

## Névtérkezelő készítésének nehézségei

*A névterek a harmadik informatikai forradalom következtében alakulnak ki. A hatalmasra duzzadt – főleg internetes – információmennyiség kereshetőségének legalkalmasabb rendező rendszere maga a természetes nyelv. Ebből a felismerésből kiindulva kezdik el az egyes nemzeti kultúrákban a nyelv szavait a jelentésükből következő szemantikai rendszerekbe szervezni, és a világháló hivatkozási lehetőségeit (ugrópontjait) felhasználva hozzákapszólják a szavakhoz mindazokat az elérhető forrásokat, melyek e szavak jelentését elmélyítik, részletezik (mint a különféle értelmező szótárak, lexikonok enciklopédiák, tanulmányok, osztályozási rendszerek és teauruszok, igazgatási nomenklatúrák és biológiai taxonómiák stb.) adatait. E fejlődésnek még csak az elején járunk. A névterek nem egyebek, mint a hagyományos értelmező nagyszótárak sokkal összetettebb, adatokban gazdagabb, az informatika lehetőségeit kihasználó változatai.*

Tárgyszavak: névtér; rendező rendszer; adatmodellezés; adatbázis-kezelő rendszer

### Bevezető

A névterekre elmondható, hogy „minden névnek az számít, ami rögzíthető”. Neveken itt nemcsak egyedek nevét, hanem közneveket (szavakat) is értünk. Ez ma nagyon speciális szakterület. A névterek a jövőjüket illetően nem egyebek, mint a hagyományos értelmező nagyszótárak webes változatai [31]. De mint ilyenek, a világháló lehetőségei következtében nagyságrendekkel gazdagabb adatállományokkal és alkalmazási lehetőségekkel [26]. Az értelmező nagyszótárak mindig csak közneveket tartalmaznak, és kerülnek bennük az összetett kifejezéseket. Névterekben nincsen területi korlát. Ezért az összetett közneveken, szakkifejezéseken kívül van olyan részük, melyben a személy- és testületi nevek, továbbá a földrajzi nevek is szerepelnek. Mindez óriási névállományokat jelent.

Ugyanakkor az informatikában a névterek még jóformán ismeretlenek, és nincs érdemi tervezési tapasztalat.

Pedig a névterek éppen a harmadik informatikai forradalom<sup>1</sup> következtében születtek meg [17], [18], [32]. A hatalmasra duzzadt – főleg internetes – információmennyiség kereshetőségének legalkalmasabb rendező rendszere maga a természetes nyelv. Ebből a felismerésből kiindulva az egyes nemzeti nyelveket kell átfogó, nem mesterséges, hanem szigorúan a nyelv szavainak jelentéséből követke-

ző szemantikai – névtérnek nevezett – rendszerbe szervezni, és ehhez a web hivatkozási lehetőségeit felhasználva hozzákapszolni minden olyan forrást, mely e szavak jelentését tartalmazza. Ez ugyan annak a fejlődésnek egy újabb, de sokkal komplexebb szintje, mint ami az informatikát közvetlenül megelőző legújabb korban, a 18. században az enciklopédiákat majd lexikonokat, a 19. század végétől kezdve pedig a nemzeti nyelvek értelmező nagyszótárait létrehozta. E fejlődésnek még csak az elején járunk [27], [28].

### Történeti áttekintés

A tapasztalat hiánya nem lenne szükségszerű. Hogy mégis így van, annak történeti oka van. Noha a névterek elődei, a teauruszok már az 1960-as évek elejétől készültek, megszakadt a folyamatosság a tervezésben. A kevés külföldi teauruszkezelő rendszert vagy maguk a készítőik hozták létre, vagy különösebb tapasztalatok nélküli, kis cégek. Ezekre a rendszerekre túlnyomórészt az egyszerűség volt jellemző. A legjobb esetben is csak a teaurusz-szabványokban rögzített szöveges megjegyzéseket és szemantikai relációt tudták kezelni [3], [4], [6], [7], [8], [11].

Magyarországon már 1974–75-ben készült teauruszkezelő rendszer, a *Schlanger László* által készített TSPR [13]. A hazai volt az irányadó. A rendszer – lévén, hogy még az online kezelés se

terjedt el – offline módban működött. Az adatokat, beleértve a módosításokat is, lyukkártyákkal kellett bevinni.

A TSPR-rel több mint egy tucat tezausz készült el. Mindennek két szempontból is jelentősége volt. Egyrészt sikerült a rendszerrel kezelt állományt mágnesszalagon megőrizni, másrészt már kezdetől fogva tapasztalatok halmozódtak fel. Ez vonatkozik mind a kezelőrendszer készítésére, mind pedig a tezauszok szerkesztésére.

A TSPR-rel készült tezauszokkal együtt 1989-ig Magyarországon 39 tezausz készült el<sup>2</sup> ezekből 13 a TZSPR-rel készült. Tucatnyit használtak. Sok tezausz nyomtatva vagy géppel olvasható hordozón rendelkezésre állt.

A nyolcvanas években az Országos Műszaki Információs Központ és Könyvtár számára *Stiegrád Gábor* készített egy MicroCDS/ISIS-alapú tezauszkezelőt. Ezzel nemcsak a Központ műszaki tezausza készült el, hanem az OSZK periodika-adatbázisának tezausza is. Legalább annyira fontos volt, hogy a két tezauszba beolvaszhatuk a korábban TSPR-rel kezelt adatállományokat.

A TSPR-t csak az ESZR gépekhez használhattuk; a fejlődés és a rendszerváltással elfújta ezeket a berendezéseket. A Mikro CDS/ISIS-alapú tezauszkezelővel nehézkesen dolgozhattunk, nem igazán online működött.

*Ungváry András*sal 1993-ra készült el saját vállalkozásban a RELEX tezausz-szerkesztő. Első változata DOS-alapon működött, vele a Word szövegszerkesztővel kezelt állományt lehetett végre ellentmondás-mentességet biztosítva kezelni [13], [25].

Ettől kezdve Magyarországon lényegében minden új tezausz és minden karbantartás a RELEX segítségével valósult meg, függetlenül attól, hogy szabad piacon végzett vagy munkahelyi munkáról volt szó. Ezzel a rendszerrel készült el 2000-ben az átfogó magyar tezausz, a Köztausz, melynek ma közel 120 ezer lexikai egysége és rendkívül sok metaadata van [7].

Az OSZK-ban folyó tezauszkészítéshez is ezt a saját, tehát nem OSZK-fejlesztésű programot használtuk. 1992-ben elkészült a C++ nyelven írt változat (4.0), 2002-ben az Access/Windows-alapú Visual Basic nyelven írt rendszer (6.0 változat).

2012-ben pedig – már az OSZK szervezésében – elkészült a több felhasználó számára hozzáférhető kliens-szerver alkalmazás, melyben maradéktalanul érvényesültek a RELEX-szel szerzett tapasztalatok. Az új rendszerrel már nemcsak tezauszok, taxonómiák, osztályozási rendszerek (például az ETO), hanem ontológiák [20], [21] és egyszerűbb névterek osztott kezelése is megvalósítható volt [26].

Mindebben az a különösen jelentős, hogy ezek a rendszerek a használatuk tapasztalatai alapján egymásra épültek és ergonómiailag fokozatosan egyre inkább közelítettek egy felhasználóbarát változathoz.

### **Informatikai kivitelező és professzionális végfelhasználó**

2017 nyaratól folyik a nemzeti névtér kezelőrendszerének tervezése. Ennek a rendszernek ugyan átfogóbb feladatokra kell alkalmasnak lennie, mint a tezauszok esetén, de mindenképpen be kell, hogy épüljenek a magyarországi tezauszkészítő rendszerekkel szerzett tapasztalatok is [15], [17], [18]. Mindez nem egyszerű feladat, mivel az informatikai rendszerek készítőinek ilyen irányú tapasztalatai nincsenek. Ezeket a tapasztalatokat tehát nem informatikus, hanem az információkereső nyelvekkel és osztályozási rendszerekkel foglalkozó szakembereknek kell közvetíteniük, akik a tezauszkészítés és -kezelés ismereteivel rendelkeznek. A munka során, az informatikailag viszonylag járatlan terület következtében kommunikációs problémák áthidalására is szükség van.

Első lépésben a kezelőrendszer adatbázisa a nemzeti névtér földrajzi részével lesz feltöltve. Ennek érdekében a kezelőrendszer kialakításával párhuzamosan, egyelőre a Relex segítségével elkezdődött az elérhető földrajzi nevek egyesítésével ennek az induló állománynak a kialakítása és szerkesztése. Az elérhető három legnagyobb állomány a Köztausz földrajzi neveit tartalmazó Geotaurusz, a Központi Statisztikai Hivatal hivatalosan nyilvántartott földrajzi nevei és a Budapest Főváros Kormányhivatalán belül működő földmérési és térinformatikai államigazgatási szerveként működő Földmérési Távérzékelési és Földhivatali Főosztályában nyilvántartott, ugyancsak hivatalos földrajzi névállomány. Az egyesítés hosszú ideig tartó egyeztetési-behasonlítási munkával jár, és a tervek szerint a névtérkezelő elkészítésével egy időben lesz kész [16], [22], [23], [29], [30].

A kulturális örökség adatait és a tudástárakat kezelő rendszerek készítésében egyes cégekben mára ugyan felhalmozódtak tapasztalatok, de ezek könyvtári, múzeumi és levéltári rendszerekre vonatkoznak. A könyvtári rendszereket készítő informatikai cégeknek pedig a szótári, szemantikai részrendszerei általában csak a tárgyszavak mutatóinak megjelenésében, továbbá a szabványos, elsősorban MARC-típusú egységesített besorolási adatok kezelésében vannak tapasztalataik. A kettő nem mindig találkozik egy tervező cégben.

Névterekben a nevekhez nemcsak sokkal több metaadat és gazdag szemantikai relációszerkezet kapcsolódik, hanem a kronologikus adatok kezelése (az időbeli rétegződés) és a webes kapcsolódások gazdagsága is jellemző.

