

Tartalom

1	A névterek fontosságáról	4
2	Létező névtérhasználati gyakorlatok.....	7
2.1	Fájlnevek.....	7
2.2	Domainnevek.....	8
2.3	Emailcímek.....	9
2.4	Telefonszámok.....	9
2.5	Földrajzi helyek.....	10
2.5.1	Postai címek.....	10
2.5.2	Földrajzi térképek, földrajzi nevek.....	10
2.5.2.1	GeoNames	11
2.5.2.2	Getty TGN.....	12
2.5.2.3	Google Maps.....	14
2.5.2.4	OpenStreetMap.....	15
2.6	Webes névazonosítási szabványok.....	16
2.7	Személynevek.....	18
2.8	Műcímek.....	18
2.8.1	Könyvek, folyóiratok.....	21
2.8.1.1	ISBN.....	21
2.8.1.2	ISSN.....	24
2.8.1.3	ISTC.....	27
2.8.1.4	DOI.....	28
2.8.1.5	NBN.....	28
2.8.2	Zeneszámok.....	29
2.8.2.1	ISMN.....	29
2.8.2.2	ISRC.....	31
2.8.2.3	ISWC.....	32
2.8.3	Filmek, audiovizuális tartalmak.....	35
2.8.3.1	ISAN.....	35
2.8.3.2	IMDb.....	47
2.9	Személyek azonosító rendszerei.....	50
2.9.1	Tudományos kutatók és cikkek azonosító rendszerei.....	51



2.9.1.1	MTMT.....	51
2.9.1.2	ODT.....	54
2.9.1.3	ORCID.....	55
2.9.1.4	ResearcherID.....	57
2.9.1.5	Web of Science.....	60
2.9.1.6	Google Scholar.....	60
2.9.1.7	Scopus.....	61
2.9.1.8	ResearchGate.....	64
2.9.1.9	Academia.edu.....	65
2.9.2	További azonosító rendszerek személyekre, helyekre.....	66
2.9.2.1	VIAF.....	66
2.9.2.2	ISNI.....	68
2.9.2.3	PIM személynévtér.....	74
2.9.2.4	ULAN.....	75
2.9.2.5	GDN.....	77
2.9.2.6	BNF.....	78
2.9.2.7	A könyvtárak besorolási állománya.....	79
2.9.2.8	ISADN.....	80
2.9.2.9	Wikidata.....	80
3	Névstruktúra.....	82
3.1	Relációelmélet.....	82
3.1.1	Relációelméleti alapfogalmak.....	82
3.1.1.1	A reláció fogalma.....	83
3.1.1.2	Relációtulajdonságok.....	83
3.1.1.3	Függvény.....	84
3.1.1.4	Művelet.....	85
3.1.1.5	Nevezetes matematikai relációk.....	86
3.1.1.6	Relációműveletek.....	87
3.1.1.7	Reifikáció.....	88
3.1.2	Ontológiai alaprelációk.....	88
3.1.3	KOS-relációk.....	93
3.1.3.1	Generikus reláció.....	94
3.1.3.2	Partitív reláció.....	96

3.1.3.3	Szinonima.....	98
3.1.3.4	Homonima.....	99
3.1.3.5	Meghatározója (előzménye, következménye) reláció.....	99
3.1.3.6	Hierarchikus alá-fölérendeltje (tartalmazás) reláció.....	100
3.1.3.7	Lexikografikus rendezési reláció.....	100
3.1.3.8	Rokona (egyéb) reláció.....	100
3.2	A használatba vehető relációk listája.....	101
3.3	Tudásszervezési rendszerek.....	102
3.3.1	Formális keretrendszer.....	102
3.3.2	Terminuslista.....	105
3.3.2.1	Formális jellemzés.....	106
3.3.3	Taxonómia.....	106
3.3.3.1	Formális jellemzés.....	111
3.3.4	Tezaurusz.....	112
3.3.4.1	Formális jellemzés.....	120
3.3.5	Folkszonómia.....	121
3.3.5.1	Formális jellemzés.....	140
3.3.6	Formális ontológia.....	141
3.3.6.1	Formális jellemzés.....	142
3.4	Dokumentumleíró rendszerek.....	142
4	Tulajdonnév és megnevezés.....	144
4.1	A névtér összetevői.....	144
4.2	A tulajdonnévtér meghatározása.....	145
4.3	A nevek egyértelműsítése.....	146
5	Lokális, nemzeti és nemzetközi névtér.....	150
5.1	A név fogalma, funkciója.....	150
5.2	A névtér fogalma, funkciója.....	150
5.3	Lokális névterek.....	153
5.4	Nemzeti névterek.....	153
5.4.1	A nemzeti névtér szűkebb értelmezése.....	153
5.4.2	A nemzeti névtér tágabb értelmezése (névtár).....	155
5.5	Globális névterek.....	156
6	A névtér fogalmának ontológiai elkötelezettségei.....	158

6.1	A természeti és társadalmi tények különbsége.....	158
6.2	Névterek és tudásszervezési rendszerek.....	158
6.2.1	A tulajdonnévterek struktúrája.....	159
6.2.1.1	A személynévtér struktúrája.....	159
6.2.1.2	A testületi névtér struktúrája.....	160
6.2.1.3	A földrajzi névtér struktúrája.....	162
6.2.1.4	Az írásműcímtér struktúrája.....	168
6.2.1.5	Az audiovizuálismű-címterek struktúrája.....	171
6.2.2	A köznévterek struktúrája.....	172
7	A névtérdefiníció fejlesztésre gyakorolt hatásai.....	174
8	Hivatkozások.....	177
8.1	Irodalom.....	177
8.2	Szabványok, szolgáltatások.....	181

