ifa.org/IV/ifa72/papers/123-Bennett-en.pdf> (Megtek. 2017. február 16.)
2. SMITH-YOSHIMURA, K.: Analysis of international Linked Data survey for implementers, In: D-Lib Magazine, 22. köt. 7/8. sz. 2016 – <http://www.dlib.org/dlib/july16/smith-yoshimura/07smith-yoshimura.html> (Megtek. 2017. február 16.)
3. ISO 27729:2012 Information and documentation, International standard name identifier (ISNI) – http://www.iso.org/iso/catalogue_detail?csnumber=44292 (Megtek. 2017. február 16.)
4. BOULET, V.–LECOMTE, H.: BnF data production & ISNI : Projects and opportunities, elhangzott a Londonban 2017. február 7-én megrendezésre került ISNI csúcstalálkozón
5. Az ONIX-ról bővebben magyarul: MOHAY Anikó: ONIX (Online Information Exchange), In: Könyv, Könyvtár Könyvtáros, 21. évf. 5. sz. 2012. p. 24-26.
6. Az extranet lényegét a legjobban úgy lehetne megfogalmazni, hogy a szervezet által használt belső hálózati szolgáltatások (intranet) azon része, amely a megfelelő korlátozások mellett hozzáférhető a szervezettel kapcsolatban lévő ügyfelek, jelen esetben mindenképp a kötelempéldány-szolgáltatásra kötelezett kiadók számára.
7. <https://bitworking.org/projects/atom/rfc5023.html> (Megtek. 2017. február 16.)
8. Bővebben Id.: DANCS Sz.: Névtér építése nemzeti könyvtári környezetben és a bibliográfiai metamorfózis, Könyvtári Figyelő, 62(1). 2016. pp. 77-85.
9. A táblázat forrásai: A Global strategy for authority metadata and identifiers (2016) – <http://www.loc.gov/aba/pcc/documents/PoCo-2016/VIAF-ISNI-position-paper.pdf>; ANGJELI, A.–MAC EWAN, A.–BOULET, V.: ISNI and VIAF – Transforming ways of trustfully consolidating identities – <http://library.ifa.org/985/1/086-angjeli-en.pdf> (Megtek. 2017. február 16.)
10. <http://opendatacommons.org/licenses/by/1.0/> (Megtek. 2017. február 16.)

Beérkezett: 2017. II. 17-én.



Dancs Szabolcs
az OSZK projektigazgatója.
E-mail: dancsz@oszk.hu