Az informatikai tervezőknek elsősorban profitorientált megrendelővel van tapasztalatuk, akik üzleti célú rendszereket igényelnek. A névtér megrendelői ezzel szemben nem profitorientált üzleti szereplők. Esetükben nincs megszokott különbség az üzleti megrendelő, a szolgáltató-működtető, és a felhasználó között. A „valódi” megrendelő egyben maga a legmagasabb szintű felhasználó, azaz az intézmény, melyben ez a névtér működni fog, és amelyet ez az intézmény professzionálisan működtetni is akar.

A magyar nemzeti névtér esetében a megrendelő, egyben a hozzáértő felhasználó az OSZK és a névtér kialakításában vele társult intézmények, mint például a PIM és a FÖMI.

Ebből viszont az is következik, hogy a megrendelési oldalon nincs laikus fél. A megrendelő mérvadó szakmai (azaz nem pusztán „üzleti” befektető) partner, aki nem afféle külső, fogyasztói használója lesz a névtérnek, hanem maga a professzionális használó is, aki például üzemeltet is, és egyben érdemi szakmai színvonalon meghatározza a rendszer adattartalmait.

Az informatikai tervezésben a korszerű agilis termékvíziók, a Scrum-módszer [12] fókuszában a megrendelő üzleti ügyfél elvárásait megkülönböztetik a felhasználói szükségletektől. Névtér esetében ez a megkülönböztetés értelmezhetetlen. Nincs érdemi eltérés ügyfél és felhasználó között. Az elvárások egybeesnek.

Sokat mondó példa, hogy a Scrum-módszer szlengjében a végfelhasználót kezdetben afféle „csirkének” nevezték. Ettől a lebecsülő szóhaszná-

lattól a későbbi módszertanokban, a reagálásokon okulva lemondtak, de azért a kezdet mégis csak sokatmondó, ami a szemléletet illeti. Az ügyfél „csirkét” érteni azt jelenti, hogy a végfelhasználón nem professzionális és egyenrangú fejlesztő partnert értenek. Azért említem ezt meg, mert fontos lélektani kérdés, és a projekt sikerét befolyásolja, hogy reális képe legyen az informatikai fejlesztőknek arról, ki is valójában – legalábbis névtér esetében – az érdemi megrendelőjük. Semmiképpen sem laikus „csirke”.

Mindennek következtében a megrendelő többnyire nem elsősorban informatikai szakembereinek azt is el kell érniük a tervezési munka során, hogy őket az informatikai tervezők egyenrangú tervező partnernek is elfogadják.

Mindez érthető. Az informatikai tervezőnek egy teljesen szokatlan területet kell megismernie.

## Néhány jellegzetes nehézség

### Magyar nyelvű felületek

A programok tervezői megszokták, hogy magában a rendszertervekben kizárólag informatikai szövegből használják például az adatok megnevezését, és ezeket a legszívesebben a nemzetközileg megszokott angol kifejezésekkel igyekszik megnevezni. Ezek az angol szavak sokszor szleng kifejezések.

Egy nemzeti névtérben – és általában minden, a kulturális örökséget kezelő adatbázisban – elvárható, hogy a magyar szakmai és köznyelvet használják. (Ahogy ez kötelező minden értelmező nagyszótárra.) Az informatikai tervező például képes az angol eredetű junior és senior lektor, továbbá superadmin kifejezést használni a rendszertervben. Ez már önmagában is elég hajmeresztő, hiszen a kifejezés ebben a formájában se nem angol (mert akkor junior lectore, vagy superadmin lenne), se nem magyar. A megfelelő magyar kifejezések, melyeknek a felületen meg kell jelenniük, a szerkesztő, névszerkesztő, rögzítő szerkesztő, beosztott szerkesztő, valamint vezető szerkesztő, irányító szerkesztő, csúcsszerkesztő, ellenőrző szerkesztő, ellenőr stb. Az informatikai szóhasználatban szokás admin-rendszer pedig vagy szerkesztő rendszer, vagy adminisztrációs rendszer. Ha a megrendelő nem vigyáz, első lépésben a felületeken az informatikai szleng (pl. admin rendszer) jelenik meg, és utólag jobb eset-

ben barátságosan elfogadják, hogy a kifejezéseket módosítsák.

Mindebből az is következik, hogy a megrendelő végfelhasználónak már a tervezés korai dokumentumaiban, például az előtervekben célszerű figyelnie a terminológiára, és jelezni kell az igényeket, hogy ne később kelljen ezen módosítani, netán vitázni.

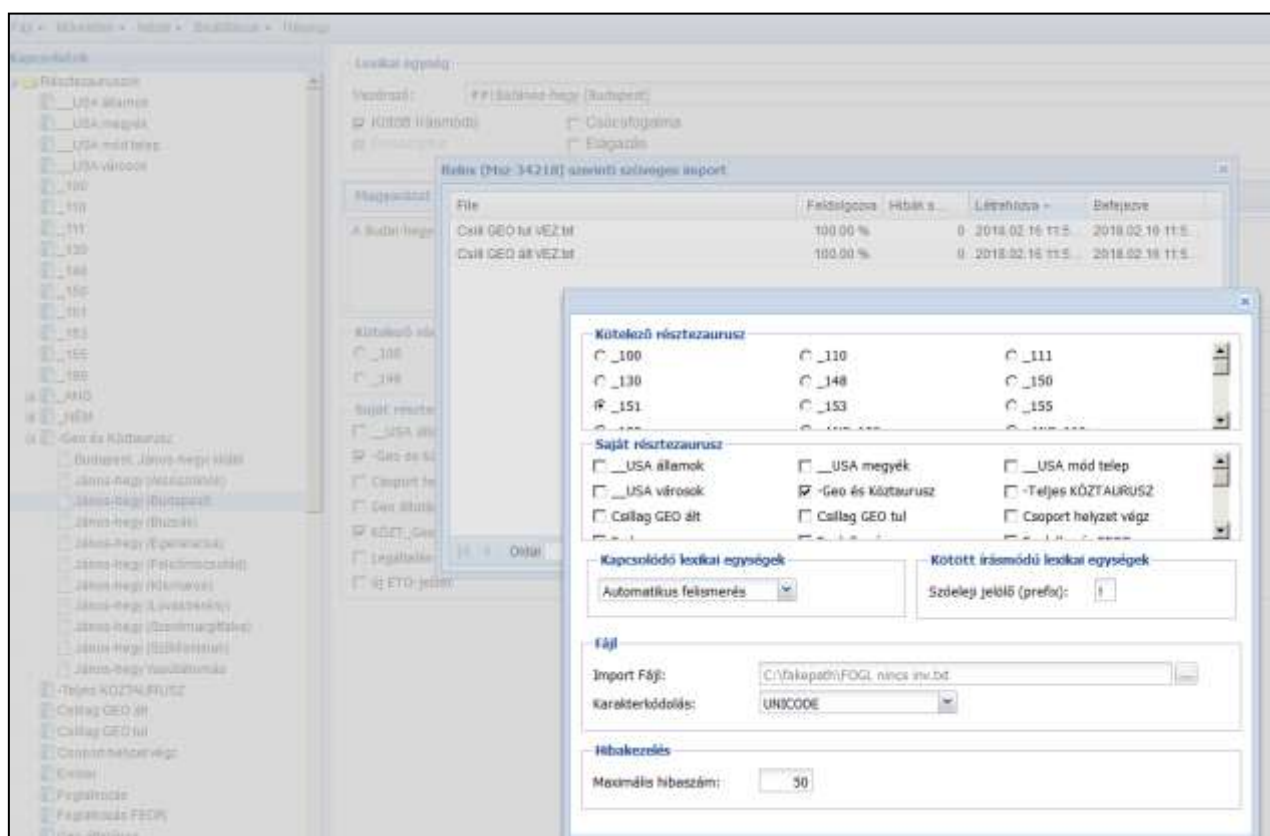
A projekttel kapcsolatos minden olyan dokumentációban (így az előtervekben és a követelményelemzésekben is), melyek nem szorosan programozói dokumentumok, a kezdetektől elvárható, hogy a megrendelő (azaz egyben a professzionális felhasználó) terminológiai igényei szerint használják a kifejezéseket (például nem „front end”, hanem végoldal, nem „analytics”, hanem statisztika, vagy forgalmi elemzés, nem „inventory”, hanem naplózás). De még az olyan esetekben is, mint „file server”, „portal server” stb. célszerű ragaszkodni a szakszerű – sokszor eleve szabványosított – magyar kifejezésekhez, mint állománykiszolgáló vagy

fájlserver, portálszerver, gyorsító kiszolgáló vagy szerver stb.

Előforduló hivatkozás, hogy az informatikus ezzel szemben csak az angol kifejezésekhez szokott, vagy netán csak azokat képes értelmezni. Ebben az esetben a dokumentációban kettős szóhasználathoz kell ragaszkodni: a megkövetelt magyar mellett zárójelben szerepelhet az angol kifejezés.

A névtéri szakterület ismeretének hiányában előfordulnak légből kapott kifejezések, amikor például az import állományra a „lexikon” kifejezést, a „relációk szerkesztése” helyett pedig a terpeszkedő „névtér-kapcsolatrendszer előállítás” kifejezést akarják használni.

Külön probléma a belső, szerkesztői felületeken és a végfelhasználók különféle megjelenítési felületein (az ablakokban) megjelenítendő szóhasználat. Az informatikában előfordul, hogy ezekre a „nézet-sablon” kifejezést használják, ami felhasználói szemszögből teljesen idegen; ehelyett a megjelenítési felület a helyes kifejezés, illetve ablak, részablak (1. ábra).



1. ábra Kezelőrendszer rögzített részablakai a háttérben, és előre ugró részablakai az előtérben

A szerkesztői felületek mindig a professzionális szerkesztés igényeinek kell, hogy megfeleljenek. Itt laikus felhasználói hozzáférésről nem lehet szó, a felületen megjelenő kifejezéseknek összhangban kell lenniük a korrekt szakmai nyelvhasználattal, és ha rövidítések szükségesek, akkor azokat ezekből kell levezetni.