1 A névterek fontosságáról

A digitális technológiák elterjedése komoly kihívások elé állítja a közgyűjteményeket, könyvtárakat, kulturális adatszolgáltatókat. A legelső feladat természetesen az, hogy a kulturális értékeket át kell emelni a digitális hálózati térbe. Amint megvalósítjuk ezt a feladatot, a kulturális dokumentumaink elérhetővé válnak bárhol, bármikor, bárki számára. Ez a hozzáférés azonban csak elvi lehetőség marad mindaddig, amíg nem biztosítjuk azt, hogy a kulturális dokumentumokat kereső felhasználók könnyen és gyorsan megtalálhassák mindazt a digitális térben, ami érdekli őket. Ehhez meg kell megtalálni, ki kell fejleszteni és fent kell tartani a hozzáférés gyakorlati biztosítékait, technológiáit. A hálózati világban kezdetektől fogva kiemelt szerepet vittek azok a szolgáltatók, amelyek ezt tudták valahogyan biztosítani. A hozzáférés gyakorlati lehetőségét a mai napig a keresőszolgáltatások képesek felkínálni a legnagyobb sikerrel. Látni kell azonban azt, hogy a népszerű globális keresőszolgáltatók (mint például a Google vagy a Bing) elsősorban olyan információs térben működőképesek, amelyek alapvetően a dokumentumok tartalmában való keresés lehetőségét nyújtják. A közgyűjteményi világ szereplői azonban már évszázadok óta olyan információs rendszereket építenek és kezelnek, amelyekben sosem a – domináns mértékben írásos – dokumentumok tartalmában való szabadszavas keresés volt a lényeg. A közgyűjtemények (könyvtárak, levéltárak, múzeumok, zene- és filmtárak, audiovizuális archívumok) kezdetektől fogva a dokumentumok, kulturális termékek, kulturális artefaktumok metaadataiból építenek rendszereket. A strukturált adatrendszerekben való kereséssel kapcsolatban azonban még sok nyitott kérdés létezik, és még több az előttünk álló, még el nem végzett fejlesztési feladat. A nyitott kérdésekre adható válaszok tudományos kutatást, a metaadatok alapján való eligazító szolgáltatások létrehozása pedig komoly fejlesztési munkát kíván meg.

A metaadatok fontossága a nem-szöveg alapú, audiovizuális archívumok világára mutatta is látható. Amilyen mértékben kerülnek bele az audiovizuális tartalmak az interaktív hálózat világába, olyan mértékben erősödik a felhasználói igény az ilyen művek kereshetőségére. Ennek kielégítésére a szolgáltatók által megvalósított, közös metaadatkezelés a jó válasz. Ebben a kollaborációban kiemelt szerepet játszanak a névterek, amelyek révén egyértelműen azonosíthatók a keresett/kereshető művek. Hogy miért van szükség az egyértelmű azonosításra? Rossz példaként hivatkozhatunk a YouTube gyakorlatára, ahol ha rákeresünk egy filmre, akkor annyi találatot kapunk, ahányszor a felhasználók feltöltötték az adott filmet, de azt, hogy vajon ugyanarról a filmről van-e szó vagy sem, azt csak akkor tudjuk megállapítani, ha egyenként és alaposan megvizsgáljuk az összes találatot.¹ Arra sem képes a YouTube (a látható szolgáltatási térben), hogy egy adott filmet összekapcsoljon más szolgáltatók (mondjuk az IMDb) adataival. Amennyiben a feltöltött YouTube videókhoz valamilyen névtérből hozzá lennének rendelve olyan azonosítók, amelyeket más szolgáltatók is hozzárendeltek már a saját adataikhoz, akkor e közös névtér-azonosító mentén a szolgáltatások összekapcsolhatókká válnának, és a felhasználók lehetőséget kapnának, hogy egyetlen kattintással „ugráljanak” a különböző oldalak között. A névterek fő funkciója tehát a szolgáltatások közti interoperabilitás megteremtése.

¹ Az más kérdés, hogy a YouTube – a „háttérben” zajló képi elemzések eredményeként – tudja, tudhatja, hogy két feltöltött videó ugyanannak a filmnek a két példánya. Ebből a felhasználó nem érzel semmit.

Természetesen a nagy keresőszolgáltatók régóta foglalkoznak azzal, hogy milyen módokon, milyen eszközökkel lehet meghaladni a szabadszavas keresők régóta ismert hiányosságait, gyengeségeit. Abban nem nagyon van vita, hogy a továbblépés lehetőségét a szabadszavas keresés szemantikai vakságának megszüntetése adhatja, de abban a kérdésben, hogy mindezt hogyan lehet megvalósítani, még nem nagyon alakult ki konszenzus. A **szemantikus keresés** kiemelt fontosságát fogalmazta meg már a 2000-ben meghirdetett **szemantikus web** projekt, ami az azóta eltelt közel két évtizedben még nem jelent meg senki sem átütő szolgáltatással e téren. Azt azért érdemes megjegyezni itt, hogy a Google már a kétezres évek elejétől fogva folyamatosan építi azokat a szemantikus keresőmodulokat, kérdés-válasz (Q&A) rendszereket, amelyek jól behatárolható, szűk területeken már képesek szemantikus szolgáltatásokra, és ezeket a modulokat be is építi a nagy keresőszolgáltatásába, de az egységes, integrált, széles információs tartományban keresni képes szemantikus keresővel még a keresőszolgáltatások világában piacvezető cég sem tudott előállni.

A kétezres évek végétől egyre divatosabbá váló és fokozatosan terjedő, RDF-triplet logikára épülő, **linked data** alapú elképzelések a szemantikus web filozófiáját viszik tovább, miközben egyre hangsúlyosabbá válnak az ontológiák, névterek építésének szükségességét hangoztató nézetek. Ebben az új megközelítésben, ezen új technikai paradigmán belül tudományos viták során kell tisztázni a közeljövőben azt a kérdést, hogy mi a pontos fogalmi és technológiai kapcsolat a névterek és az ontológiák (illetve más tudásszervezési rendszerek) között. A névterek esetében viszont már most egyértelmű, hogy azok legfontosabb funkciója az lehet, hogy segítségükkel biztosítani lehet a hálózatra került nagy adatrendszerek közti keresztlinkek elhelyezését, és ezáltal a rendszerek közti átjárhatóságot, interoperabilitást.

A jelen projekt felfogásában a – szűkebb értelemben vett – névterek rendszerezett, egyértelműsített, egyedi azonosítóval ellátott, tipizált tulajdonnév-állományok. A közgyűjtemények, a tudományos, oktatási és társadalmi szervezetek által fenntartott kulturális adatbázisok névállományai (személynevek, testületi nevek, földrajzi nevek stb.) közösek, nagy mértékben egymást átfedők. Egy személy, egy intézmény, egy földrajzi hely sokféle módon kapcsolódhat a könyvekhez, újságokhoz, filmekhez, zenékhez, képekhez, egyszóval a kulturális dokumentumokhoz. A nevek egységes, integrált kezelését, egyértelmű azonosítását névterek segítségével lehet biztosítani. A névterek lehetővé teszik a kulturális intézmények közti kollaborációt, biztosítják a kulturális adatbázisok, szolgáltatások közti interoperabilitást, az átfedő adatszegmensekkel rendelkező információs rendszerek közti sokszálú kapcsolatrendszer kiépítését, a nevek egységes és integrált kezelésével összekapcsolják és könnyen bejárhatóvá teszik a magyar kultúra dokumentumait, lexikonjait, kézikönyveit.

