

Simon András

Rekord vagy háló – tudásrepresentációs eszközök személynévtér-építéshez gépi katalógusokból kinyert adatok alapján

Névtér: új fogalom régi megoldás

Ügyfeleink, a könyvtárak és múzeumok katalógusainak és leltárkönyveinek számítógépes kezelése során újra meg újra szembesülünk azzal, hogy mennyire fontos a besorolási állomány, és különösen a személynévállomány egységesítése, karbantartása. Tapasztalataink azt mutatják, hogy a jelenleg alkalmazott, MARC-rekordokra épülő adatmodell alkalmas egy intézmény névtérrel kapcsolatos igényeinek kiszolgálására. Ugyanakkor viszont érzékeljük, hogy az egyre nagyobb egységekbe szervezett adatbázisok már olyan kihívásokat támaszthatnak a rendszerrel szemben, amelyek a hagyományos rekordalapú adatszerkezet helyett szemantikus tudásrepresentációra való áttérés szükségességét indokolhatják. Az adatbázisok mind magasabb fokú integrációja, illetve a rekordalapú megközelítésről szemantikus hálóra való áttérés számos nehézséget rejt magában, és érdekes kérdéseket vet fel.

Tárgyszavak: névtér; személynév; azonosítás; metaadat; tudásrepresentáció

A névtér fogalma, definiálása

Korunkban a közgyűjteményi és informatikus szakmában már nem kérdés, hogy névterek építésére, fenntartására szükség van. Abban sincs nézetkülönbség, hogy a névterek felépítése és karbantartása hatalmas feladat, és ebben a munkában kulcsszerepe van a közgyűjtemények gépi nyilvántartásainak, a könyvtári számítógépes katalógusoknak és a múzeumi digitalizált leltárkönyveknek. E tanulmány szerzői a névtér meghatározása esetében *Ungvári Rudolf* a TMT „A névtér, mint kulturális szükséglet” c. cikkében felsorolt ismérvek szerint meghatározottan fogják fel: „a magyar nevek (tulajdonnevek és köznevek) rendszerre szervezett, minden közvetlen gyakorlati alkalmazástól független, weben megvalósuló szemantikai rendszere.”¹

A névtér tehát, jóllehet adatait ma (még) adatbázis kezelő rendszerekben tárolják, több mint adatbázis, és természetesen, több mint az alkalmazásokban betöltött szerep, funkció, melynek keretében valahol, valamilyen formában hasznát látjuk. Mindamelltt jelen tanulmány szerzői alapvetően az alkalmazásfejlesztés gyakorlati oldala felől közelítik meg a kérdést. Egy közgyűjteményi informatikai cég alkalmazottaiként a megrendelők számára fejlesztett és szolgáltatott gyűjteménykezelő és

szakirodalmi információs rendszerek kiszolgáló háttereként tekintenek a névterekre, tudatában lévén természetesen annak, hogy a névtér önmagában is érték, nem csak abban a tekintetben, amennyire az egyes informatikai megoldások azt használni tudják.

Miért van szükség névterekre?

Egy korszerű gyűjteménykezelő rendszer már nem elégedhet meg egyszerűen besorolási állományok építésének a lehetőségével, különösképpen nem szótárak biztosításával. Háttér állományaik elemeinek adatbázisszinten, – sőt lehetőleg tágabb szinten – azonosítottak, tartalmukban szegmentáltak, rendszeren belül és rendszeren kívüli kapcsolataikban pedig pontosan meghatározottak kell lenniük. Ebben a tekintetben, ahogyan a besorolási állomány is, több mint egy kötött szójegyzék, és a teaurusz is több mint egy hálózatba rendezett besorolási állomány, a névtér is több mint egy kicsit bonyolultabb és átfogóbb teaurusz. Mind-egyik szint egy lépcsőfok azon a nehéz fejlődési úton, ahol a számítógépre vitt szöveges tudást egyre jobban rendezve és meghatározva az emberi tudás asszociatív logikájához akarjuk közelebb vinni, hogy az egyes szakemberek által megszerzett tudást a lehető legjobban (gyorsabban és pon-

tosabban) visszakereshető módon kívánjuk a többi szakember számára hozzáférhetővé tenni.

Személynévtér, földrajzi névtér, egyéb névterek

A névtereknek több fajtájuk lehetséges, ismerünk személynévtereket, földrajzi névtereket, és más köz illetve tulajdonnevekre vonatkozó névtereket is. Gyakorlati okokból a gyűjteménykezelő rendszerek használói és tervezői, illetve menedzselői egyaránt legtöbbit a személynévterekkel kénytelenek foglalkozni. Ennek oka, hogy egyrészt különösen a könyvtári dokumentumok kezelésénél, besorolásánál a személynévbázisra épülő szerzői nevek megkülönböztetett jelentőséggel bírnak, másrészt, akár a helytörténeti gyűjtemények akár a különböző szakmák, írói, művészi, tudósi munkásságokat feltáró gépi nyilvántartások, adatbázisok rekordszerkezetének fő vázát a személyi adatok adják. Elmondható, hogy a gépi katalógusok adatait bővítő szakértői rendszerek alapvetően a személyi tevékenységére és életpályájának bemutatására vannak megtervezve.

Ugyancsak, a közönség érdeklődése is leggyakrabban a személyek felé fordul, és a személyek felől közelíti meg az adatbázist annak nyilvános felhasználói interfészén keresztül. Igaz ez a helytörténeti gyűjtemények esetében az illető személy elődei, hozzátartozói személye iránti érdeklődésre, de az egyes adatbázisokban szereplő írók, művészek, tudósok életpályája iránti érdeklődésre egyaránt. Emellett, kellő alkalmat kínál az adatbázisok tervezőinek, kezelőinek (pl. migrálóinak), hogy a névtér tovább építésében tapasztalatot szerezzenek. Egyrészt ugyanis legalább alapegysége, (egy élő, vagy már elhunyt embertársunk) egyértelműen meghatározható (ugyanaz a testületi vagy földrajzi nevekről egyáltalán nem mondható el), legalább a családi kapcsolatok mentén kétség nélkül hierarchiába rendezhető (melyben az állat- és növényfajok és esetleg a földrajzi nevek hasonlíthatók csak hozzá), másrészt adatok hallatlan bőségét kínálja ahhoz, hogy egy adatbázisban szereplő névtérelmet az összetett kapcsolatrendszer valamennyi féleségébe beilleszthessük.

Mint szállítók kiemelten személynévtérrel foglalkozunk jó néhány ügyfelünknel, így a *Petőfi Irodalmi Múzeum*, az *Országos Színháztörténeti Múzeum és Intézet*, illetve a *Magyar Művészeti Akadémia*, de a legtöbb megyei könyvtár, illetve megyei hatókörű városi múzeum esetében és természetesen a helytörténeti gyűjtemények esetében is. Cikkünk felvetései mindamellett általánosan érintik a többi

névtérrel és összefüggnek bármilyen közgyűjteményi nyilvántartó rendszerre épülő adatbázis összes többi névtérének kérdésével.