Például nem „authority” állomány vagy adat, hanem egységesített vagy besorolási vagy ellenőrzött vagy szabványos adat.

A megjelenítési felületek egyik változata a professzionális felhasználói igényeknek kell, hogy megfeleljen, a másik pedig a laikus felhasználókat szolgálja ki. Ez utóbbi esetben célszerű választhatóvá (opcionálissá) tenni, hogy kutatói felhasználóról van-e szó, vagy érdeklődő felhasználóról. Egy névtér ugyanis előre láthatóan – többek között – az értelmező szótárak feladatait is elláthatja idővel, és akkor számítani kell arra az egyszerű igényre, hogy valaki pusztán tájékozódni akar kifejezések jelentéséről, összefüggéseiről, továbbá a rájuk vonatkozó, ugrópontokon keresztül is elérhető webes forrásokról (népszerű lexikonokról, térképi helyekről stb.), melyekben a keresett név előfordulhat.

Mindez emlékeztet a Wikipédiára, mégse ilyesmiről van szó. A névtér nem tartalmaz részletes, hosszú szöveges információkat az adott kifejezéssel kapcsolatban, hanem csak viszonylag rövid szöveges megjegyzéseket. Ezzel szemben sokkal gazdagabb egy-egy név metaadatokban és ugrópontokban.

Ami a névhasználatot illeti, nem terminológiai szórszálhasogatásról van szó. A magyar nemzeti névtér a magyar nyelv neveinek és szavainak a távlatilag teljességre törekvő állománya. Távlatilag, amikor majd évtizedek múlva megközelítőleg teljesen (de soha sem egészen teljesen) feltöltődik, a weben átveszi a nemzeti nyelv értelmező nagyszótárának szerepét. Ez magát a kezelőrendszert is nyelvileg kötelezi.

### Megjelenítési nehézségek

Névtérre – ha nagyjából feltöltődnek – a rendkívül nagy névállomány jellemző. Ez elérheti a több száz ezret, személy- és testületi nevek esetén akár milliós nagyságrendet is. Ami a közneveket illeti, a nagy értelmező szótárakban csak az egyszerű (pl. hang) és az összetett (pl. hangmagasság) szavakat tartalmazzák (az utóbbiak közül is csak a köznyelvben elterjedten használt kifejezéseket); az

összetett szakkifejezéseket (pl. adatbázis, adatbázis-kezelő) és a különírt szavakból álló szókapcsolatokat (pl. abszolút hangmagasság, adatbázis-kezelő rendszer, agresszió elleni védelem) nem.

### A több jelentésű nevek megjelenítése

A hatalmas állomány azzal is jár, hogy rendkívül megnő a több jelentésű szavak száma. Köznevek esetén ezek lehetnek tartalmilag egymással összefüggő polisziemiák (pl. körte) vagy nem összefüggő homonimák (pl. ár). Személy- és testületi nevek esetén valójában mindig homonimákról van szó (akár tízezres nagyságrendű is lehet például a „Kis” családnév, melynek hordozói semmiféle összefüggésben nincsenek se egymással, se testük nagyságával). Földrajzi nevek esetében részben homonimákról van szó (pl. János-hegy), részben polisziemiákról (pl. Öreg-hegy – az ilyen nevű hegyek általában valamilyen értelemben nagyot, jelentőset képviselnek a környezetükben).

Az azonos **személynevek** megkülönböztetésére nem hogy a keresztnév, a kronologikus adatok (születési és halálozási hely és idő), a foglalkozás nem elég, de előfordulhat, hogy még az anya neve sem elegendő (mert ikrek esetében lehetséges, hogy nem közös apától vannak). Ez azt eredményezi, hogy az adott családnévhez egész sor – előre nem meghatározható – metaadatot kell kapcsolni, amikor adott családnévvel keresve az azonos nevű találatokat megkülönböztethetően kell megjeleníteni. Például:

Nagy Ferenc (Bisse, 1903–Herndon, 1979), politikus

Nagy Ferenc (Győr, 1903–Budapest, 1979), politikus

Nagy Ferenc (1903–1979), fényképész, anyja: Stein Mária

Nagy Ferenc (1903–1979), fényképész, anyja: Elsbeth ?

Nagy Ferenc (1903–1979), fényképész, anyja: Elsbeth ?, apja Kovács János

Nagy Ferenc (1903–1979), fényképész, anyja: Elsbeth ?, apja Alister Frank

A nem európai, például kínai nevek kérdését itt meg se említjük – noha magyar névtérben ilyenek is lehetségesek.

Mindebből egyrészt következik, hogy

– a névhordozót természetes nyelven azonosító, összekapcsolt metaadatok összessége egy névtípuson belül is rendkívül különböző lehet (például az ókori nevek esetében elég akár

egyetlen adat, mások esetében a születési időponttól, helytől kezdve mindkét szülő nevén keresztül a foglalkozáson át nagyon sok metaadat szükséges, hogy a hasonlótól meg lehessen különböztetni;

- a metaadatokkal való azonosítás nem formalizálható, hanem intellektuális szerkesztést igényel;

Az azonos földrajzi tulajdonnevek esetében a megkülönböztetés nem különféle metaadatokat igényel, hanem azonos típusú földrajzi tulajdonnevek partitív (rész-egész reláción alapuló) tranzitív láncolatát. A Geonamesben ilyen láncokkal azonosítanak (2. ábra):


The screenshot shows the Geonames search interface with the search term "Óreg-hegy" and a dropdown menu set to "all countries". The search results are displayed in a table with 48 records found for "Óreg-hegy". The table columns are Name, Country, Feature class, Latitude, and Longitude. The results list various locations across Hungary, including Veszprém, Pest, Tolna, Heves, and Komárom-Esztergom, with their respective coordinates.

Name	Country	Feature class	Latitude	Longitude
1 Óreg-Futónai Futónai-Köves-hegy, Futónai-Köves-hegy, Óreg-Futónai, Óreg-Futónai-hegy, Óreg-Futónai, Óreg-Futónai-hegy	Hungary, Veszprém	mountain	N 47° 13' 31"	E 18° 1' 58"
2 Óreg-Kereszt Óreg-Kereszt, Óreg-Kereszt-hegy, Óreg-Kereszt, Óreg-Kereszt-hegy	Hungary, Pest	hill	N 47° 28' 11"	E 19° 29' 11"
3 Magas Hegy Magas-hegy, Óreg-hegy, Óreg-hegy	Hungary, Tolna	hill	N 46° 30' 0"	E 18° 36' 0"
4 Zsongo-tető Óreg-hegy, Óreg-hegy, Zsongo-hegy, Zsongo-tető, Zsongo-hegy, Zsongo-tető, Óreg-hegy, Óreg-hegy	Hungary, Veszprém	mountain	N 47° 21' 22"	E 17° 50' 48"
5 Óreg-hegy	Hungary, Pest	hill	N 47° 55' 21"	E 18° 47' 31"
6 Óreg-hegy	Hungary, Heves	hill	N 47° 46' 0"	E 19° 58' 0"
7 Óreg-hegy	Hungary, Heves	hill	N 47° 45' 26"	E 19° 50' 47"
8 Óreg-hegy	Hungary, Pest	hill	N 47° 44' 48"	E 19° 22' 34"
9 Óreg-hegy	Hungary, Heves	hill	N 47° 42' 47"	E 19° 42' 25"
10 Óreg-hegy	Hungary, Somogy	hill	N 46° 49' 46"	E 17° 55' 0"
11 Óreg-hegy	Hungary, Somogy	hill	N 46° 47' 46"	E 17° 47' 33"
12 Óreg-hegy	Hungary, Somogy	hill	N 46° 42' 38"	E 18° 1' 52"
13 Óreg-hegy	Hungary, Tolna	hill	N 46° 42' 2"	E 18° 28' 51"
14 Óreg-hegy	Hungary, Somogy	hill	N 46° 36' 0"	E 17° 24' 0"
15 Óreg-hegy	Hungary, Tolna	hill	N 46° 31' 56"	E 18° 12' 37"
16 Óreg-hegy	Hungary, Somogy	hill	N 46° 29' 29"	E 17° 24' 11"
17 Óreg-hegy	Hungary, Tolna	hill	N 46° 23' 11"	E 18° 28' 23"
18 Óreg-hegy	Hungary, Somogy	hill	N 46° 34' 31"	E 17° 22' 38"
19 Óreg-hegy	Hungary, Tolna	hill	N 46° 15' 50"	E 18° 3' 0"
20 Óreg-hegy	Hungary, Budapest	hill	N 46° 7' 0"	E 18° 46' 0"
21 Óreg-hegy	Hungary, Heves	hill	N 47° 50' 0"	E 20° 28' 0"
22 Óreg-szőlős Óreg-szőlős-hegy, Óreg-szőlős-hegy, Óreg-szőlős, Óreg-szőlős-hegy, Óreg-szőlős-hegy, Óreg-szőlős	Hungary, Komárom-Esztergom	hill	N 47° 20' 0"	E 18° 3' 0"
23 Óreg-szállás-hegy	Hungary, Nógrád	hill	N 48° 0' 13"	E 19° 2' 4"
24 Óreg-Petra-hegy	Hungary, Tolna	hill	N 46° 31' 16"	E 18° 39' 51"
25 Óreg-hegy	Hungary, Nógrád	hill	N 47° 37' 14"	E 19° 18' 31"
26 Óreg-hegy	Hungary, Nógrád	hill	N 47° 51' 5"	E 19° 27' 49"
27 Óreg-hegy	Hungary, Veszprém	hill	N 47° 19' 0"	E 17° 30' 0"
28 Óreg-hegy	Hungary, Veszprém	hill	N 47° 7' 8"	E 17° 38' 11"
29 Óreg-hegy	Hungary, Veszprém	hill	N 46° 51' 6"	E 17° 38' 38"

2. ábra Azonos nevek azonosítása a Geonames rendszerben


Jól látható, hogy elégtelen a megyék szerinti megkülönböztetés, mivel egyetlen megyén, de még járáson belül is lehetnek azonos földrajzi hevek (még az is előfordulhat, hogy egyetlen település

határában is létezik két azonos név). Maguk az államok már nincsenek összekapcsolva a földrészekkel.