Vannak már névterek a világban globális, nemzeti és lokális-intézményi szinten egyaránt, de ezek mind típusonként külön állnak, nincsenek érdemi kapcsolatban egymással, ezek a rendszerek még nem integrálódtak.² A cél most az, hogy ezeket közös platformra hozni és egységes infrastruktúrával kell őket kiszolgálni – egyelőre nemzeti keretekben gondolkozva, de természetesen minden fejlesztést úgy tervezve és végezve, hogy a globális névterekhez való kapcsolódást is biztosítani tudjuk az első lehetséges alkalommal.

² A VIAF, az ISNI a személyek, testületek, helyek globális azonosítására szolgál. Az összerendelt azonosítókat, a hozzájuk kapcsolható adatokat a nemzeti könyvtárak által beküldött besorolási adataiból (authority állományokból) állítják össze. Az ISAN az audiovizuális művek nemzetközi azonosítását végzi.

A névterek használata kompatibilis a szemantikus web, a linked data világával is, ahol az az elvárás, hogy a dokumentumokat, ágenseket (személyeket, szervezeteket), helyeket mind egyetlen azonosítóval (URI-val) kell leírni.

A hazai jogszabályok az Országos Széchényi Könyvtárat jelölik meg a nemzeti névterek építéséért felelős intézményként. A nemzeti névterek felépítésének munkálatai megindultak, és már a kezdetek kezdetén megmutatkoztak a közös névtérhasználat előnyei és eredményei. Bármilyen filozófia mentén is építsenek archívumot a saját maguk számára a hazai könyvtárak, levéltárak, múzeumok, audiovizuális archívumok, mindenképpen lokális névtereket kell építeniük, és érdemes csatlakozniuk a nemzeti névterek adatállományaihoz is. A nemzetközi névterekhez való kapcsolódást a nemzeti névtereken keresztül lehet megvalósítani, mert a globális kooperáció megteremtését (az érthető okokból felmerülő kontrolligény miatt) a nemzeti névtereken keresztüli együttműködések rendszerétől lehet várni (nem pedig a lokális névterek globális kollaborációjától).

A gyakorlatban működőképes globális kollaborációs lehetőségek megteremtéséhez szükség van a nemzeti (és természetesen globális) névterekre, de ezek mellett kellenek a lokális, intézményi névterek is. A háromszintű névtérrendszerre azért van szükség, mert könnyen felmerülhet az a helyzet, hogy nem kell minden adatot a nemzetközi, globális névterekben kezelni, mert ez szétfeszíthetné azok kereteit. A nemzetközi névterekbe nem érdemes minden, az adott nemzet kultúrája számára fontos névelemet felvenni, mint ahogy a nemzeti névterekbe sem érdemes (nem is lehet) minden, a lokális, intézményi névterekben tárolt adatot átemelni. Konkrét példákkal élve: egy mozgóképes nemzeti névtérbe fel lehet venni az összes olyan dokumentumfilmet, kisfilmet, amelyet valamilyen szempontok szerint megőrzendő hungarikumnak minősítünk, de ezek mindegyikét nem kell (sőt, nem is lehet) beregisztrálni a globális névtérbe, vagy egy médiaképzésben keletkezett vizsgafilmet fel lehet venni a képzést végző intézmény névtérébe, de nem feltétlenül kell betenni (befogadni) a nemzeti névtérbe is.

2 Létező névtérhasználati gyakorlatok

A névtérek építésével kapcsolatos fogalmak, definíciók, elvárások, szabályok áttekintése előtt érdemes alaposabban feltárni, bemutatni, hogy milyen névtérek, névtérhasználati gyakorlatok léteznek szerte a világban, azokat milyen célokra használják, milyen módszerek mentén üzemelnek, mit lehet tanulni tőlük. A névtérek jelentőségét egyre jobban el- és felismerik a hálózati világ, a digitális kultúra terjedésével. Globális, nemzeti és lokális névtérek épülnek szerte világban. Mielőtt tisztáznánk, mit is jelent (és mit nem) a névtér fogalma, bemutatunk pár speciális névtérrel, hogy megnézhessük, milyen célok mentén, hogyan működnek.

2.1 Fájlnevek

Talán a legegyszerűbb, de mindenképpen a legtöbbször használt névtér a számítógép fájlrendszere, amelynek egyetlen feladata azt biztosítani, hogy a számítógépen minden állomány egyedi névvel rendelkezzen. Az operációs rendszertől függ, hogy milyen szabályok mentén lehet neveket létrehozni, de ez másodlagos a névtér működése szempontjából. Az egyedi nevek kiosztását viszont megkönnyíti az a lehetőség, hogy a neveket komponensekből lehet felépíteni. Az állományokat csoportokba (folderekbe, könyvtárakba) lehet gyűjteni, az ilyen gyűjtőmappákat egymásba lehet ágyazni, és a tárolóknak is nevet adva úgy lehet biztosítani az állomány nevének egyediségét, hogy az állományt tartalmazó gyűjtőmappák neveit is beépítjük a teljes név összeállításába. Ez egy hierarchikus névazonosítási módszer – a hierarchikus vezérlés minden előnyével és hátrányával együtt. A hierarchikus szerveződés nyelvi szempontból azt jelenti, hogy a neveket névelemekként értelmezhetjük, vagyis egy nevet mint névelemet többször is felhasználhatunk egy konkrét, fájlnev megkomponálásakor. Erre azért van lehetőség, mert az állománynevek összetett nevek, vagyis a teljes – és a teljességében egyedi – állománynév legalább annyi névkomponensből áll, ahány könyvtárba ágyazódik maga az állomány a fájlrendszer gyökeréhez képest. Ezek a névkomponensek azonban nem kell, hogy önmagukban egyediek legyenek, ezek ismétlődhetnek. Bármilyen névelemet többször is felhasználhatunk a hierarchikus névkomponálás során (minden egyes könyvtáron belül újra és újra előfordulhat ugyanaz a névelem).

Azzal a követelménnyel, hogy minden állománynak legyen neve és ez a név egyedi legyen, azt lehet biztosítani, hogy minden állomány egyértelműen azonosítható és biztosan megtalálható legyen a számítógépen. Ez az elvárás megengedi azt, hogy a fájlrendszerben egy állománynak lehessen több neve is, hiszen ha a névegyediség biztosítva van, és minden név össze van kapcsolva egy állománnyal, akkor a nevek felől mindig minden állomány elérhető (fordítva ez nem lenne igaz, de olyan igény nem jelentkezik, hogy egy állomány felől indulva kellene egy adott nevet megtalálnunk). Nyelvtani szempontokat érvényesítve azt mondhatjuk, hogy ha két név ugyanahhoz a dologhoz (itt: állományhoz) tartozik, akkor ebben a névtérben a szinonimitás megengedett, de ha egy név csak egyetlen dologhoz (állományhoz) tartozhat, akkor a homonimitás nem megengedett.