Miért van szükség paradigmaváltásra?

Az emberi gondolkodás, a tudomány történetében újra meg újra előáll az a helyzet, hogy a növekvő információmennyiség bővülése az információmennyiség birtokosait, kezelőit szemléletváltozásra készítetik. Természetesen maga az információ-tömeg eddig is megvolt, sőt a szakma látókörébe is került, nem merült azonban fel, hogy ez az adat-tömeg egyszerre egy közvetlen hálózatba kötött rendszerben, adatelemek és kapcsolataik sokaságában legyen elérhető, kereshető, böngészhető. A változás logikájának megértéséhez mi könyvtárosok gyakran fordulunk történelmi példákhoz, *Gutenberg* találmányához, a könyvnyomtatáshoz, majd az ősnymtatványok korát lezáró újításhoz, az *Aldus Manutius* velencei nyomdászhoz kötött, új formájú, olcsóbb és könnyebben megszerezhető és kezelhető modern könyvekhez, végül pedig *Panizzz*hez, aki talán a leginkább tekinthető a modern katalogizálási gyakorlat megalapítójának. Annyiban jól megvilágítják a mai folyamatokat a könyv- és könyvtártörténetből vett példák, hogy a technikai, ipari újítást mindig kicsit lemaradva követik a módszertani változások. Jelenleg az elektronikus információforrások robbanásszerű megjelenésével és elterjedésével azonban már többről van szó, mint egyszerűen a hordozók megváltozásáról, és a korábbi lexikonokba, enciklopédiákba, szótárakba, és egyéb kézikönyvekbe, továbbá a bibliográfiákba és könyvtári katalógusokba illetve múzeumi leltárkönyvekbe bevitt információk digitalizálásáról, mely korlátlan, egyidejűleg több felhasználó számára lehetséges távoli használatot, és veszteségmentes és szinte költségmentes sokszorosítást tesz lehetővé.

A jelenlegi változások, melyek előzményei egyébként évtizedekre nyúlnak vissza, már magának az emberi tudásnak az írásbeliség által nyújtotta lehetőségeket meghaladó reprezentációját vetítik előre, az adatbázisok, adatbankok rekordtömegeinek irányított, azonosított, minősített, és definiált kapcsolatai révén. A MARC szerkezet ezt egyre bővülő relációs adatkészlettel oldja meg. A MARC szerkezetet leképező könyvtári adatbázisok ezeket a relációkat is tárolják.² Ezzel az adatbázisfejlesztési gyakorlattal és alkalmazásfejlesztéssel úgy tűnik, végre a lassú és ostoba számítógép valóban segítségére lesz a gondolkodásban, és az összefüggések levonásában is az okos és gyors embernek.

Hogyan épülnek fel a névterek?

Névterek sok úton sokféleképpen létrejöhetnek és létre is jöttek. Jelen cikkben nem célunk a többféle névtér létrehozási megoldásának összevetése, elemzése. Mi csupán azzal a helyzettel foglalkozunk, amikor a névterek könyvtári gépi katalógusok adataiból, bibliográfiai és besorolási állományából jönnek létre. Ez a megoldás igen elterjedt, mind nálunk, mind külföldön. Előnye ennek, hogy az adatokat már valaki számítógépre vitte, vagy a kártyaalapú katalógusrekordok retrospektív bevitelével, vagy már közvetlenül az integrált könyvtári rendszerben való katalogizálással bekerültek. Ezeket az adatokat a nyilvános olvasói felületen és a katalógusrekordokat más könyvtári rendszerekbe való átvételekor sokan használják, így jó esély van a szakmai és gépi hibák felszínre bukkolásának. Az állomány a rendszeres használat miatt jól karbantartott, általában a szállító által támogatva könnyen elérhető, emellett a már elterjedt közös katalogizálás miatt már nemzeti szinten is egységes. Mivel a magyar könyvtárak anyagai nagy részben magyar nyelvűek, magyar eredetűek, vagy legalábbis kapcsolatban állnak a magyar kultúrával, ezért a magyar személynévállomány teljes körűen és jól feltártan fellelhető bennük.

A meglévő névterek kapcsán jelen cikkünkben elsősorban a nemzeti könyvtárak adatállományai-val foglalkozunk. A várhatóan egyre nagyobb szerepet kapó VIAF (*Virtual International Authority File*), ezek közös reprezentációja.

Mink van most? Besorolási rekordok tömege, metaadatok metaadatai

Párhuzamos rögzítés – amitől az információ-áradat még áradóbb

A különféle adatbankok, adatbázisok építése és a papíralapon létrejött adattartalmak (könyvtári katalógusok, múzeumi leltárkönyvek, levéltári nyilvántartások, bibliográfiai kiadványok, lexikonok, adattárak) számítógépre vitele már jó néhány évtizede zajlik. Ezek a munkák jóval azelőtt elindultak, ráadásul minden országban és nyelvterületen külön-külön, mielőtt az internet létrejött, sőt jóval azelőtt, hogy a számítógépek egyáltalán hálózatba lettek kötve. Ennek eredményeképpen egymástól független adattömegek keletkeztek, sok esetben még csak egységesített háttérállományok sem voltak használatban a felépítésükkor. A retrospektív katalogizálással létrejött digitalizált katalóguscédula-állományok is sok esetben ellenőrzés nélkül, pusztán

a cédulaadatok betöltésével jöttek létre. Ha volt is ellenőrzés, az sok esetben nem volt teljes körű. Emellett a digitalizált, papíralapú szövegállomány gyakran karakterfelismeréssel vált szöveggé. Még ha az egyes adatbázisok építésekor már használtak is valami segédállományt, szótárakat, értéklis-tákat, vagy éppen besorolási rekordok adatbázisát, akkor is az egyes tételek nem rendelkeztek univerzális azonosítóval, és a mindenkori igényektől, lehetőségektől, és az adatrögzítőtől függött, hogy mennyi információt rögzítettek például egy személy esetében.

A személynév-adatbázisok esetében, bár potenciálisan minden ember része lehet ezeknek, aki valaha megszületett, valójában több gyűjteménytípus esetében egy behatárolt csoport az, akiknek neve a névtérben megjelenik. A különféle integrált könyvtári rendszerek személynév-besorolási állománya általában igen rendezett, lényegében azok szerepelnek benne, akik szerzőként vagy esetleg más minőségben az egyes kiadványok létrejöttében közreműködtek. Ez a kör persze kiegészül azokkal a nevezetes személyekkel, akikről művek szólnak, audiovizuális dokumentumok esetében pedig az előadóművészekkel. Nagyjából ezt a névállományt ismerik és használják a könyv- és lemezkiadók is. A magyar nevek mellett persze minden ilyen adatbázisban (a gyűjtemény gyűjtőkörétől függő mértékben) találkozunk külföldi nevekkel is, melyek azonosítása különösen a nem latin betűkkel író népekhez tartozó alkotók esetében még nehezebb feladat.