Click the  icon to view the hierarchy.

[Semantic View \(JSON, JSONLD, RDF, N3/Turtle, N-Triples\)](#)








**ID: 1105063** **Record Type: physical**

 **János-hegy (mountain)**

*Coordinates:*  
 Lat: 47 31 00 N *degrees minutes*    Lat: 47.5167 *decimal degrees*  
 Long: 018 58 00 E *degrees minutes*    Long: 18.9667 *decimal degrees*

**Names:**  
 János-hegy (**preferred,C,V**)

**Hierarchical Position:**

-  World (facet)
-  .... Europe (continent) (P)
-  ..... Hungary (nation) (P)
-  ..... Budapest (special city) (P)
-  ..... Budapest (inhabited place) (P)
-  ..... Buda (neighborhood) (P)
-  ..... János-hegy (mountain) (P)

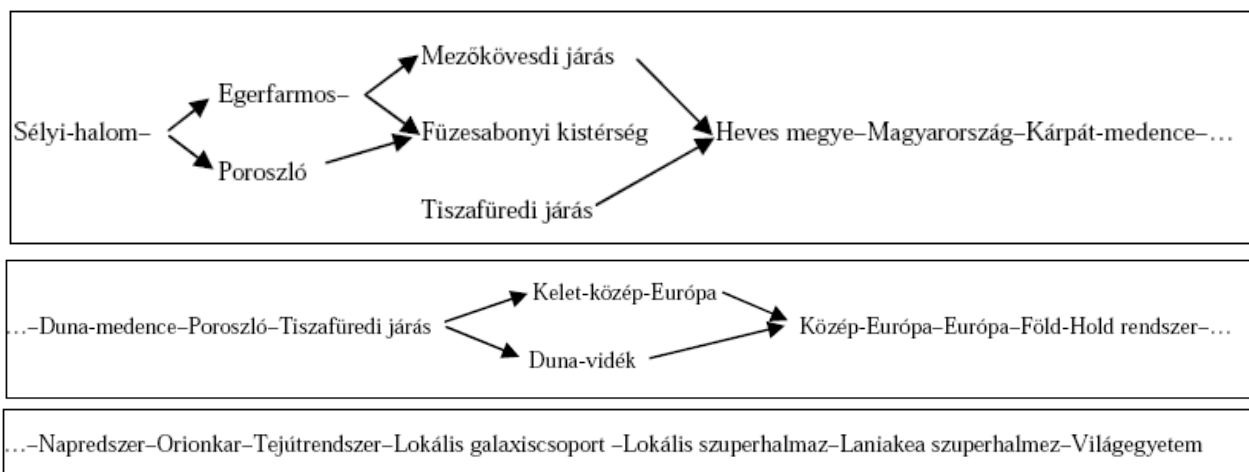
**Place Types:**  
 mountain (**preferred, C**)

3. ábra A világegyetemig vezető azonosító partitív hierarchialánc

A Getty földrajzi tezaszban már a világegyetemig vezet a partitív hierarchialánc (3. ábra).

A nemzeti névtérben – hiszen „minden névnek, amely rögzíthető” – nagyságrendekkel több a név, mint elődeiben, a tezaszokban, ennél fogva a

közbenső partitív fölrendeltek száma is nagyobb. Más szóval a teljes lánc rendkívül hosszú. A földrajzi névtéri részben nem csupán a megye, az ország, a földrész szerepel, hanem minden ennél kisebb és nagyobb képződmény. Például:



A Sélyi-halom (Saj-halom, Sej-halom) két település határában fekszik (a valóságban három, de egyszerűsítettünk), és a két település a legújabb közigazgatási beosztás szerint más-más járáshoz tartozik. Mivel a Duna-medence egyszerre része Kelet-közép-Európának és a Duna-vidéknek (mely

utóbbi viszont nem része Kelet-közép-Európának), ezért a hierarchialánc megint szétágazik és öszszesemegy. A Földön kívül pedig még egész sor csillagászati képződmény létezik, mely a világminőségig bezárólag egyre átfogóbb egész.

Ilyen hosszú azonosító láncok nem szükségesek. Ehhez viszont a rendszerbe be kell építeni a megjelenítendő lánc hosszúságára vonatkozó feltételeket. Például:

*Ha a földrajzi tulajdonnév nem a „folyóvíz” tranzitív generikus alárendeltje és a név a „Föld” közvetlen vagy tranzitív része, akkor az azonos nevek megkülönböztetésére a partitív hierarchialáncban csak a „Föld” névig tartó láncot kell megjeleníteni ÉS ha partitív elágazás (polihierarchia) van, az azonos földrajzi tulajdonnevek közül a betűrendben az elsőt kell figyelembe venni.*

Különleges eset a folyóvizéké (erek, patakok, folyók, folyamok; csatornák). Ezeket nem lehet a nagyobb területrésszel azonosítani, mert többnyire sok területen folynak át. Igaz ugyan, hogy a kisvízfolyások egy része csak egy megyén belül vezet, de egy megyén belül sokszor lehet azonos neve egy pataknak.

Ezért az azonos nevű folyóvizeket a folytatásukkal – vagy azzal az állóvízzel, amelybe torkollik – kell azonosítani, de még így is előfordulhat, hogy egy nagyobb vízfolyásba két azonos nevű kisvízfolyás ömlik, többnyire persze nem egymás közelében. A vízfolyást jelentő homonimák és poliszemiák esetében tehát szükség van a gyakori esetek formális szabályozására:

*Ha a földrajzi név generikus hierarchialáncának valamelyik tranzitív eleme „folyóvíz”, akkor az azonos nevű folyóvizeket a folytatásukat képviselő folyóvíz nevével kell megkülönböztetni.*

Valójában ez a szabály nem minden esetben érvényesíthető. Vannak olyan folyóvizek, melyek nem állóvízbe torkolnak (melybe beleértjük a mocsarakat és lápokokat is), hanem száraz területen, sivatagban enyésznek el. Ilyen például Dél-Afrikában az Okavango. Az elmúlt évszázadban Közép-Európában számos ér és patak alsó folyása megszűnt. Számítani kell arra, hogy ezek elenyészési helyét azonosítani kell, és az nem állóvíz. Ezek automatikus kiszűrése csak akkor lehetséges, ha az összes víztestet (folyó- és állóvizeket) jelentő földrajzi nevet résznévtér is tartalmazza. Ekkor automatikusan megjeleníthetők azok a folyóvizek, melyeknek nincs folytatás-folyóvízük.

A vádik (aszóvölgyek) többnyire sivatagban érnek véget, noha előfordul, hogy folyóvíz alsó szakasza vádi.

Az azonos nevű **vádik** esetében az automatikus azonosítás alapja a sivatag, melyben végződnek, noha előfordulhat több ugyanabban a sivatagban végződő azonos nevű vádi. A magyar nemzeti névtérben ezek az esetek ritkák, noha egy igazán jól szervezett névtérkezelőben ez is algoritmizálható.

Az **azonos köznevek** esetében a megkülönböztetés ugyancsak nem különféle metaadatokat igényel, hanem – ellentétben a földrajzi tulajdonnevek esetével – általában csak egyetlen generikus fölérendeltet (nem-fogalmat). Például:

orgona–olajfaféle

orgona–levegőbefúvósos billentyűs hangszer

Látható, hogy itt is érvényesül az a következmény, hogy „minden névnek, amely rögzíthető”, mert az orgona–növény

orgona–hangszer

közvetlenül többnyire nem létezik. Túl sok közbeneső, egyre általánosabb jelentésű név esik a közgondolkodás szerinti egyszerű értelmező kifejezés elé. Ezen csak manuális beavatkozással lehet segíteni: ún. hátravetett értelmező alkalmazása szükséges, amely metaadata lesz a névnek:

orgona (növény)

orgona (hangszer)

Ezért biztosítani kell, hogy a fenti – és valójában az ebben a fejezetben tárgyalt összes – következtetési szabályt adott esetekben felül lehessen írni (azaz a szabályt lehessen adott földrajzi tulajdonnév esetében kiiktatni).

### **A közbenső fölérendeltek szükségessége**

A névteret egyaránt strukturálják szemantikailag a generikus és a partitív hierarchialáncok (együtt más relációtípusokkal). Ennek ellenére nem lehet például megjeleníteni a Pest megyei kistélepedéseket Pest megye partitív alárendeltjeiként, vagy az európai államokat Európa részeiként. Ennek is a névtér rendkívüli névgazdagsága az oka. Pest megye részei a járások, azok részei a települések, és csak ezek részei a kistélepedések. Európa több földrajzi és politikai részre tagolódik, ezeknek további részei vannak, és az államok csak ezek után következnek.

Természetesen kialakíthatók olyan keresési műveletek, melyekben következtetési algoritmusok épülnek be. Például a Szentendrei járás külterületi településrészei úgy találhatók meg a „Szentendrei járás” ÉS „külterületi településrész” kérdéssel,



hogy egyben a következő utasítás is megadható: ha a Szentendrei járásnak vannak települései, akkor jelenjenek meg találatként e települések külterületi lakott helyei.

Ez így rendben is volna, csak éppen szinte senki se fog ilyen keresést használni. Az átlag felhasználónak (akit érdekelhet, hogy például milyen kis-települései vannak Pest megyének) még az ÉS és a VAGY kapcsolatokkal végzett keresés is „túl bonyolult”.