Fontos még megjegyezni, hogy a fájlrendszer, mint névtér, lokális abban az értelemben, hogy minden számítógépen van (legalább) egy fájlrendszer, tehát ilyen típusú névtérből legalább annyi van, mint amennyi számítógép. Ezek a lokális névtérek persze – hálózati kapcsolat mentén

- egybeszervezhető, és ezáltal terjedelmüket tekintve nagyobb névterek is kialakíthatók, de az egyediség ekkor is könnyen fenntartható, csak az összekapcsolt gépeket kell ellátni egy egyedi azonosítóval. A fájlrendszer esetében egyszerű a válasz arra a – minden névtér esetében fontos – kérdésre, hogy ki kezelheti (hozhatja létre, törölheti, változtathatja meg) a neveket. A válasz (a dolgokat kissé leegyszerűsítve): a számítógép kezelője.

2.2 Domainnevek

A fájlrendszerhez hasonló elvárás, ti. a nevek egyediségének és nevekké jelölt dolgok (számítógépes erőforrások) egyértelmű megtalálhatóságának elvárása fogalmazódik meg a webes domainnevek rendszerével mint névtérrel kapcsolatban. Akár a világháló bejegyzett domainneveire, akár a weboldalak címeinek hierarchikusan strukturált halmazára gondolunk, mindkét esetben névtérről beszélhetünk (úgy, hogy utóbbi tartalmazza az előbbi). Mi a domainnevek fő funkciója? Egyértelműen hivatkozható, ember által könnyebben memorizálható karaktersorozatokkal jelölni a weboldalak címeinek közös gyökereit (index.hu, microsoft.com, mit.edu). A domainnevek képzésére vannak szabályok (mit lehet, mit nem lehet névelemként alkalmazni, a névelemek között szeparátorokat kell alkalmazni stb.), és van egy alapvető elvárás a nevek egész halmazára vonatkozóan: minden névnek egyedinek kell lennie. Ezt azáltal lehet biztosítani, hogy a feladatot földrajzi területek szerint szétosztva, regisztrátor szervezetek kezébe teszik le a domainnév-bejegyzés jogát, akiktől elvárják az egyedi névkiosztás elvének betartását a felelősségükbe tartozó névtartományban (ez a névadási rendszer is hierarchikus szerveződésre épül).

A domainnévtér fenntartására első szinten azért van szükség, hogy egy új névbejegyzési igény esetében lehessen tudni, létezik-e az igényelt domainnév a névtérben vagy sem. Ha a névtérben már bent levő névről van szó, akkor nem lehet bejegyezni, ha az igényelt név még nincs a névtérben, akkor bejegyezhető (ha a névadási szabályoknak megfelel). Mit jelent a bejegyzés? A domainnevek és az őket bejegyző (és innentől fogva birtokló) névtulajdonosok közti kapcsolat rögzítését. A másik fontos funkciója a domainnévtérnek az, hogy egyértelmű kapcsolatot teremtsen a domainnevek és a gépek által kezelt, gépekhez tartozó IP-címek között.

A domainnévtér tehát hierarchikus rendbe szervezett, egyedi nevek, valamint a nevek és a tulajdonosok, illetve nevek és a fizikai címek közti dinamikus kapcsolatok naprakész nyilvántartása. A domainnevek egyedisége egyfelől biztosítja az általuk "jelölt" IP-címek egyértelmű beazonosíthatóságát, másfelől a regisztráció a domainnév használati jogát is egyértelműen rögzíti. Mivel a domainnévtér legfőbb funkciója a világháló összes kommunikációs csomópontjának egyértelmű azonosítása, nyilvánvaló, hogy ez egy globális névtér. Ez a névtér engedélyezi a szinonimitást, de tiltja a homonimitást, és – a fájlrendszerhez képest kisebb mértékben ugyan, de – használja az összetett név komponálásának eszközét is, hiszen a domainnevek összetevőkre bonthatók, és egyes összetevőkben lehet ugyanazokat a névelemeket is újra alkalmazni.

2.3 Emailcímek

Ahogy a távközlési szolgáltatások döntő részében, úgy az emailcímek világán belül is egy egyértelmű címzési rendszer fenntartására van szükség. Egy-két szemponttól eltekintve mondhatjuk, hogy az emailcímeteket értelmezhetjük a domainnév-rendszer részeként is, hiszen az egyedi emailcímek kiosztását, illetve a levelek eljuttatását a címzethez a hierarchikus módon egyértelműsített domainnevekhez kapcsolódva lehet biztosítani, csak más protokoll mentén lehet az adatokat küldeni és fogadni a hálózaton keresztül. Az emailcím felépítéséhez egy szinttel tovább kell bővíteni a domainnevek hierarchiáját, hiszen amikor az emailcímet létrehozunk, és a domainnév elé egy '@' jelet teszünk, majd az elé még egy karaktersorozatot illesztünk, akkor ezzel az adott domain alatt hozunk létre egy új nevet egy garantáltan egyedi nevekből álló lokális névtéren belül.

i@syi.hu = username: 'i' + szeparátor: '@' + domain: 'syi.hu'

Mivel az emailcímek anoním módon is használhatók, így ez a rendszer csak azt képes biztosítani, hogy az emailcímek egyediek, ezáltal egyértelműen hivatkozhatók legyenek, de más dolgok (például emberek) azonosítására nem alkalmasak. Ennek ellenére a web világában egyre több szolgáltatás alkalmazza a felhasználók azonosítására az általuk megadott és használt emailcímet.

2.4 Telefonszámok

A telefonszámok rendszerére is ráhúzható, hogy általában egy egyértelmű címzést lehetővé tevő címteret alkotnak, amelyben két tetszőleges pont közötti privát, címzett kommunikációt pont az teszi lehetővé, hogy a számok egyediek, bár az egyedi címek kiosztásának elvét és gyakorlatát megsértő ritka kivételként hivatkozni lehet az államszocializmus idején létezett ikertelefonok jelenségére, amikor ugyanazt a telefonszámot két végpontra is bekötötték. A telefónia történelmében azonban ez csak egy ritka kivételnek tekinthető, hiszen az elvárt szolgáltatásminőség nyilván csak az egyedi címek kiosztásával volt biztosítható. A telefonszámok rendszere – általában – háromszintű, hierarchikus címzési rendszer, amelyben az országok, majd az azon belül definált régiók, körzetek, végül a végpontokba bekötött egyedi telefonszámok különíthetők el.