Egy múzeumi gépesített leltárkönyv személyi adatbázisa ennél már lényegesen nagyobb lehet. Gyűjtők, tulajdonosok, képző- és iparművészek illetve népi iparművészek egészítik ki a személyek körét. Családfa-adatbázisok, illetve helytörténeti adatbázisok még sokkal több nevet tartalmaznak. Egy-egy településen a helytörténeti adatbázisok építői még megpróbálkozhatnak a nevek besorolásával, azonosításával, de nagyobb egységek esetében, például egy városi levéltári állomány, vagy egy katonai nyilvántartás esetében, már szinte keresztülvihetetlen feladat a személynevek ilyen módon való kezelése. A személynevek besorolásánál a név mellett egyéb személyes adatok, születési és halálozási év, születési és halálozási hely, és néhány egyéb adat már elegendő a személy egyértelmű azonosításához, de ezek az adatok nem mindig állnak pontosan rendelkezésre. Igen nagy e további adatok megadásakor a hibalehetőség is. A dátumokat könnyű hibásan begépelni, a települések esetében pedig a különböző írásmódok, eltérő nyelvi változatok miatt állhat elő az a

helyzet, hogy két személyi besorolási rekord betűhíven tekintve különböző, de tartalmilag valójában azonos. A foglalkozásoknál a sokféle hasonló értelmű fogalom párhuzamos alkalmazása okoz zavart, de gondot jelenthet a különféle, nemesi és tudományos címek, tiszti rendfokozatok, vagy rokonsági kiegészítők és azok rövidítéseinek nem egységes alkalmazása is.

A különféle gépi nyilvántartások tehát eleinte, a kezdeti időszakban még intézményi szinten sem voltak tartalmuk tekintetében egységesek. Amikor a közgyűjtemények egyre szélesebb körben – a könyvtárak a múzeumokat időben jelentősen megelőzve – megkezdték az integrált rendszerek használatát, és gépen tárolt adattartalmukat elkezdtek összetölteni, jelentős redundancia keletkezett. Ugyanazok a személyek egynél több névalakban is jelen voltak az összetöltött adatbázisokban. Az integrált rendszerekben, vagy ma már korszerűbb nevükön gyűjteménykezelő rendszerekben az összetöltésekkel létrejött duplikációkat a rendszerek szállítói és az intézmények munkatársai, fáradságos munkával igyekeztek felszámolni. Később az egyes intézmények integrált rendszerei is nagyobb egységekbe kezdtek szerveződni, egyre másra jöttek létre a területi (megyei vagy városi szintű) illetve intézményi hálózati, vagy ágazati együttműködések, melyek egyre gyakrabban építettek közös keresőket. Ezek vagy virtuális közös adatbázisra támaszkodtak, vagy már közösen épített egybetöltött nagy katalógusokat tartalmaztak.

A különféle IKR-ek, múzeumi és könyvtári illetve helytörténeti adatbázisok még ma is külön léteznek, de egyre másra épülnek az országos vagy európai szintű közös adatbázisok, például nálunk a MOKKA és az ODR adatbázisa, múzeumi szinten a *Muzeummap* adatbázisa, vagy európai szinten az *Europeana*. Minél inkább igyekszünk az adatbázisok összeépítésekor a redundanciát csökkenteni, és adateltérés esetén a validáltat elfogadni, és a kétséges, bizonytalan információkat egymástól megkülönböztetni, szétválasztani, valamint az azonos dologra vonatkozó külön rekordokat emberi munkával vagy számítógépes segédlettel összeolvasztani, annál értékesebb eredményt kapunk.

Időközben a közös katalogizálás teljesen elterjedté vált, a korábbi könyvtári katalógusok szinte teljes egészében, mind nálunk, mind külföldön gépre kerültek. A bibliográfiai rekordok átvételével a besorolási rekordokhoz tartozó tartalmak is átvételre kerülnek, így egymástól különböző adatok egyre ritkábban jönnek létre, mindamelllett az illető katalógusba máshonnan betöltött besorolási adat, például egy személynévrekord az új gépi katalógusban már önálló életet él, módosítható, átírható. A bibliográfiai rekordok letöltése csak a katalogizálási munka gyorsítását szolgálja, és maga a besorolási állomány, annak további adataival már nem kerül a letöltéskor a saját rendszerbe át. Így a redundancia mértéke csökken ugyan, de a közös katalogizálás a duplumokat ettől még nem számolja fel. A gépi katalógusrendszerek az import során gépi behasonlítást végeznek, mely nem teljes egyezés esetén új rekordként hozza létre az importált bibliográfiai rekorddal érkező információkból képződő besorolási adatot. Emberi erőforrás kérdése, hogy a könyvtár katalógusában az újonnan létrehozott besorolási rekordokat van-e mód átnevezni, jóváhagyni.

A kiadói adatbázisok, a könyvtári katalógusoktól, és főként a múzeumi gépi leltárkönyvektől ma még jórészt teljesen függetlenül épülnek. Lényegesen kevesebb adattal dolgoznak. A teljes szöveges tartalmak megjelenésével a redundancia még tovább bővült. Itt a digitalizált szöveget vagy automatikus karakterfelismeréssel, és esetleg későbbi emberi ellenőrzéssel teszik gépi úton olvashatóvá, vagy eleve begépelik. Mindkét esetben csak kevés figyelem juthat egy-egy név számára. Rendszerint a szövegben csak a név maga áll rendelkezésre, tehát az ilyen szövegek névállományának besorolása, csak a szöveggörnyezet, a dokumentum kezelkezési körülményei, vagy a teljes szöveghez kapcsolódó bibliográfiai adatok ismeretében lehetséges, sőt gyakori vezeték és keresztnevek esetében még ilyen helyzetben is szinte lehetetlen.

Adattisztítás, adategységesítés

A dokumentumok mind teljesebb körű digitalizálása és a tartalmak interneten való elérhetősége azonban új lehetőségeket is teremt az új feladatok mellé. Kereséseink során nehéz megtalálni a releváns adatot, hiszen az internet a világ legnagyobb és legrendetlenebb íróasztala, mégis az azonnal és kényelmesen elérhető adatforrások korábban elképzelhetetlenül nagy köre egyben lehetőséget is teremt az adatrendszerre, adattisztításra.

Beviteli hibák szűrése

Az adattisztítás első eleme, az adatbevitelkor előforduló redundancia minél szűkebb körre való szorítása. Az adatbevitelt a lehető legtöbb esetben és a legkényelmesebb, de hibákat nem megengedő módon végzett ellenőrzéssel, szótárak és háttérál-

lományok alkalmazásával lehet pontosabbá tenni. Ez az ellenőrzés az alkalmi felhasználói felületről történő adatbevitelt érinti. A rekordimport során szerencsés esetben a rendszernek elegendő információ áll rendelkezésére a behasonlítóhoz, amennyiben szükséges, emberi ellenőrzés során lehet validálni (jóváhagyni) a betöltött rekordokat. Tömeges import – konverzió esetében célszerű utólagos ellenőrzést végezni, illetve a gépi konverziónál redundáns rekordok keletkezését amennyire lehet megakadályozni.