Ezért a névterekben a keresettebb esetekre ún. közbelső fölérendelteteket alakítanak ki. Ezek megválasztása komoly tájékozottságot, szerkesztői hozzáértést igényel. Fel kell ismerni, milyen közbelső fölérendeltetekre van igény. Ezek végleges kialakulása akár évtizedeket is igénybe vehet. Más szóval, a névtér állandó tökéletesítésre szorul. Például:

Pest megyei kistelepülés  
Békés megyei csatorna  
Európai Unió állama (és az Európai Uniónak csak az „Európai Unió állama” lesz a része, és ennek lesznek fajtái az egyes államok).  
Magyarországi természetvédelmi terület  
Moldvai magyar település  
Barlang–Európa barlang–Kárpátmedencei barlang–Magyarországi barlang  
délvidéki vár

Mivel magyar nemzeti névtérről van szó, ebbe számos külföldi földrajzi név is bekerül, ha bármilyen módon kapcsolódik a magyar kultúrához. Például magyar szerző ír róla, külföldi szerző ír róla, de Magyarországon adták ki a művét; összefoglalóan: ha hungarikum. De bekerülhetnek a névtérbe pusztán kiemelkedő jelentőségük következtében is külföldi nevek. A külföldi földrajzi nevek gyakorisága teljesen véletlenszerű. Ezért nincs értelme például a „franciaországi barlang” közbelső generikus fölérendeltnek, elég az „európai barlang”, mivel jelenleg csak az Altamira-barlang szerepel a névtérben. Ha az ilyen nevek – például az európai barlangok – száma növekszik, ennek megfelelően kell a szerkesztői munka során a közbelső fölérendeltet számát is bővíteni.

Mindebből két következtetés adódik.

- Egyrészt a felsorolt különleges esetek informatikailag semmiféle különleges, nehezen megoldható feladatot nem jelentenek. Csak éppen a kereskedelmi vagy közigazgatási, egészségügyi stb. célú programrendszerek megoldandó feladataihoz képest nagyon speciálisak. Már a

## Ungváry R.: Névtérkezelő készítésének nehézségei

rendszertervezés kezdetén tekintettel kell lenni ezekre a nagyon különleges feladatokra.

- Egy induló nemzeti névtér már eleve rendkívül nagy névállományt tartalmaz, de ez nagyságrendekkel kisebb, mint amekkora idővel, a feltöltődés során lesz. Ezért nemzeti névtér készítésének semmiféle értelme nincs, ha nincs biztosítva az intézményes, folyamatos szerkesztői munka.

## Értelmezési nehézségek

Jellegzetes probléma, hogy a névtér tartalma számos afféle „szemantikai meglepetést” is okoz az informatikai tervezőknek. Bizonyos tartalmi-szemantikai megoldások szükségessége ezért hosszabb magyarázatokat, ismételt fordulókat igényel.

## A nevek szenzitivitása

Az egyik legalapvetőbb kezdeti nehézség, hogy ugyan nevek, többek között személy- és testületi nevek kezeléséről és tárolásáról van szó, ami megtévesztő. Hagyományosan ugyanis az ilyen nevek szenzitív (azaz kényes, és ezért szigorúan védett személyes) adatok. A személy- és testületi nevek vonatkozó informatikai tapasztalatok a közigazgatásban, a banki és egészségügyi rendszerekben szükséges adatkezelésekre vonatkoznak, melyek rendkívül szigorúak. Ezekben a rendszerekben például a nevek felvételére, módosítására és törlésére nagyon kötött eljárásokat programoznak. Minden lépés csak az előző lépés ellenőrzött visszaigazolása után tehető meg. Ha például új nevet vesznek föl, a név nem jelenhet meg a felhasználói felületen mindaddig, amíg nem ellenőrizték, és erről a névszerkesztő, aki a nevet fölvette, magán a rendszeren belül értesítést nem kap.

Erre névtérben semmi szükség nincsen. Ha egy nevet a névszerkesztő fölvesz, annak automatikusan függő (még nem véglegesített) státuszba kell kerülnie, és maga a név nyomban megjelenhet minden végfelhasználói felületen, természetesen együtt azzal, hogy jelezve van, ez a név még nincsen véglegesítve. A felvett neveknek ugyanis nagyon különböző, és eleve nyilvános státusa lehet, például ideiglenes, emlékeztető, teljes, előzetes, mely állapotoknak különböző készültségi szintek, illetve szerepek felelnek meg.

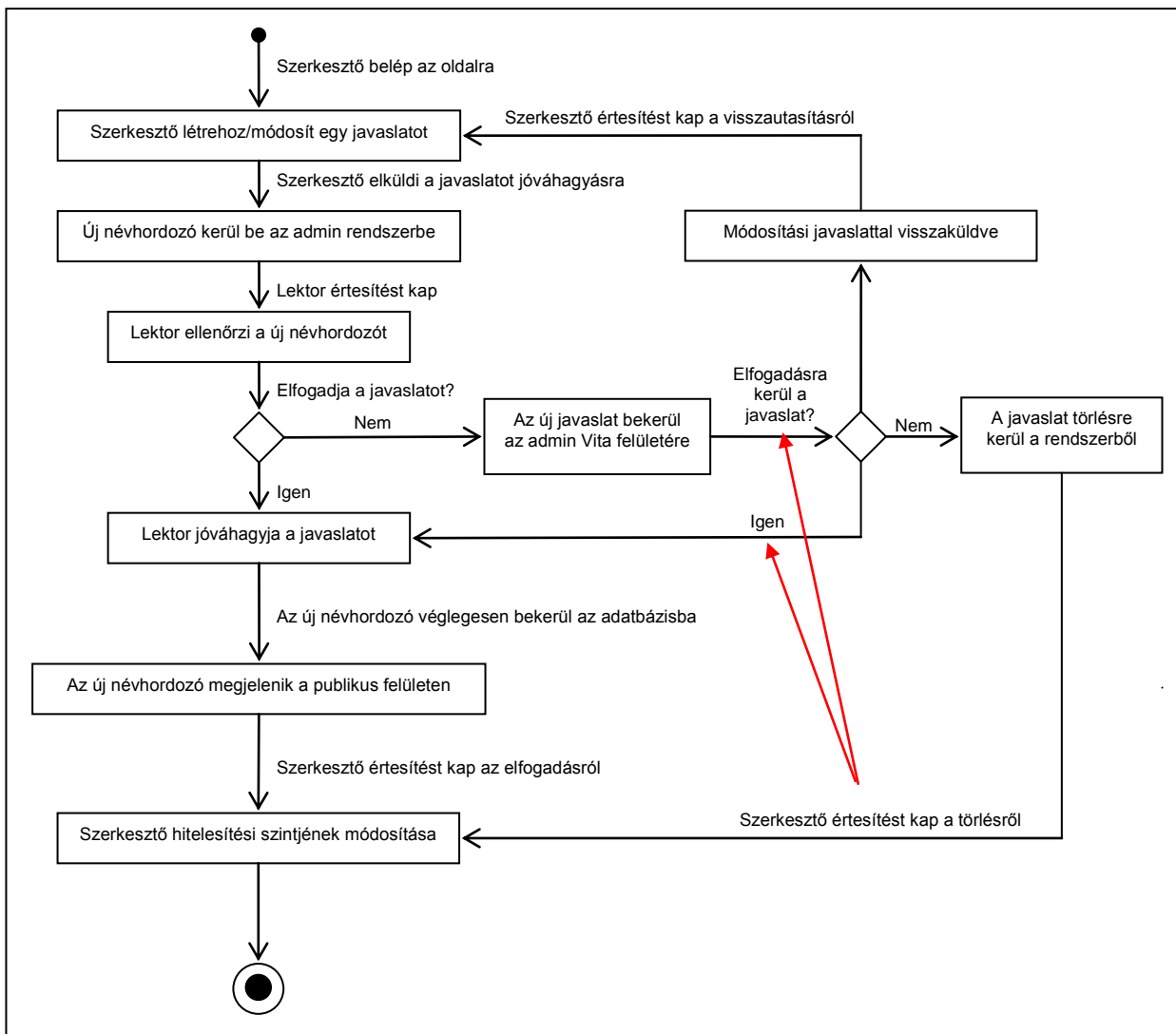
Ezért már a tervezés kezdetén figyelni kell arra, hogy az informatikai rendszerben ne érvényesüljön

mechanikusan (az egyébként nagyon hasznos) szigorú államigazgatási szenzitív adatkezelés. A névtér adatai ugyanis semmiben sem különböznek egy értelmező szótár, vagy egy történelmi névtár neveitől. Kezelésükhöz nem a szenzitivitás, hanem kizárólag a tartalmi (történelmi, irodalmi stb.) hozzáértés szükséges. Az előbbit programozhatóan ki kell kényszeríteni (ami például az egészségügyben

törvényi kötelesség is), az utóbbi azonban lehet kizárólag kezelőrendszeren kívüli tudás.

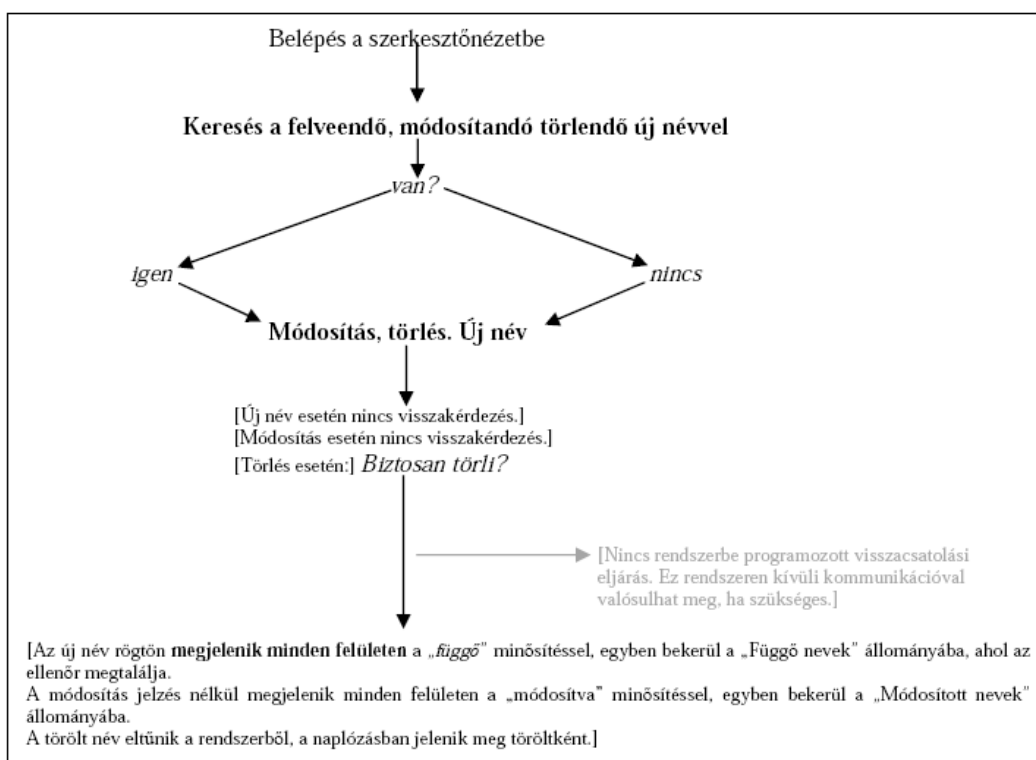
Összegezve: névterekben érdemben nincs szükség szigorú, programhoz kötött eljárás alapú kezelésre.