36 30 333 1234 = 36 (országkód) + 30 (körzetkód) + 333 1234 (telefonszám)

A címteret, a telefonszámok címtérében „megbúvó” hierarchiát tovább lehet bővíteni azzal, hogy telefonmellék felvételét engedjük meg az adott telefonszám alatt, aminek elérését valamilyen továbbkapcsolási technikával biztosítunk.

A telefonszámok egyértelműen nem kapcsolatok össze a telefont használó emberrel, de a lehetőség adott (különösen a mobiltelefonok esetében), és egyre több olyan kezdeményezést, kísérletet láthatunk a világban, amelyeknek célja az, hogy a mobilszámokat személyek, illetve az általuk kezdeményezett hálózati tranzakciók azonosítására használják.

2.5 Földrajzi helyek

A földrajzi helyeket a legelvontabb szinten három típusra bonthatjuk: beszélhetünk pontszerű, vonalszerű és poligonszerű objektumokról. Ez a felosztás csak gyakorlati szempontból tartalmaz. Elméletileg minden lokáció poligonszerű, és csak a felbontás adott pontossága mellett beszélhetünk pontszerű vagy vonalszerű objektumokról, a felbontás mélységét növelve mindig kiderül, hogy ezeknek is kétdimenziós kiterjedése van. A földrajzi térben való tájékozódás megkönnyítése végett az ember ősidőktől fogva nevet adott a környezetének. Mindhárom geotípus esetén található neveket, amelyeket névterekbe lehet rendezni.

2.5.1 Postai címek

A földrajzi, postai címek, valamint az ingatlannyilvántartás rendszere is értelmezhető címtérként, hiszen ezek kialakítását, fenntartását a földrajzi helyek, a lakcímek, az ingatlanok egyértelmű azonosításának igényével magyarázhatjuk. Ezekre az azonosítási, egyértelműsítési feladatokra azonban elég erős kizárólagossági igény fogalmazódik meg az államigazgatási szféra irányából, hogy ezen a terepen nem várható valódi együttműködési szándék, így nem érdemes elmélyülni ezen a területen.

2.5.2 Földrajzi térképek, földrajzi nevek

A földrajzi navigáció kiegészítő eszköze régóta a térkép, amely egyfelől geometriai adatokat, másfelől névadatokat tartalmaz. A predigitális világban a térképekhez olykor kapcsoltak földrajzinév-tárakat (gazetteereket), amelyek a földrajzi helyek kétdimenziós vizuális reprezentációját kiegészítő alfanumerikus adatokat, neveket, számokat, statisztikákat tettek elérhetővé. A térképek, földrajzinév-tárak névszerű adatkomponensei esetében nyugodtan beszélhetünk névtérről, még ha ezek kortárs használói nem is így gondoltak rájuk. A földrajzinév-tárak, a térképi segédletek a térképeken, a földrajzi térben való eligazodást a helyek és nevek egyértelmű kezelésével, azonosításával tudták igazán támogatni. Ezek a földrajzi névterek műfajilag lehetnek többfélék. Van köztük „sima” adatbázis, van teaurusz, van térképszolgáltató.

2.5.2.1 GeoNames

A GeoNames földrajzi adatbázisban több, mint 11 millió név szerepel a világ összes országából. Az adatok letölthetők, szabadon használhatók (geonames.org). Az adatokhoz egy egyszerű keresődoboz segítségével elindított kereséssel lehet hozzájutni.

The GeoNames geographical database covers all countries and contains over eleven million placenames that are available for download free of charge.

enter a location name, ex: "Paris", "Mount Everest", "New York"

Browse the names	Information	Download
<ul style="list-style-type: none">CountriesPostal codesWikipediaCountry statisticsRecent modifications	<ul style="list-style-type: none">About GeoNamesData SourcesUser manualAmbassadors and TeamForumBlogMailing listCommercial Support and Consulting	<ul style="list-style-type: none">InfoFree Gazetteer DataFree Postal Code DataPremium Data

Sponsoring

<ul style="list-style-type: none">Dealspotr Promo CodesChhabra Law - IP Attorneys	<ul style="list-style-type: none">PromoCodeWatch, LLC	<ul style="list-style-type: none">Activeinstafollowers.comDonations and Sponsoring
--	---	---

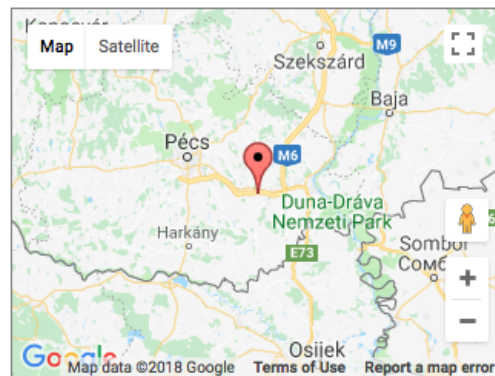
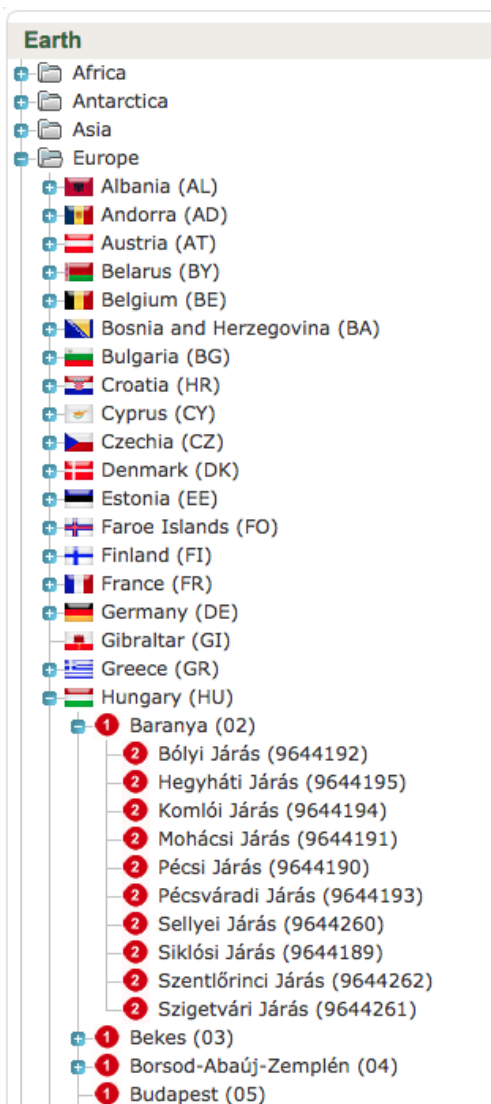
Web Services