Mit tekinthetünk besorolási rekordnak? A szükséges és elégséges adatkör kérdése

A besorolási jellegű információ, a leírásban szereplő harmadlagos metaadatok, vagy más néven tulajdonságadatok halmaza még nem besorolási rekord. Jó kérdés, hogy mennyi adat kell például egy személynév esetében ahhoz, hogy azt már önálló, más rekordoktól jól megkülönböztethető, az érkező személynevekkel behasonlítható rekordnak tekintsük. Ráadásul hiába határozzuk meg adatminimumot, ha más, integrálásra kerülő adatállományokból rendre információszegényebb tételek érkeznek. Amennyiben létrejönnek nemzeti névterek, közös katalógusok, a szakma már meg tud egyezni adatminimumban, legalább a továbbiakban felvételre kerülő nevek esetében. A személynevek azonosításához persze egyéb nem szorosan vett életrajzi adatok is jól használhatók, például a VIAF logikáját követve a művek is azonosíthatják a személyt. Sok esetben maga a személynév már annyira egyedi, ritka család- és keresztnévből áll, hogy magában is azonosítja az egyént, de a legtöbb név esetében ez nem áll fenn.

Paradoxon: „egységesített” névalakok különbségei

A különböző névtereknel, még ha nagy munkával önmagukban egységesek és konzisztensek is, az egymással való összehasonlításban mégis mindig lesznek különbségek. Különösen igaz ez a különböző országokban készített névterekre, és főként azokra a (közép Európában különösen nagy számú) személyiségekre, akiket több nemzet is saját kultúrája részének tekint.

Jó példa erre a régi Habsburg Birodalom számos alattvalója, például *Nikolaus Lenau* költő.

Adatait megtekinthetjük a Petőfi Irodalmi Múzeum webes katalógusának névtér felületén:

– <https://opac-nevter.pim.hu/>

- Adatbázis: Magyar életrajzi index (Név); Magyar családtörténeti adattár
- Egységesített név: Lenau, Nikolaus
- Születés/Halálozás: 1802-1850
- Névvariáns: Lenau Miklós (névvariáns), Niembsch Von Strehlenau Miklós
- Foglalkozás: költő
- Születési hely: Csatád
- Születési megye: Torontál vm.
- Pontos születési dátum: 1802. VIII. 13.
- Halálozási hely: Felső-Döbling
- Pontos halálozási dátum: 1850. VIII. 22.
- Rokoni kapcsolatok: Maygraber Mária TERÉZ Antónia 1771–1829: anya
- Adatforrások: Sátoraljaújhely: Sátoraljaújhely lexikona. Szerk. biz. elnöke Fehér József. Sátoraljaújhely, Kazinczy Ferenc Társaság, 2001. 548 o., ill. Szinnyei: Szinnyei József: Magyar írók élete és munkái. Bp., 1891-1914. Hornyánszky Viktor 14 db.

A költő műveire a Petőfi Irodalmi Múzeum katalógusa 68 találatot ad.

Mivel német nyelven író költőről van szó, a *Deutsche Bibliothek* természetesen ugyancsak tartalmazza rekordját, mint német költőt:

Nikolaus Lenau

Link zu diesem Datensatz

<http://d-nb.info/gnd/118571508>

Person	Lenau, Nikolaus
Adelstitel	Edler
Geschlecht	männlich

Andere Namen

Lenau, Nicolaus

Lenau, N.

Lenaŭ, Nikalaŭs

Lēnau, Nīqōlaus

Lenau, Miklós

Niembsch von Strehlenau, Nikolaus Franz

(Wirklicher Name)

Strehlenau, Nikolaus Franz von

Strehlenau, Nikolaus F. von

Niembsch, Nikolaus F.

Lenau, Nikolus

Niembsch von Strehlenau, Nikolaus F. (Wirklicher Name)

Lenau, Nīkalauš

Leñau, Nīqōlaus

Niembsch von Strehlenau, Nicolaus

Strehlenau, Nikolaus Franz Niembsch von

Strehlenau, Nicolaus Niembsch von

Niembsch von Strehlenau, Nicolaus Franz

Strehlenau, Nicolaus Franz Niembsch von
 Niembsch, Nicolaus
 Niembsch von Strehlenau, Nikolaus
 Strehlenau, Nikolaus Niembsch von
 Von Strehlenau, Nikolaus Franz Niembsch
 Niembsch, Nikolaus Franz, Edler von Strehlenau
 Lenau, H.
 Quelle M LCAuth, OGND
 Zeit Lebensdaten: 1802-1850
 Land Österreich (XA-AT);
 Ungarn (XA-HU); Deutschland (XA-DE)
 Geografischer Bezug Geburtsort: Lenauheim
 Beruf(e) Schriftsteller
 Funktion(en) sonstige Person (s) ;
 Textverfasser (Text)
 Weitere Angaben
 Österr. Dichter aus Csatad bei Temesvar; Edler;
 Österr. Dichter
 Dichter, Ungarn, Oesterreich
 Beziehungen zu Organisationen Internationale
Lenau-Gesellschaft
 Systematik 12.2p Personen zu
 Literaturgeschichte (Schriftsteller)

Integráció = Zűrzarav

Mindebből látható, hogy a két névtér integrációja itt redundanciát eredményezne.

Csatad? Ilyen község nincs. Látható, hogy a német szövegben születési helynek először Lenauheim van megadva, a „Weitere Angaben” címkénél pedig Csatad. A német (és sok más nyugat-európai) térképeken, kézikönyvekben gyakori jelenség, hogy a magyar ékezetet elhagyják, vagy tévesen tüntetik fel. A német név sokszor amúgy sem egyéb, mint a magyar névnek a németben nem jelenlevő ékezetek elhagyásával való feltüntetése. Különösen igaz ez arra a korra, amikor hazánk hivatalos nyelve a latin volt, és sok magyar községnek nem is létezett hivatalos magyar név-

alakja. Nézzük tehát a régi Torontál vármegyéhez tartozó, Trianonban elcsatolt Csatád község sorsát, nevének változásait: Csatád, neve románul Cetad illetve újabban Lenau iránti tiszteletből Lenauheim. Német neve újabban ugyancsak Lenauheim, régi neve Schadat, Schaddat láthatóan a magyar névből átalakítva. Ma község Romániában, a Bánságban, Temes megyében. 1920 előtt Torontál vármegye Csenei járásához tartozott.