A 4. ábrában egy szigorúan az informatikai rendszeren belül – valójában feleslegesen – megkötött eljárás folyamatábrája látható:



4. ábra Új név felvételének feleslegesen kötött felvételi eljárása. A piros nyillakkal jelzet visszacsatolásra az informatikai rendszeren belül eljárás-kötötten nincs szükség. Elég az informatikai rendszeren kívüli kommunikáció. Az új név felvételének pillanatától kezdve látható lesz a rendszer megjelenítési felületein, feltüntetve, hogy mi a feldolgozottsági állapota.

Az 5. ábrán a névtérben szükséges, ergonómiaailag egyszerűbb folyamatára látható:



5. ábra Új név felvétel névtérnek megfelelő egyszerű módon

Ha a névtérhez külső, például katalogizáló, könyvtári, levéltári, múzeumi vagy egyéb nyilvántartó rendszerek csatlakoznak, akkor természetesen be kell építeni megfelelően programozott visszacsatolásokat, hogy e rendszerekben a változások megjelenjenek. De ez névtéren kívüli kérdés.

### Típus és rész

Értelmezési nehézséget okozhat a típus és a rész megkülönböztetése. Az informatikai és egyéb szakmai szóhasználatban a részrendszert szokás alrendszernek nevezni, holott itt a rendszer részéről van szó, és nem a rendszer valamiféle altípusáról. Más szóval különbséget kell tenni a partitív és a generikus összefüggés között.

A névtérek altípusai a személy-, testületi név, cím-, földrajzi- és köznévtér. Ezek nem résznévtérek. Ezzel szemben egy altípuson belül például a személynévtér esetében a „politikusok” a névtér része, és nem altípusa, a „határon túli magyar település” pedig a földrajzi névtér egyik résznévtére. A tervezés szempontjából ez nem terminológiai kérdés, mert egy speciális (azaz alnévtér) és egy résznévtér kezelése két teljesen különböző feladat. Az előbbiben a rendszer maga is módosulhat, specializálódhat, az utóbbiban a kérdés nem egyéb, mint hogy a résznévtérbe kerülő nevek kapnak egy speciális minősítőt (metaadatot), amelynek alapján állományuk az egésztől elkülönítve kezelhető.

Névtérfajták	Névtérterek példái
Személynévtér	Határon túli magyar személyiségek
Testületi névtér	Civil szervezetek; Rendőrségek; Magyar külképviseletek
Művek címeinek névtére	Szent szövegek címei; Krimicímek
Földrajzi névtér	Moldvai magyar helynevek; Csillagászati képződmények; Ókori helyek
Köznévtér	Foglalkozások; Betegsétünetek; Ismerettünetek; Dokumentumtípusok

**MARC21, ISAAR, CIDOC adatok és a tőlük független kezelés**

Mivel a névtéri nevek egyik professzionális gyakorlati felhasználói a könyvtárak, levéltárak és múzeumok, azaz a kulturális örökségi intézmények, és mert ezekben az intézményekben egyre inkább a MARC-típusú formátumokban cserélik az adatokat, ezért minden olyan adat esetén, mely egyben a MARC-formátumokban is szerepel, biztosítani kell a MARC szerinti kezelést is [9], [10], [24].

Ugyanakkor lehetővé kell tenni a levéltári és múzeumi igényeknek megfelelő szabványos adatok kezelését is [1], [2], [5.]

Attól azonban, hogy ez szükséges, ezeknek az adatoknak szinte mindegyike a MARC-formátu-

moktól és a levéltári és múzeumi szabványoktól függetlenül is létezik és fontos.

Például egy Magyarországon és Izraelben egyaránt írói nevén (Uri Asaf) publikáló szerző eredeti nevének (Asztalos Iván) elsődleges leírása latin, az alternatív leírása héber (אורי אסף).

<p><b>Asaf, Uri</b> eredeti neve Asztalos Iván héber neve אורי אסף</p>
--

Mindhárom név esetében informatikailag kezelhető módon kell metaadatokkal minősíteni a név típusát és írásrendszerét. A MARC21 szerint az alábbi módon kell ezt minősíteni:

066 ## \$c(2	<i>Az elválasztó karakter a karakterláncot egy bájt per karakterként azonosító hexa 29 (ASCII szerint. A nyitó kerek zárójel a karakterláncot egy bájt per karakterként azonosító hexa 29; a záró 2 karakter hexa 32; az így azonosított G1 karakterkészlet a héber ábécé.</i>
100 1# \$aAsaf, Uri	
400 1# \$aAsztalos Iván	
880 1# \$6100-2/(2)r \$a אורי אסף	<i>A 100-as szerzőnév mezőben a (2 adattal azonosított héber név szerepel).</i>

A névtérben a bevitel ugyan MARC21-től teljes független módon valósítható meg:

<i>Írásrendszer (a 100-as mezőbe bevitt második név esetén): héber</i>
--

Ha azt akarjuk, hogy a nevet a könyvtári, levéltári és múzeumi világban is lehessen fölhasználni, akkor egyben biztosítani kell, hogy létrejöjjön a MARC21 szerinti minősítés is. Itt a probléma az, hogy ha a MARC-hívójelekkel végezzük a bevitt adat azonosítását, akkor ezek ismeretében a bevitt felület egyszerű: egyetlen bevitt mezőben a különféle MARC-hívójelekkel ellátva végezhető a rögzítés. Ehhez azonban jól kell ismerni a MARC21-et.

Ha azonban ez felhasználóbarát módon történik, akkor vagy annyi könnyen érthető bevitt mezőt kell rendelkezésre bocsátani, ahány bevitt MARC-hívójelekkel lehetséges (és ezek száma rendkívül nagy), vagy meg kell oldani egy MARC-bevitelhez hasonló, valamivel érthetőbb bevittet.

Kevésbé ismert tény, hogy valójában alig képzelhető el név olyan metaadata, mely a MARC21 szerint ne lenne rögzíthető. Például a nevek etimológiai tulajdonságai:

Használati típusa:	pl. „köznyelvi” [ma élő] (János-hegy), „táj- és rétegnyelvi” (Tobruk [A vörösberényi Csebere-dűlőben 1941-ben, Tobruk ostroma idején nehéz körülmények között felépült papírgyári lakótelep])
Megértéstörténeti típusa:	pl. „értett” (Erdély), „elhomályosult” (Pozsonyi-hegy, máj), régies (Erdőelve, bürü).
Nyelvkörnyezeti típusa:	pl. „endonima” ([nemzeti nyelven belül keletkezett] János-hegy), Lő [Légrád Árpád-kori neve], Paris [francia endonima]; „exonima” ([nemzeti nyelven kívül keletkezett] Párizs, Óperenciás-tenger)
Nyelvtörténeti típusa:	pl. „ősmagyar kor” (Óperenciás-tenger), „Árpád-kor” (Lő), „középmagyar kor”; (Légrád).

Mindez például a török idők előtt létező szlavóniai Árpatarló esetében:

**Árpatarló** (köznyelvi, elhomályosult; endonima, Árpád-kor)

Egy ilyen megjelenítés esetén biztosítani kell, hogy ha a felhasználó a zárójelben szereplő jellemzők valamelyikére kattint, akkor egy felugró részablakban megjelenik a minősítés típusa (pl. megértés-történeti) és a lehetséges értékei (pl. történelmi elhomályosult, történelmi értett, értett).

A felhasználóbarát bevitel esetén biztosítani kell, hogy az új névhez a grammatikai minősítéseket megfelelő részablakon keresztül hozzá lehessen kapcsolni.

Ugyanakkor biztosítani kell a MARC szerinti minősítést is. Példánkban:

072 #7 \$aelhomályosult \$2Megértés-történeti osztályozás  
151 ## \$aÁrpatarló  
072 #7 \$aelavult \$2Használat/elterjedés  
151 ## \$aÁrpatarló  
072 #7 \$aegyéb \$2Nemzeti igazgatási osztályozás  
151 ## \$aÁrpatarló

Egy másik példa ugyanarra a MARC-mezőhívójelre:

072 #7 \$aMammalia \$2Zoológiai taxonómia  
150 ## \$aprérikutya

Ez a helyzet merőben szokatlan. A MARC21 első pillantásra egy nagyon szűk szakterület adatcsere-formátuma, ezért kevesen tudják, hogy ettől még a formátum rendkívül rugalmas. Valójában nincs olyan lehetséges névtéri metaadat – legyen bármilyen szakterületé –, melyet a MARC szerint nem lehet rögzíteni. Ezért minden, MARC-ban rögzített adatelem valójában MARC-független.