- [Overview](#)
- [Documentation](#)
- [Client Libraries](#)
- [Premium Web Services](#)



a GeoNames nyitóoldala

A geoNames a találatokat sokféle formában képes megmutatni: ki tudja tenni térképes felületre, mutatni képes a kiválasztott földrajzi hely kapcsolatrendszerét, de hierarchikus nézetbe is be tudja rendezni a földrajzi helyeket azok tartalmazási kapcsolatai alapján. Hasznos szolgáltatás ezen az oldalon az, hogy az egérrel a hierarchikus lista elemeire mutatva a rendszer megjeleníti a lista mellett elhelyezett térképen a kiválasztott lokáció pozícióját a hely tágabb környezetén belül. Érdekes még azt is megjegyezni, hogy a megjelenítést mennyire képes feldobni az a megoldás, hogy az országokhoz, régiókhoz, adott esetben az alacsonyabb szintű helyekhez a nevek mellé elhelyezik a helyhez tartozó színes zászlókat, címereket.



GeoNames GeoTree hierarchikus névmegjelenítője térképes gyorsmutatóval

2.5.2.2 Getty TGN

A Getty földrajzi nevek tezaurusza (Getty Thesaurus of Geographic Names – Getty TGN) szintén a világ teljes földrajzi névkészletét igyekszik összegyűjteni. Ebben a rendszerben az elemek közti struktúrát nem a földrajzi-geometriai adatok alapján felépíthető partitív kapcsolatok adják (amit térképen lehet a legjobban reprezentálni). Ahogy a neve is mutatja, a Getty TGN tezaurusz, vagyis az egységek között többféle relációt megengedő, szabványos tudásszervezési rendszer, amiben túlsúlyban vannak a földrajzi tulajdonnevek.

Explore the Getty | Getty360 | Calendar | Blog | Connect with Us | Shop | Support Us

The Getty Research Institute

Exhibitions & Events | Special Collections | Library | Search Tools & Databases | Scholars & Projects | Publications | About the GRI

Print | Share

Search Tools & Databases

- Primo Search
- Getty Research Portal
- Collection Inventories & Finding Aids
- Photo Archive
- Research Guides & Bibliographies
- Digital Collections
- Article & Research Databases
- Collecting & Provenance Research
- BHA & RILA
- Getty Vocabularies

Getty Thesaurus of Geographic Names® Online

Search the TGN [Help](#)

Find Name or ID: [Search](#)

Place Type: [Clear](#)

[Lookup](#)

Nation:

[Lookup](#)


[Pop-up Search](#) | [Browse the TGN hierarchies](#)

[Have a Question?](#)

[Contact the Vocabulary Program](#)

a Getty TGN kezdőlapja

A szabványos tezaurusrelációkat nem jelentik meg az egyébként szokásos módon, bár nyilván ki lehet találni, mikor milyen reláció mentén kapcsolják össze az adataikat. A hangsúlyt a partitív relációra teszik, a keresési feltételek alapján kiadott találati oldalon a legszélesebb terjedelmű 'világ' (Föld) fogalmától kiindulva felsorolják a partitív hierarchiába tagolható egységeket, a legutolsó elemként a keresett egységet megjelenítve.



Research

Research Home | Tools | Thesaurus of Geographic Names | Hierarchy Display


Getty Thesaurus of Geographic Names® Online






Hierarchy Display

[New Search](#) | [Previous Page](#) | [Help](#)

[Vernacular Display](#) | **English Display**

[View Selected Records](#) | [Clear All](#)

Click the  icon to view the hierarchy.
Check the boxes to view multiple records at once.


-  Top of the TGN hierarchy (hierarchy root)
-  World (facet)
-  Europe (continent)
-  Hungary (nation)
-  Pécs (urban county)

[New Search](#)

[Back to top](#)

a Getty TGN hierachikus névmegjelenítője 'Pécs'-re keresve

Az adott lexikai egység teljes kapcsolatrendszerét mutató oldalon látszik, hogy felveszik a földrajzi hely geometriai adatait, neveit, típusbesorolásait, a partitív hierarchián belüli helyét (ebből többet is megadhatnak), az adatok forrását, olykor megjegyzéseket, további rokonsági kapcsolatokat.

Click the  icon to view the hierarchy.

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

ID: 7594106

Record Type: administrative

 **Pécs (urban county)**

Coordinates:

Lat: 46 06 00 N *degrees minutes* Lat: 46.1000 *decimal degrees*

Long: 018 15 00 E *degrees minutes* Long: 18.2500 *decimal degrees*


Names:


Pécs (preferred, C, V)


Pecs (C, V)


Pécs Megyei Város (C, V)

Hierarchical Position:

 World (facet)

 Europe (continent) (P)

 Hungary (nation) (P)

 Pécs (urban county) (P)

Place Types:

urban county (**preferred**, C)

first level subdivision (C)

Sources and Contributors:

Pécs..... [VP Preferred]

..... [NGA, GEONet Names Server \(2008-\)](#) accessed 17 July 2014

Pécs Megyei Város..... [VP]

..... [NGA, GEONet Names Server \(2008-\)](#) accessed 17 July 2014

Pecs..... [VP]

..... [CIA, World Fact Book \(2006-\)](#) accessed 17 July 2014

Subject: [VP]