A német kollégák nem tüntettek fel halálozási helyet. A PIM-ben megadott Alsó Döbling mint halálozási hely valójában soha nem volt Magyarország területén. A település mint *Széchenyi István* gyógykezelési és halálozási helye ismeretes lehet a történelem iránt érdeklődőknek. Valaha Unter- és Oberdöbling néven önálló falu volt Alsó Ausztriában, ma több más településsel együtt Bécs XIX. kerületét alkotja.

Lenai Miklós – Nikolaus Lenau – Lenau, Nikolaus – Lenau, Nicolaus

Láthatóan a PIM helyesen, mint német nemzetiségű személyt nevének német alakjában tüntetik fel. A MOKKA-ban keresve ugyancsak német név-alakkal szerepel, de kétféle módon, és mindkettőhöz tartoznak találatok.


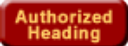

név(születés-halálozás)	találatok száma
Lenau, Nicolaus (1802-1850)	369
Lenau, Nikolaus (1802-1850)	352

Szerzőségi közlésben természetesen számos esetben szerepel a Lenau Miklós névváltozat is, de ismert költőről lévén szó, a könyvtáros kollégák mindig a német alakhoz sorolták be a MOKKA adatszolgáltató könyvtáraiban.

Csak kitekintésképpen megmutatom az érintett találatokat Lenau névre a *Library of Congress* katalógusából is (1. táblázat):

1. táblázat

Találatok a Library of Congress-ben

	3	1	Lenau, M. [from old catalog]	personal name
	4	0	Lenau, N. (Nicolaus), 1802–1850	personal name
	5	85	Lenau, Nicolaus, 1802–1850	personal name
	6	16	Lenau, Nicolaus, 1802–1850 [from old catalog]	personal name
	7	2	Lenau, Nicolaus i.	personal name
	8	1	Lenau Nicolaus i. e. Nicolaus Franz Numbsch, edler von Strehlenau, 1802–1850. [from old catalog]	personal name
	9	0	Lenau, Nikolaus, 1802–1850	personal name

Ugyancsak érdemes tanulságképpen megnézni rekordját a VIAF-ban is:

[https://viaf.org/viaf/36931332/#Lenau, Nicolaus, 1802-1850](https://viaf.org/viaf/36931332/#Lenau,_Nicolaus,_1802-1850)

Mindenesetre, ha bármilyen integrációt meg akarunk valósítani és rekordokat akarunk egymással összevetve, gépi és emberi erővel egységesíteni, igen nagy fába vágjuk a fejszénket.

Névterek és gyűjteménykezelő rendszerek integrációjának kérdései

Önmagában relációs adatbázis, vagy adatbázisok szemantikus hálója

A gyűjteménykezelő rendszerekben a rendszer logikájától függően rendszerint megkülönböztetünk bibliográfiai rekordot, példányrekordot, besorolási rekordot, ezekből az egyes rendszerektől függően többfélé is, a múzeumi rendszerekben pedig, múzeumi leltáritétel-rekordot. A relációs adatbázisokban, kapcsolatok sokasága definiálанд az egyes rekordok között. Ez a hagyományos szemlélet évtizedek óta uralkodó a gyűjteménykezelő rendszerekben, és mivel a bibliográfiák tekintélyes részét könyvtárak építik, és a bibliográfia maga is már nem más, mint a gépi katalógus valamilyen formában való reprezentációja, illetve a közös adatbázisok is a könyvtári gépi katalógusok összeépítésével jöttek létre, a rekordalapú szemlélet az adatbázisokban általánosan uralkodóvá vált.

A bibliográfiai rekordok, miközben rendszerről rendszerre konvertálásra kerülnek, tovább léteznek akkor is, ha az intézmény lecseréli gyűjteménykezelő rendszerét. Továbbra is fennmaradnak akkor, ha – rendszerint szervezeti változások eredményeképpen – a gépi katalógusok részeként összetöltik őket. Emellett, rekordmásolással, le- és betöltéssel sokszorozódnak, és sorsuk, miként az osztódással szaporodó élőlényeké, egy-egy másolással újra meg újra el is válik, ezért a kezdeti tökéletes hasonlóság az idők folyamán, mivel nem egyszerre és nem egyformán változnak, egyre csökken. Ugyanez a sors éri a besorolási rekordokat is, ráadásul itt a különféle áttöltésekkel még a rekord létének folytonossága sem feltétlenül áll fenn.

A gépi katalógusokat támogató relációs adatbázisok csak annyiban nem ismétlik az adatokat, és annyiban tekinthetők szemantikus jellegűnek, amennyiben a besorolási rekordok nem másolódnak be a bibliográfiai rekordba, hanem csak hivat-

Simon A.: Rekord vagy háló – tudásreprezentációs ...

kozásként, egy azonosítóval vannak jelen, így a különféle felületeken és nyomtatott outputokban a bibliográfiai rekordhoz kapcsolt más bibliográfiai rekordok, vagy besorolási rekordok mindig a pillanatnyi állapotukban látszódnak, mutatkoznak meg.

Egy szemantikus hálóként felépülő rendszerben, ez az elv van még következetesebben megvalósítva. Az idegen rendszerből érkező információ itt csak linkelve van, nem ténylegesen begépelve. A fél szemantikus rendszerekben, a ma még korszerű integrált rendszerek relációs adatbázisaiban, amikor importálunk egy besorolási rekordot, az ugyanúgy létező rekorddá válik, mintha ténylegesen begépelnénk, vagy importtal bejuttatnánk az adatbázisba. Az adat tehát, ha az import program új besorolási rekordként érzékeli, azaz a behasonlítás során a saját adatbázisban valamiért nem találja meg, az import során duplikálódik, és amennyiben a behasonlítás hibásan nem észlelt egyezést egyetlen létező rekorddal sem, redundancia állhat elő. A redundancia oka lehet a behasonlító program működési hibája, gépelési hiba valamelyik rekordban, vagy valamilyen adatkülönbség, mely a rekordokat létrehozók eltérő adatfelfogásán, vagy szakmai nézetkülönbségen is alapulhat.

Tökéletesen relációs rendszerekben a kívülről jövő adatot linkként integráljuk a gyűjteménykezelő rendszerekbe. Nem elég, ha a gyűjteménykezelő rendszerben lévő személyneveket kapcsoljuk a névtérhez, a gyűjteménykezelő rendszereknek közvetlenül kell használniuk a névteret a munkafolyamataik során.

Ehhez az alábbi feltételeknek kell(ene) érvényesülniük:

- a kívánt besorolási adatot a névtérből kell választani (pl. szerző műhöz, illetve kiadványhoz rendelésekor),
- ha a rekordon változtatásokat kell eszközölni, a névtérben kell a szükséges módosításokat elvégezni,
- a keresési szolgáltatásokat integrálni kell,
- tökéletes relációs szerkezetnek kell érvényesülni, tehát MARC adatfolyam helyett csak címkékkel ellátott, önálló, egymással összekapcsolt adatelemek lehetnek az adatbázisban jelen.