A névtérnek mindenképpen kell saját formátumának lenni, azaz nem épülhet a MARC-ra. Mégse lehet eltekinteni attól, hogy a MARC-formátumot ne kezelje, mivel a különféle külső felhasználók számára egy magyar nemzeti névtér formátuma nem tekinthető kötelezőnek. Más szóval a névtérkezelő-rendszerben kell alkalmazkodni a lehetsé-

## Ungváry R.: Névtérkezelő készítésének nehézségei

ges felhasználói csoportok adatcsere-igényeihez. A névtér saját formátumának elemeit meg kell feleltetni például a MARC21 elemeivel.

Ezáltal a névtérbe teljesen felhasználóbarát módon (tehát nem alkalmazva a MARC21 mező-indikátor–almező szerkezetét) kell tudni rögzíteni a nevek metaadatait. A névtérkezelőt kell kialakítani úgy, hogy ezeket egyszerű parancsra MARC-formátumban is strukturálni tudja.

Ráadásul számítani kell arra, hogy a levéltári, ill. múzeumi igények nem felétlenül csak ugyanazt a formátumot fogják igényelni, mint a könyvtárak. Ezért a névtéreket úgy kell tervezni, hogy különböző utólagos számítástechnikai-programozási igénybevétel nélkül a későbbiekben maga a professzionális felhasználó tudjon további adatcsere-formátumok szerinti export-import lehetőségeket létrehozni a rendszerben.

Mindez átvezet a metaadatok kérdéséhez általában.

### A metaadatok

A nemzeti névtéreket egyszerre, de még rövid néhány év alatt se lehet feltölteni. A nagy értelmező szótárak készítéséhez hasonlóan ez a munka több évtizedet is igénybe vehet.

Ugyanakkor azt se lehet előre meghatározni, hogy a neveknek milyen metaadatai legyenek. Előre láthatóan számos lehetséges érdekelt szakma és intézménye csak idővel csatlakozik a névtérhez. Előbb ugyanis be kell látnia, hogy számukra is hasznos, ha neveiket, szavaikat a nemzeti névtéren keresztül kezelhetik. Ehhez azonban az kell, hogy a névtér nyitott legyen az új, lehetséges felhasználók irányába, miközben pontosan nem is ismerhető még, ki mindenki lehet az.

A neveknek ezért rendkívül sok, előre nem látható metaadata lehet. Ha például a névtérhez nyelvészeti intézmények csatlakoznak, akkor több tucat grammatikai (szófaj, szótípus stb.), nyelvtörténeti-etimológiai (Árpád-kori név, ugor eredetű név, endogén név stb.), nyelvhasználati (köznyelvi, nyelvjárási, szleng stb.) metaadatra van szükség.

Az államigazgatási felhasználás esetén igazgatási (hivatalos, normatív stb.), nemzetközi igazgatási (NUTS, LAU), postai (irányítószám, telefon-körzet-szám stb.) metaadatok szükségesek.

Mindez csak töredéke a lehetséges résztvevői igényeknek.

Ehhez biztosítani kell, hogy a névtérbe új metaadat-típusokat lehessen könnyen, informatikai segítség nélkül fölvenni, módosítani, törölni.

Ez azt jelenti, hogy a professzionális felhasználó számára kellenek felületek, melyeken egyszerű módon meghatározhatja

- a) a metaadatok névtéren belüli megjelenítését mind a szerkesztő felületen, mind a különböző megjelenítési felületeken (a végfelhasználó számára); és
- b) az előző fejezethez kapcsolódóan a metaadatok migrációs (export-import) formátumát.

A 6. ábrán a Relex tezaszuszkezelő metaadat-definíciós ablaka látható.

### A tervezés részletessége

A fenti problémákból az következik, hogy már az előtervek és a követelményelemzések szintjén is nagyobb részletesség szükséges, mint a hagyományos, kereskedelmi célú tervezések esetén, amikor

például a Scrum-módszerrel általánosabb, majd fokozatosan speciálisabb tervek készülnek. A névterek sajátos komplexitása következtében a tervezendő kezelőrendszer bizonyos szempontból egyszerűbb, mint számos igazgatási, pénzügyi stb. rendszer, ugyanakkor speciálisabb azoknál. Ez ugyan minden rendszerre igaz, de névterek esetében ez a specifikusság alig ismert, az informatikában pedig – egyelőre – szokatlan. El kell kerülni, hogy az előtervek alapján első lépésben egy olyan kiinduló kezelőrendszer keletkezzék, amelyben bizonyos eljárásokat túlságosan „bekábeleztek” a megszokott gyakorlat alapján. Ilyen esetben ugyanis később derülnek ki a félreértések. Ezáltal a szükséges változtatások a rendszer alapvető szerkezetét érintik. Az utólagos módosítás mindig többletmunka. Ez nagyon kellemetlen anyagi következményekkel, és nem utolsósorban pszichológiai megterheléssel járhat. Az is előfordulhat, hogy anyagiak hiányában már nem lehet változtatni.

Már az előtervekben kialakított konkrét modell (struktúra) egyes részeinek is pontosan tükrözniük kell a névtéri realitásokat.

mező	indik.	almező	almező ho...	elvál.jel el...	elvál.jel ut...	ismétlődő ...
046	##	t	25	Bef.:		
046	##	u	25	<	>	
046	##	v	25	URLfor:		;
046	##	2	25	Forrás:		;
052	#*	a	25	Földro.:		;
052	#*	d	25	Hely:		
052	#*	2	25	(Forrás: )		
065	##	a	80	Oszl. r:		
065	##	c	80	Indkif:		
065	##	2	25	(Forrás: )		

6. ábra Metaadatok felhasználóbarát definíciós ablaka. Meghatározhat az adattípus azonosítóját, ha kell indikátorait, almezőit, ezek hosszát, továbbá megjelenítésük sorfolytonos formáját az elválasztójelekkel

### Irodalom

- [1] ISAAR (CPF) : International Standard Archival Authority Record for Corporate Bodies, Persons and Families. – Second ed. – International Council

on Archives. – ICA Committee on Descriptive Standards: Canberra, 2004.

<<http://www.ica.org/en/node/30230>>;

<[http://bfl.archivportal.hu/id-1037-](http://bfl.archivportal.hu/id-1037-isaar_cpf_szervezetek_testuletek.html#521)

[isaar\\_cpf\\_szervezetek\\_testuletek.html#521](http://bfl.archivportal.hu/id-1037-isaar_cpf_szervezetek_testuletek.html#521)>

- [2] ISAD/G: A levéltári anyagleírás nemzetközi szabványa. – Ottawa, Nemzetközi Levéltári Tanács, 1994. – Magyar fordítása: 2009.
- [3] ISO 2788–1986 ; Documentation – Guidelines for the establishment and development of monolingual thesauri
- [4] ISO 5964–1985 ; Documentation – Guidelines for the establishment and development of multilingual thesauri
- [5] ISO 21127. Definition of the CIDOC Conceptual Reference Model. Produced by the ICOM/CIDOC Documentation Standards Group, continued by the CIDOC CRM Special Interest Group. Version 4.2.1. Szerk.: Nick Crofts, [et al]. – International Council of Museums (ICOM), CIDOC CRM, Special Interest Group, 2006.  
<[http://cidoc.ics.forth.gr/docs/cidoc\\_crm\\_version\\_4.2.1.pdf](http://cidoc.ics.forth.gr/docs/cidoc_crm_version_4.2.1.pdf)>
- [6] ISO 25964–2011 ; Information and documentation – Thesauri and interoperability with other vocabularies
- [7] Köztaurusz. 2018. március 1. Szerk. Ungváry Rudolf. Budapest, Országos Széchényi Könyvtár, 2000–[1992–]  
<<http://www.oszk.hu — Útmutató/Tezaurusz>> és  
<<http://mek.oszk.hu/adatbazis/thes.htm>>  
<<http://mek.oszk.hu/adat-bazis/thes.htm>>
- [8] KSZ/5. Földrajzi nevek mint adatbázisrekordok tárgyi hozzáférési pontjai. Közreadja a Könyvtári és Szakirodalmi Tájékoztatási Szabványosítási Bizottság. Kiadja az Országos Széchényi Könyvtár, Budapest. A jóváhagyás időpontja 2005. június.  
<<http://www.ki.oszk.hu/szabalyzatok/foldrajzinevek/ksz5.pdf>>
- [9] MARC21 Format for authority data. [Washington] : Library of Congress. Network Development and MARC Standards Office; Cataloging Distribution Service. – 1999 Ed.. Update No. 1 (October 2001) through Update No. 9 (October 2008). – Washinton, Library of Congress, Network Developoment and MARC Standards Office.  
<http://www.loc.gov/marc/authority/ecadhome.html>  
Magyarul: MARC21 Besorolási adatok adatcsereformátuma : OSZK. 2018. [kézirát]
- [10] MARC Format for Classification Data. [Washington] : Library of Congress. Network Development and MARC Standards Office; Cataloging Distribution Service. – 2000 Edition. Update No. 1 (October 2001) through Update No. 9 (October 2008)  
<<http://www.loc.gov/marc/classification/eccdhome.html>>
- [11] MSZ 3418–87 Magyar nyelvű információkereső tezauruszok. Szerkezete, részei és formái.
- [12] SCHWABER, Ken; SUTHERLAND, Jeff.. The Scrum Guide - The Definitive Guide to Scrum: The Rules of the Game. 2011.  
<<https://www.scrum.org/resources/scrum-guide>>
- [13] UNGVÁRY Rudolf: Az OSZK tezaurusz és a KÖZTAURUSZ. = Könyvtári Figyelő, Új folyam 11. (47.) évf. 2001. 1. szám. p. 11–40.  
<<http://primula.oszk.hu/kiadvany/kf/2001/1/ungvary.html>>
- [14] UNGVÁRY Rudolf: Tezaurusz és ontológia, avagy a fogalmi ismertetőjegyek generikus öröklődésének formalizálása. = Tudományos és Műszaki Tájékoztatás, 51. évf. 2004. 5. sz. p. 175–191.  
<[http://tmt.omikk.bme.hu/show\\_news.html?id=3615&issue\\_id=450](http://tmt.omikk.bme.hu/show_news.html?id=3615&issue_id=450)>
- [15] Rudolf UNGVÁRY, Tamás RADNAI: Thesaurus and Ontology: a Formalism of Generic Inheritance of Conceptual Characteristics, 9th IEEE International Conference on Intelligent Engineering Systems, INES, 2005.  
<[http://ieeexplore.ieee.org/xpl/freeabs\\_all.jsp?arnumber=1555149](http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=1555149)>
- [16] UNGVÁRY Rudolf; PÁSZTI László: A földrajzi nevek mint az adatbázisrekordok hozzáférési pontjai. Egy könyvtári szabályzat elvei és megoldásai. = Könyvtári Figyelő, 51. köt. 2005. 4.  
<<http://epa.oszk.hu/00100/00143/00057/ungvary.html>>
- [17] UNGVÁRY Rudolf: Tezauruszok mint „kisvilág”. Kapcsoltság a fogalmak között. = W3C Szemantikus Web Műhelykonferencia. 2006. április 13. Budapest, MTA SZTAKI, 2006.  
<[http://www.w3c.hu/rendezvenyek/2006/szemweb/eak/oszk\\_ungvaryr.pdf](http://www.w3c.hu/rendezvenyek/2006/szemweb/eak/oszk_ungvaryr.pdf)>
- [18] UNGVÁRY Rudolf: A tezaurusz mint „kisvilág”. Az optimális fogalomelérés hálójá. = Networkshop 2006. Miskolc, Miskolci Egyetem, 2006. április 19. – április 21. : elektronikus dok. : kongferencia anyag. / [rend., közread. A Nemzeti Iformációs Infrastruktúra Fejlesztési Iroda] . – [Budapest] : NIIF Iroda, cop. 2006. CD-ROM.  
<<https://nws.niif.hu/ncd2006/>>
- [19] UNGVÁRY Rudolf: Az ontológiák legfelső generikus szintje, a csúcsfogalmak természetes rendszere és a DOLCE kritikája. = IV. Magyar Számítógépes Konferencia. MSZNY 2006. Szeged, 2006. december 7–8. Szerk.: Alaxin Zoltán, Csendes Dóra. Szeged, Szegedi Tudományegyetem Informatikai Tanszékcsoport, 2006. p. 95–96.
- [20] UNGVÁRY Rudolf: A Köztaurusz és az ontológiák csúcsfogalmi. = Networkshop 2007. Eger, Esterházy Károly Főiskola, 2007. április 11. – április 13. : elektronikus dok. : kongferencia anyag. / [rend., közread. A Nemzeti Információs Infrastruktúra Fej-