..... [NGA, GEONet Names Server \(2008-\)](#) accessed 17 July 2014

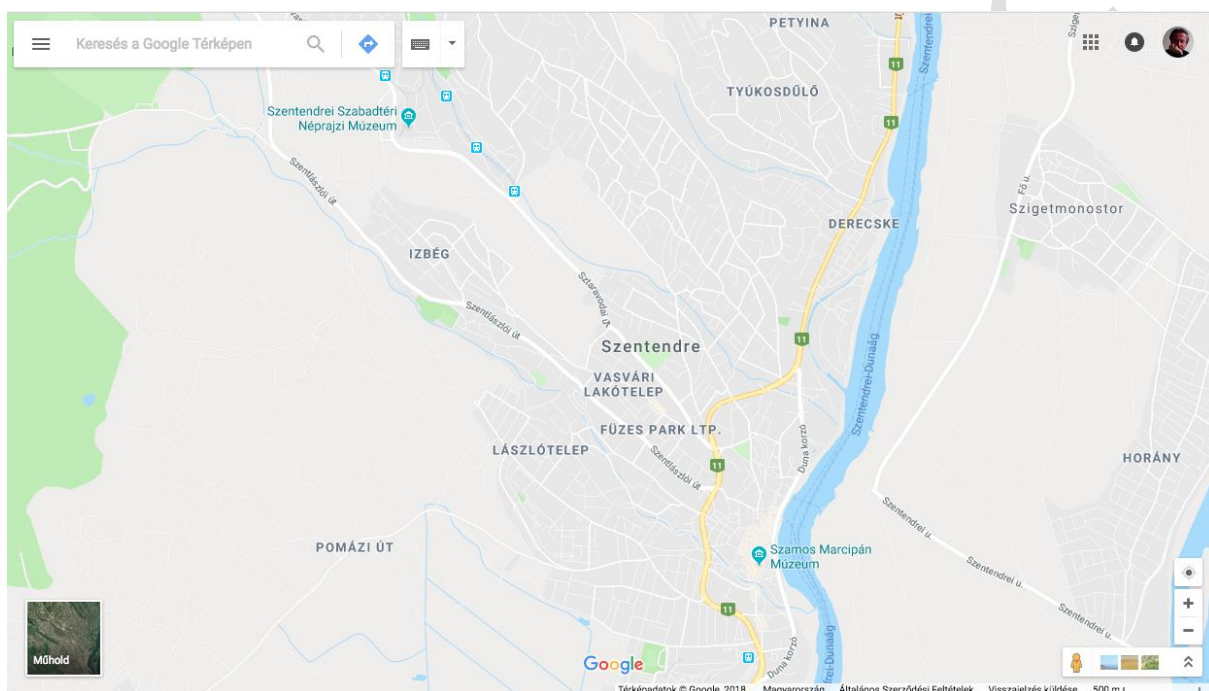
Pécs adatai a Getty TGN megjelenítő felületén

Mind a GeoNames, mind a Getty TGN alapvetően nyelvi alapon, a karakteres adatokra támaszkodva (a gazeteer mintájára) stuktúrálja a földrajzi helyekre vonatkozó adatokat. A hagyományos világ hagyományos technológiái csak ezt a rendezést tudták biztosítani. A digitális világ azonban felkínálja a rendelkezésre álló adatok geometriai alapú összerendelését és ezzel megteremti annak lehetőségét, hogy a földrajzi nevek gyűjteményét térképes alapon szervezzék egybe, amikor már a karakteres adatok lesznek a kiegészítő jellegűek. Két példát érdemes itt említeni, a Google Maps (maps.google.com) és az OpenStreetMap (osm.org) szolgáltatásait.

2.5.2.3 Google Maps

A Google Maps – az anyavállalat brutális pénzügyi és intellektuális erejére támaszkodva – rövid idő alatt piacvezető szolgáltatássá nőtte ki magát a saját műfajában. A szolgáltatás célja minél

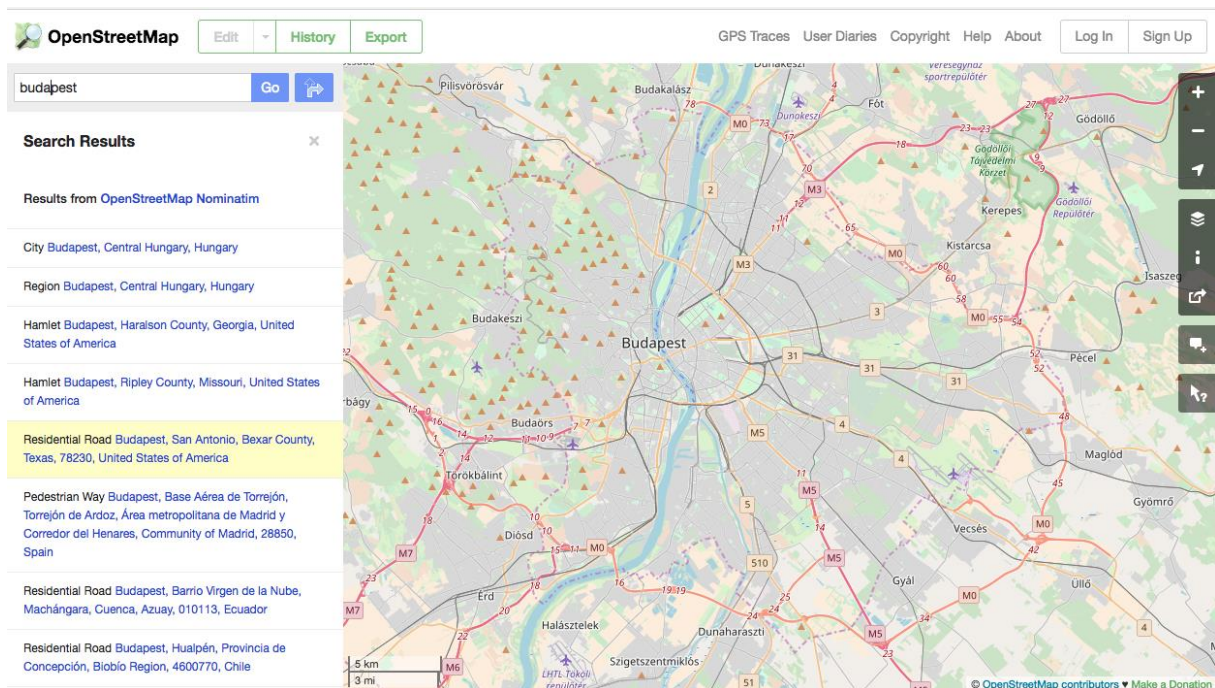
többfajta, minél alaposabb, minél megbízhatóbb navigációs segítséget adni a földrajzi térben eligazodni vágyók számára. Amennyire ebből a szempontból fontos lehet a földrajzi helyek, nevek egyértelműsítése, annyira fontos a vállalatnak a földrajzi névtér építése, de ezt – vélhetőleg – nem tekintik önmagában vett célnak, és a vállalat nem törekszik arra sem, hogy más szolgáltatókkal összerendelje saját adatállományát. Ennek ellenére hatalmas névállomány áll rendelkezésükre, amelyek a térképhez is vannak kötve. Ez nyilván kiváló alapja lehetne egy nemzetközi földrajzi névtérnek is, de hogy lenne-e bármilyen készítés egy kollaboratív alapon álló, nemzetközi földrajzi névtér felépítésére vonatkozóan a vállalat döntéshozóiban, arról semmi információ nem áll rendelkezésükre, feltételezni pedig inkább azt érdemes, hogy nem lenne túl nagy céges hajlandóság erre.