A legfőbb építőelem

A gyűjteménykezelő rendszer igényei mások, mint egy névtéré. A gyűjteménykezelő rendszer a gépi katalógus igényeit követve lett tervezve. Ezzel

szemben a névtér a szakértői rendszer logikája szerint épül fel. Egy gyűjteménykezelő rendszerben természetesen jelen vannak a személynév besorolási rekordok, de nem jelent kulcsproblémát a rendszer számára, ha nem érvényesül tökéletesen az egy rekord egy ember elve. A gyűjteménykezelő rendszer alapegységei a bibliográfiai rekordok, melyeket elsődleges metaadatoknak is nevezünk. A névtér esetében azonban alapelv, hogy meghatározzuk azt az entitást, mely a rendszer alapegysége, és amelyhez a többi információt hozzáköthetjük. Egy személynévtér esetében a rendszer alap építő eleme a személy, az ember, és egy rekordnak mindig egy személyt lenne szabad csak azonosítania. De mi azonosít egy embert? Tekinthejtük besorolási névalaknak azt, ami a személyi igazolványban van, vagy ami az irodalomban ismert és elfogadott név. Gond van persze a régi korokkal, amikor még nem volt személyi igazolvány, esetleg népszerűnyilvántartás, sőt akár még egyházi anyakönyvek sem. Probléma támadhat a többnemzetiségű személyeknél, a nem latin betűs átírásoknál, a nemesi előneveknél, a többes keresztnéveknél. Mindenesetre az egy ember egy névtér rekord elv mellé odatehetjük, hogy a besorolási névalak az a (születéskori) név, amely az illető személyi igazolványában áll, vagy állt volna, ha lett volna személyi igazolványa. Kivételt képez hölgyeknél az asszonynév, ha írói munkássága azon ismertebb, illetve egy személy név-változtatása esetében az a névalak, melyen írói tevékenységét kifejtette. Többnemzetiségű személyeknél, amennyiben az illető legalább valamennyire magyarnak tekinthető, (pl. Kitaibel Pál), hazai készítésű névterekben a magyar nevet érdemes kitéüntetettnek tekinteni, mivel nálunk, a magyar nyelvű könyveken általában a magyar névalak szerepel a dokumentumokra rányomtatva, és hazai körökben általában ez az ismertebb névalak is. Ellenpéldának tekinthető a már említett Nikolaus Lenaun kívül a pozsonyi születésű Johann Nepomuk Hummel.

Rekordalapú szemlélet, állításlapú szemlélet

A gépi könyvtári és múzeumi katalógusokban a rekordalapú szemlélet uralkodik. A bibliográfiai rekord a katalóguscédula illetve a leltárkönyvi bejegyzés utóda. A besorolási rekord ugyancsak a személynévhez besorolt könyvtári bibliográfiai leírás utóda, nevét is ebből a funkcióból örökölte.

Igazából az adatbázisok tervezésekor egymástól némileg függetlenül, két kérdés merül fel. Egyrészt, hogy állandó adatszinkron vagy közvetlen

linkek formájában való adatmegjelenítés érvényesüljön-e, másrészt, hogy rekordalapú, az eminens és nem eminens értékeket megkülönböztető, vagy minősített állításokra épülő, azokat alternatíván megjelenítő rendszert képzelünk-e el. Ezek egymástól függetlenül megválaszolhatóak, mégis, alapvetően két modell állítható fel logikusan. Az egyik, a rekordalapú szemlélet, kiválasztott eminens értékekkel, és ténylegesen létező adatbázisokkal, melyeket valamilyen módon szinkronban tartunk a külső adatforrásokkal. A másik, a szemantikus háló elveit következetesen érvényesítő adatszerkezet, mely állítások halmazaként fogja fel az adatbázist, melyben a kívülről jövő állítások együtt jelennek meg „saját” állításainkkal. A MARC rekordok között a hagyományos katalogizáló rendszerben létrehozott relációk itt ugyancsak állításként fogalmazódhatnak meg.

Ez utóbbi jövőbe mutatóbbnak tűnő megoldás persze egy sor gyakorlati kérdést is felvet.

Meg kell oldani a besorolási rekord jelenlétét a bibliográfiai rekordban. Ezek terei: az – egyre ritkábban nyomtatott – ma inkább már csak a nyilvános felületen létrehozott katalóguscédula, a különféle, képernyőre megjeleníthető, letölthető, vagy bibliográfiákba másolható megjelenítési formátumok, az adatcsere-formátumok, (ha még valaki adatokat akar cserélni), a böngésző listák, és természetesen maga a katalogizáló felület. Ez utóbbi a legfontosabb, hiszen a felhasználói felület meghatározza a gondolkodást, döntő befolyást gyakorol mint sorvezető, és mint a gyakorlati munka meghatározója a feldolgozó könyvtárosra. Visszahatásként, a feldolgozás módja döntően hozzájárul a szolgáltatott rekord milyenségéhez, azzal együtt is, hogy ma már az adatrögzítő általában igyekszik a felhasználó fejével gondolkodni.

Névtérállításból besorolási rekord

Amennyiben a névteret nem adatbázis-rekordokból építjük fel, hanem RDF-állítások halmazaként képzeljük el, ezzel megoldjuk azt a kérdést, hogy mi történik akkor, ha az adattömeget több gépi katalógusokból feltöltve, sőt összetöltve, egymástól különböző, sőt egymásnak ellentmondó információk érkeznek. Ezek egy része lehet persze gépelési hiba, más része lehet eltérő adatfelfogásból eredő különbség. Megint más része már lehet eltérő szakmai felfogás, meggyőződés. Ugyancsak ki lehet iktatni a rendszerből azt a döntést, hogy mit tekintünk besorolási névalaknak a többféle változat közül.

Az adateltérések egy része lehet olyan információ, melyet egységesítéssel ki lehet iktatni. Például lehet standardizálni a dátumformátumokat. Ugyancsak közmegegyezéssel lehet kezelni a településneveket. Például Lenau szülőhelye esetében Csatád községet fogadjuk el besorolási névalakként, mint egy hajdan Magyarországhoz tartozott település magyar hivatalos nevét, minden más alakot névváltozatként rögzítve. Az Alsó Döblinget meg lehet feleltetni a német névalaknak, Unterdöbling, de persze Döblingnek vagy akár Wien-nek (Bécs) is. E sorok szerzője is támogatja azt a magyar szakmában uralkodó közmegegyezést, hogy a települések besorolási névalakja a mindenkori hivatalos közigazgatás szerinti egység akkori hivatalos neve, Magyarország esetében meghagyva a Trianon utáni időszakban besorolási névelemként a települések hivatalos magyar nevét. Tehát Lenau esetében születési helye Csatád, halálozási helye Unterdöbling. Csatád további névváltozata a se-regnyi névalak, míg Unterdöblinget keresztthivatkozással lehet Wien-hez kapcsolni. Amennyiben a névtér állításokra csak linkként hivatkozunk, akkor belátásunk szerint használjuk és jelenítjük meg a különféle változatokat.