- lesztési Iroda] . – [Budapest] : NIIF Iroda, cop. 2007. CD-ROM.  
<<https://nws.niif.hu/ncd2007/>>  
<[http://videotorium.hu/hu/recordings/details/1769\\_A\\_Koztaurusz\\_es\\_az\\_ontologiak\\_csucsfojalmai](http://videotorium.hu/hu/recordings/details/1769_A_Koztaurusz_es_az_ontologiak_csucsfojalmai)>  
<<https://nws.niif.hu/ncd2007/docs/ehu/044.pdf>>
- [21] UNGVÁRY Rudolf: Ontológiai és tezauszok csúcsfogalmai. = TMT, 54. évf. 2007 10. sz.  
<[http://tmt.omikk.bme.hu/show\\_news.html?id=4800&issue\\_id=487](http://tmt.omikk.bme.hu/show_news.html?id=4800&issue_id=487)>
- [22] UNGVÁRY Rudolf: A földrajzi nevek és szerepük az információkeresésben. Nemzetközi fejlődés, problémák és tapasztalatok a Köztauruszban és Geotauruszban. = Könyvtári Figyelő. 2008. 3.  
<<http://ki.oszk.hu/ki/2010/10/a-foldrajzi-nevek-es-szerepuk-az-informaciokeresesben/>>
- [23] UNGVÁRY Rudolf: A földrajzi nevek és szerepük tárgyi kifejezésként : Nemzetközi fejlődés, problémák és tapasztalatok a Köztauruszban és a Geotauruszban. = Networkshop 2008. Dunaújváros, Dunaújvárosi Főiskola, 2008. március 17. – 19. : elektronikus dok. : kongferencia anyag. / [rend., közread. A Nemzeti Információs Infrastruktúra Fejlesztési Iroda] . – [Budapest] : NIIF Iroda, cop. 2008. CD-ROM.  
<<https://nws.niif.hu/ncd2008/>>  
<[http://videotorium.hu/hu/recordings/details/517\\_Foldrajzi\\_nevek\\_es\\_szerepuk\\_targyi\\_kifejezeskent\\_Nemzetkozi\\_fejlodes\\_problema\\_es\\_tapasztalatok\\_a\\_Koztauruszban\\_es\\_a\\_Geotauruszban](http://videotorium.hu/hu/recordings/details/517_Foldrajzi_nevek_es_szerepuk_targyi_kifejezeskent_Nemzetkozi_fejlodes_problema_es_tapasztalatok_a_Koztauruszban_es_a_Geotauruszban)>
- [24] UNGVÁRY Rudolf: MARC21/HUNMARC. A besorolási adatok adatsere-formátuma. Főbb jellemzők, fejlődés és problémák. = Könyvtári Figyelő, 20. (56.) köt., 1. sz. 2010. p. 4–70.  
<<http://ki.oszk.hu/ki/2010/10/marc21hunmarc-a-besorolasi-adatok-metaadat-formatuma-fobb-jellemzok-fejlodes-es-problema/>>
- [25] UNGVÁRY Rudolf: Névtér és névtérkezelő rendszer: követelmények, jellemzők. 44. Könyvtáros vándorgyűlés, 2012. Győr. Széchenyi István Egyetem, 2012. június 12–14.  
<[https://videotorium.hu/hu/recordings/details/6003\\_Nevkeze\\_lo\\_rendszerek\\_kovetelmenyek\\_jellemzok](https://videotorium.hu/hu/recordings/details/6003_Nevkeze_lo_rendszerek_kovetelmenyek_jellemzok)>
- [26] UNGVÁRY Rudolf: A webes Relex. Szabadon használható névtér-, tezausz- és ontológiaszerkesztő. Társelőadók Király László, Szabó Gyula. Networkshop 2012. Veszprém, Pannon Egyetem, 2012. április 11.–13.  
<[https://videotorium.hu/hu/recordings/details/4166\\_A\\_webes\\_Relex\\_Szabadon\\_hasznalhato\\_nevter-tezausz-es-ontologiaszerkeszto\\_rendszer](https://videotorium.hu/hu/recordings/details/4166_A_webes_Relex_Szabadon_hasznalhato_nevter-tezausz-es-ontologiaszerkeszto_rendszer)>  
Networkshop 2013. [elektronikus dok.] : [Pécs], 2013. március 26. – 28. / [rend., közread. a Nemzeti Információs Infrastruktúra Fejlesztési Intézet, Hungarnet Egyesület]
- <[http://nws.niif.hu/ncd2012/index.htm?\\_ga=1.49608870.1256184114.1424339593](http://nws.niif.hu/ncd2012/index.htm?_ga=1.49608870.1256184114.1424339593)>
- [27] UNGVÁRY Rudolf: A névterek és az adatok tulajdonságai. A történet, avagy a fogalmi és a nyelvi szint. = Tudományos és Műszaki Tájékoztatás, 59. köt. 3. sz. 2012. p. 91–105.  
<[http://tmt.omikk.bme.hu/print.html?id=5594&issue\\_id=535](http://tmt.omikk.bme.hu/print.html?id=5594&issue_id=535)>
- [28] UNGVÁRY Rudolf: A névtér mint kulturális szűkséglet. A nemzeti aggregátor kitüntetett feladata. = Tudományos és Műszaki Tájékoztatás, 59. köt. 3. sz. 2012. p. 320–326.  
<[http://tmt.omikk.bme.hu/issue.html?issue\\_id=540](http://tmt.omikk.bme.hu/issue.html?issue_id=540)>
- [29] UNGVÁRY Rudolf: A földrajzi névtér jellemzői. A meglévő földrajzi névtér kritikája. = Networkshop 2013. Sopron, Nyugat-Magyarországi Egyetem, 2013. március 26.–28.  
<[https://videotorium.hu/hu/recordings/details/6179\\_A\\_foldrajzi\\_nevter\\_jellemzoi](https://videotorium.hu/hu/recordings/details/6179_A_foldrajzi_nevter_jellemzoi)>
- [30] UNGVÁRY Rudolf: Névtér és földrajzinév-tárak. Észrevételek a Magyarország Földrajzinév-tára jövőbeni továbbfejlesztéséről. = Tudományos és Műszaki Tájékoztatás. 63. köt. 4. sz., 2016. p. 13–157.  
<<https://tmt.omikk.bme.hu/tmt/article/download/50/51>>
- [31] UNGVÁRY Rudolf: A nyomtatott és webes szótárakról. = Orvosi Könyvtárak, 2017. 1. sz pp. 15–17.  
<<https://drive.google.com/file/d/0B5tgn-v7x9zedEtTa2IOR0FGaTg/view>>
- [32] UNGVÁRY Rudolf: A névtér értelme. Filozófiai szerkezeti jellemzők. Tudományos és Műszaki Tájékoztatás, 2018. 1. sz.  
<<https://tmt.omikk.bme.hu/tmt/article/view/1720>>

## Megjegyzések

<sup>1</sup> Az első informatikai forradalom az információk közösségi átadására megszületett nyelv, vele a beszéd volt, mely a társadalmi cselekvés alapja lett. A második az írás, mellyel az információk nemzedékek közötti átadása vált hatékonyabban lehetségessé. A harmadik pedig a rendszerezés és a rendezés forradalma.

<sup>2</sup> A Magyarországon 2000 végéig időrendben elkészült tezauszok jegyzékét lásd [13].

Beérkezett: 2018. VII. 30-án.



**Ungváry Rudolf**  
informatikus mérnök.  
E-mail: [ungvaryr@gmail.com](mailto:ungvaryr@gmail.com)