Ha nem közvetlenül az adatbázisba illesztett külső hivatkozással élünk, akkor meg kell oldanunk a konverziót és adategységesítést, egyszer a névtér felől, majd pedig ugyanezt a munkát automatikus adatszinkron formájában is végeznünk, végeztetnünk kell, hogy a két adatbázis, a saját katalógusunk és a névtér állandóan aktuálisan egyforma legyen.

Katalógusrekordból névtérállítás

Amennyiben elvben elfogadtuk, hogy a könyvtári illetve múzeumi katalógusból vagy azok már korábban létrejött közös összetöltött állományaiból építjük fel névterünket, néhány gyakorlati kérdéssel is szembe kell néznünk.

A gépi katalógusok besorolási rekordjai, köztük a személynév besorolási rekordok is a katalogizálás során egyedi rögzítés, vagy más – akár külföldi – rendszerből történt letöltés során jöttek létre. Egyszeri rögzítés után sok felhasználás történt, mely egyrészt alkalmat kínált az ellenőrzésre, de a gépelési és esetleg tartalmi hibák szétterjedésére is. A sok felhasználás során a rekord egyszerűsödhetett, hiszen az importok során nem mindig minden adatot töltöttek be, a különféle célú és szerkezetű katalogizáló rendszerekbe. A rekord természetesen a későbbi szerkesztések során meg is változ-

hatott. A katalogizáló munka során tömegével lehetnek, „kvázi” besorolási rekordok, vagy más néven „skeleton – csontváz” rekordok igen kevés adattal. Ezeket egyes rendszerek megkülönböztetik az „igazi” besorolási rekordoktól, másik viszont nem. Különösen jellemző ez a sors a kevésbé ismert szerzőkre. A besorolási állomány névtérre építése során szigorúan meg kell határozunk azt az adatminimumot, mely alapján alkalmasnak tekintünk egy rekordot arra, hogy egy névtérbe legyen betöltve. A műszaki feltételeken túl tartalmi szempontokat is fel kell állítanunk, hogy megállapíthassuk, mit tekintünk besorolási rekordnak, van-e adatminimum (baj, ha nincs).

Integráció idegen rendszerekkel

A könyvtári katalógusokon túl felmerül, hogy más rendszerekből is be kell adatokat fogadnunk.

Külföldi adatbázisok esetében itt nehézséget jelent az eltérő felfogás, a nehéz azonosítás, és természetesen a sok nem magyar név. Igazából nemzeti névterek építésében kell gondolkodni, és az idegen szerzőket idegen könyvtári katalógusokra épített névterekből, vagy éppen a VIAF-ból átvenni. A nemzeti névterek ma, ha egyáltalán léteznek és annak nevezhetők, alapvetően még relációs adatbázisok. A fentebbi példákban is látható volt, hogy számos átfedés van, sok személy több nemzethez is tartozik. A könyvtári katalógusokban minden országban a rengeteg személynév valójában külföldi szerző. Ha ezeket a névtér építése szempontjából figyelmen kívül hagyjuk, tisztább képet kapunk. Az egyes érintett könyvtárak emellett saját katalógusukban önálló névállománnyal is dolgozhatnak, nem csak a nemzeti névtérrel.

A múzeumi gépi leltárkönyvek nagyon sokféle személyi adatot tartalmaznak. Előadó és alkotóművészek, népi iparművészek, gyűjtők, muzeológusok, előző tulajdonosok, ábrázolt személyek, tudósok sokasága kerül ezekben felvételre. Nagy részükből nem is képződik soha igazi besorolási adat. Gondot okozhatnak az adatszegény leltárkönyvi bejegyzések, nagy redundancia lehet az egymástól függetlenül épített gépi leltárkönyvekben. A helytörténeti adatbázisok önmagukban igen jól szerkesztettek lehetnek, sokszor tartalmaznak családfákat, de egymással való összevetésük halatlan feladat. Emellett az adatok a helyi személyekről önkéntes adatközlőktől érkeznek, ezeket az adatokat valakinek a többivel össze kell vetnie, és ellenőriznie kell. Erre sokszor nincs az adatbázis építése során idő.

A levéltári adatok, családfák, digitalizált anyakönyvek állománya még több nevet eredményez. Valahol határt kell szabnunk az állomány bővülésének, hiszen jelenlegi technikai viszonyaink között elképzelhetetlen egy olyan adatbázis felépítése, melyben ellenőrzöttek és azonosítottan jelen van szinte mindenki, aki valaha megszületett.

Adatbázisaink jövője

Miből lesz végül adatbázisunk, avagy a névtérrekordok nem potyognak az égből

Jó kérdés, hogyan lesz akkor, miből épül fel a névtér adatbázisunk.

Amennyiben tiszta lappal indulunk, a névtér-állomány csak sokára lesz használható. Csak fizetett személyzet munkája során töltődik fel, nagyon nagy költséggel.

Jó megoldás lehet a kiindulás egy(néhány) kitüntetett adatbázisból, de persze a szakmának, vagy valamilyen hatóságnak kell eldöntenie melyik(ek)ből? Ha szakítunk a rekordalapú szemlélettel, és állítások sorát képezzük le a több helyről is beérkező adatokból, ezzel kihúzzuk a méregfogát annak a vitának, hogy kinek az adatai jobbak, pontosabbak. Egy közös adatbázis igazából jó terepe lehet annak, hogy szakmailag ellenőrzött, egységes és ellentmondásoktól szinte mentes adattömeget nyerjünk, de ha megengedjük párhuzamos állítások létét, ennek az egységes adatbázisnak a létrejöttét lényegesen lelassítjuk, mert közvetlen gyakorlati szükséglet hiányában minden adattisztítás lassabban halad.

Nem mindig a központi adatbázis a legjobb. Egyes szakterületek, városok, megyék esetében a helyi adat használhatóbb, pontosabb. Mivel a világban a hatékony munkavégzés az egyszemélyi felelősségre épül, igen fontos az adatgazda intézménye. Mindig jó, ha egy szakember, vagy intézményvezető személy szerint felelősséget vállal a feltöltött rekordért. Ha egy rekordért senki sem vállalja az adatbevitel és a jóváhagyás tekintetében a felelősséget, az az illető rekord esetében szerencsés, ha egyértelműen látszik. Célszerű, már csak történeti szempontból is a helyesnek és helytelennek tartott adatok, tehát az eminens és a nem eminens rekord egymás melletti megőrzése. A múzeumokban a leltárkönyv szerkesztésekor előírás, hogy az adatok megváltoztatása esetén a változás tényét, idejét és az eredeti adattartalmat továbbra is fel

kell a leltárkönyvben tüntetni. A számítógéppel szerkesztett múzeumi leltárkönyvekben, a gyűjteménykezelő rendszerek ehhez az úgynevezett „historia” – „történet” funkciót biztosítják. A rendszer itt nemcsak feltünteti a megváltoztatott adat eredeti állapotát, a változtatás végzőjét és idejét, hanem megfelelő jogosultság esetén annak visszaállítását is lehetővé teszi. Ez a gyakorlat a nagyobb könyvtári gépi katalógusokban ugyancsak segítheti a szakmai munkát, és a viták nyomán követését. Az alternatív állítások egymás mellett tartása azonban bizonytalanságot szülhet a szakemberekben is, tehát itt ezeket minősíteni kell, fel kell jegyezni, honnan származnak, és amennyire lehet, minden esetben valamelyiket eminensként meg kell jelölni. Amennyiben szakítunk a katalóguscédula/rekordalapú szemlélettel, és a kapott rekordokból állításhalmaz építésébe kezdünk, az adatközlő megjelölése különösképpen elengedhetetlen. Az adatokra épülő nyilvános felületeken való szolgáltatások kialakításánál, az eminens rekordokra, illetve az adatközlő, vagy adatgazda szerint leválogatott értékekre figyelve kell összeállítani a nyilvánosságra hozandó információk körét.

Az adatbázisok összetöltésük után, ha nincs duplumellenőrzés, rengeteg párhuzamos adatot fognak tartalmazni. A hagyományos kézzel végzett duplumszűrés végtelen történet. Közéletben jó megoldás lehet a folyamatos szakmai ellenőrzéssel, és a nyilvános felület használóitól érkezett jelzéseket figyelembe vevő karbantartással párosuló rendszerkönyvtárosi munka, mely a gépi környezet eszköztárát hasznosítja. Kiegészítve ezt a gépi duplumszűréssel, mely az emberi gondolkodást és asszociációt modellező, és az ilyen lépéseket hallatlanul gyorsan elvégző, adatbáziskezelővel végrehajtott lépések sorát jelenti, már igen jó eredményre számíthatunk.

Alapvetően sokféle gépi megoldás képzelhető el. Vannak ismert jó gyakorlatok. Az alapvető adatokból duplumkulcs készíthető, ezekkel duplum párok jönnek létre. Itt a kapcsolódás erőssége pontoszhozható, a bizonytalan pontszámúak mind átnézhetőek, a biztosnak vélteket pedig érdemes szűrőpróbaszerűen áttekinteni. Természetesen bizonytalan esetekben mindig embernek kell kimondania a végső szót.

Amennyiben állításhalmazt építünk, tudatában kell lennünk annak, hogy e mögött is, a mai látókörünkbe eső technika esetében még adatbázisrekordok állnak. A technikán persze túlléphet, és mindig túl is lép az idő, és akkor a névteret tartal-

mazó adattömeget migrálni kell valamibe, akár egy rendszerbe, ami mögött már nem is relációs adatbázis áll.

Hol lesz az adatbázis?

Azt kell mondanunk, hogy igazából mindegy. Valahol egy lemezen ott lesznek a byte-ok, mert minden digitális tartalom bit-ek és byte-ok formájában ott van egy számítógép merevlemezén valahol. Fontos, hogy legyen róluk mentés, biztonságban. Országos hatókörű, nemzeti kincsnek számító adatbázis esetében mindenképpen belföldön, és mindenképpen egy másik városban.

A felhőalapú megoldás önmagában még nem hatékonyabb, nem jobb és nem korszerűbb. Egy-egy felhasználó esetében (valószínűleg a többségük-nél amúgy) racionális döntés informatikai szolgáltatásaik jó részét nem saját szerverről, hanem külső eszköztől szolgáltatni. Ez akár tekinthető felhőalapú megoldásnak is, de ha valamiről nem tudjuk, hogy hol van, attól az még van valahol.

A szolgáltatások és az adattárolás központosításáról is e sorok szerzőjének az a véleménye, hogy sok esetben hatékony és jó, de önmagában nem megoldás semmire, még csak az sem biztos, hogy minden helyzetben ezt érdemes választani, és hogy biztosan korszerűbb vagy előremutatóbb. Amennyiben egy központi eszközre összeépítünk szolgáltatásokat anélkül, hogy műszaki és szakmai szempontok szerint összeépítenénk, összehangolnánk őket, akkor azok nem együtt, csak egymás mellett fognak működni. Amikor adatbázisokból közös katalógusokat építünk, anélkül hogy az összetöltést, összeolvasztást, erőteljes szakmai felügyelettel kísérnénk végig, a közös adatbázis nem lesz jobb. Illúzió az, hogy az összetöltéseket lehet automatizálni, rengeteg emberi felügyelet, emberi döntés kell ma még, ameddig gondolkodásunkat, asszociációs rendszerünket nem tudjuk algoritmizálni, hogy azt programozók alkalmazásokba tudják építeni.

Simon A.: Rekord vagy háló – tudásreprezentációs ...

Adatbázis lesz?

Véleményem szerint még sokáig igen. Minősített, címkézett, eredetmegjelölt, és egymással sok éllel definiáltan összekapcsolt adatsomópontok sokasága fogja alkotni a Tudás Szervező Rendszeret, halvány utáztatként végezve el azokat a feladatokat, melyekre igazából a „világegyetem legcsodálatosabb anyaga” az emberi agy hivatott.

Hivatkozások

¹ UNGVÁRY Rudolf: Névterek és földrajzinév-tárak. = TMT 63. évfolyam (2016) 4. szám

² UNGVÁRY Rudolf: A besorolási adatsere-formátum bővülése. A legutóbbi két évtized fejleménye. = TMT 58. évfolyam (2011) 9. szám

Egyéb felhasznált irodalom

BÁNKI Zsolt – MÉSZÁROS Tibor – NÉMETH Márton – SIMON András: Azonos személyekre vonatkozó név besorolási rekordok automatikus felderítése a PIM adatbázisában. = TMT 63. évfolyam (2016) 12. szám

DANCS Szabolcs: A Bibliográfiai Átállás programja – az RDA franciaországi átültetése. = TMT 63. évfolyam (2016) 9. szám

DANCS Szabolcs: Bemutatkozik az ISNI, a nemzetközi szabványos névazonosító. = TMT 64. évfolyam (2017) 4. szám

LENGYEL Mónika – ANDAHÁZI SZEGHY Viktor – MAGYAR Gábor: Mindenki másképp csinálja! A retrospektív konverzió két útja. = TMT 59. évfolyam (2012) 10. szám

UNGVÁRY Rudolf: Metaadatok összehasonlító vizsgálata gyűjteményi rendszerekben. = TMT 58. évfolyam (2011) 6. szám

UNGVÁRY Rudolf: A névterek és az adatok tulajdonságai. = TMT 59. évfolyam (2012) 3. szám

UNGVÁRY Rudolf: A névtér mint kulturális szükséglet. = TMT 59. évfolyam (2012) 8. szám

Beérkezett: 2017. IX. 5-én.



Simon András

a MONGUZ Informatiótechnológiai Kft. ügyfélmenedzsere.

E-mail: asimon@monguz.hu