a Google Maps alapfelülete

2.5.2.4 OpenStreetMap

Amíg a Google Maps (GM) a legnagyobb for-profit térképes szolgáltató, addig az OpenStreetMap (OSM) crowdsourcing alapon állva próbálja meg létrehozni a non-profit világtérkép-szolgáltatást – ígéretes eredményekkel. Az OSM nagy valószínűséggel mindig kevesebbet fog tudni a Google Maps szolgáltatásaihoz képest a két vállalkozás eltérő háttere miatt. A hálózati kollaboráció azonban eddig biztosítani tudta azt, hogy az OSM követni tudja a GM-t a szolgáltatási minőségében, adatgazdagságában, és minden amellettszól, hogy ez így maradhat a jövőben is. Mivel az OSM teljes adatállomány szabadon letölthető és felhasználható, ez a rendszer kiváló alapot biztosíthat akár a földrajzi helyek nemzetközi névtérének felépítéséhez, de az OSM-adatok bármely nemzeti névtér számára is megfelelő kiindulási pontot jelenthetnek. Érdeemesnek tűnik állandó, kétirányú kapcsolatrendszert kiépíteni az OSM és a nemzeti földrajzi névtér között, hogy az adatváltások mindkét irányban nyomon követhetők és átvezethetők legyenek mindkét rendszeren belül.



az OpenStreetMap keresőfelülete

Az OSM mindegyik térinformatikai entitástípuson (pontszerű, vonalszerű vagy poligonszerű³) belül egyedi azonosítót rendel a földrajzi helyekhez, ami miatt lenne esély arra, hogy ezek alapján földrajzi névtérként lehessen használni, értelmezni a rendszert, de ennek a lehetőségnek útjában áll az a tény, hogy minden OSM-eintitást bármikor ki lehet törölni, és persze bármikor újra lehet definiálni, vissza lehet venni a rendszerbe, de egy ilyen működésében semmi sem garantálja, hogy ugyanazok az egyedi azonosítók rendelődjenek ugyanazokhoz az entitásokhoz. Az OSM nem tudja rendszerszerűen biztosítani az azonosítók perzisztenciáját, ezért önmagában nem alkalmas arra, hogy földrajzi névtérként működjön.

2.6 Webes névazonosítási szabványok

A hálózati kommunikáció az adatok szintjén, a gépek közötti közvetlen kommunikációban mindig is címzett, pont-pont jellegű kapcsolatra támaszkodott. Az emberek számára látható, befogadható web világában a hálózaton keresztül elérhető erőforrások egyértelmű azonosítására nem lehetett időben állandó, perzisztens címzési rendszert fenntartani, mert a gyakorlatban mind a kommunikációt kiszolgáló domainnevek, mind a weboldalakra kitett dokumentumok nevei bármikor megváltozhattak. Természetesen a dokumentumok, erőforrások adott időpillanatban való egyértelmű címzését, elérhetőségét meg lehet oldani az erre szolgáló URI-techikák segítségével.

³ Az OSM-ben a poligonokat relációnak nevezik.

A Uniform Resource Identifier (URI) segítségével a webes erőforrásokhoz – meghatározott szabályok alapján – egyedi címet (karakter sorozatot) lehet rendelni, amely egyedisége révén azonosítani tudja a címmel jelölt erőforrást, dokumentumot.

```
scheme://[user[:password]@]host[:port]][/path][?query][#fragment]
```

Az URI maga csak egy fedőfogalom, az általános séma rögzítésén túl nincs konkrét szabályrendszer hozzákapcsolva, a tényleges azonosítást csak az URI két típusához, az URN-hez és az URL-hez kidolgozott technikák segítségével lehet megvalósítani.

A Uniform Resource Locator (URL) segítségével a domainnevek rendszerére támaszkodó webcímek a megjelölt protokollon (pl. http- vagy ftp-protokollon) keresztül azonosítják egyértelműen és teszik elérhetővé az erőforrásokat. Ez az azonosító technika az erőforrások elérhetőségén, pontosabban az erőforrásokról szóló dokumentumok címén keresztül biztosít hozzáférést az erőforráshoz. A dokumentumok címe itt a digitális tér adott pontjára utal. Az URL egy példája lehet a következő cím.

```
https://en.wikipedia.org/wiki/URL
```

A másik URI-technika, az egységes forrásmegnevezés, Uniform Resource Name (URN) segítségével nem a digitális tér valamely pontjára, tehát az erőforrás helyére tudunk rámutatni, hanem az erőforrás nevét tudjuk megadni. Ehhez definálni kell egy névtérre, és meg kell adni azt az azonosítót, amellyel egyértelműen rá lehet mutatni az erőforrásra a névtéren belül. Az URN-be bármilyen azonosítórendszer beilleszthető. Az RFC 2141 internet-szabvány értelmében minden URN azonosító három részből tevődik össze: az „urn” karaktersorból, a helynév-azonosítóból (Namespace Identifier, NID) és a helynév egyedi megjelöléséből (Namespace Specific String, NSS). Az NSS tartalmazza a felhasznált azonosítót, pl. az ISMN-t, a NID pedig a felhasznált azonosító típusának szabványos elnevezését. A könyvek világában használt ISBN-számokra az alábbi példát lehet adni.

```
urn:isbn: 963-05-3268-9
```

Ez az URN rámutat az ISBN névtérre, és azon belül a Kenyeres Ágnes főszerkesztő vezetésével összeállított 'Magyar életrajzi lexikon' című könyv 1984-es, 4. kiadását azonosítja.

Az URN segítségével vissza lehet keresni magát a dokumentumot, annak leírását, vagy azon URL-ek listáját, amelyekben a dokumentum megtalálható. A felhasználó szemszögéből az URN használata egyszerű: a böngésző keresőablakába az URL helyett az URN-t kell begépelni. Az URN is a HTTP-protokollra és a DNS-szolgáltatásra épül.

2.7 Személynevek

A technikai példák után nézzük meg, hogy a személyneveket miért és hogyan használjuk, és hogyan lehetne személynevek esetében névteret építeni. Egy családon belül az eltérő keresztnévek használatának nyilván ugyanaz az értelme, mint a fenti két példában: az egyedi nevek használatával az egyértelmű hivatkozás lehetőségét teremtjük meg a nevekkel jelölt dolgokra, a gyerekekre a családon belül. Egy család esetében ehhez elég abból a párezres keresztnév-állományból neveket választani, amelyek az idők során elfogadottakká váltak, de a családi közösségeken túllépve már nyilvánvalóan alkalmatlan lehet egy ilyen keresztnévtér minden gyerek azonosítására. Ebből a szempontból kicsit segíthet az összetett nevek képzése, amikor a személy teljes neve a családnév és a keresztnév kettőséből áll, de tudjuk jól, hogy ezzel a megoldással sem biztosítható az, hogy minden embernek egyedi neve legyen. A személynevek segítségével csak akkor tudjuk egyértelműen azonosítani az embereket, ha az emberek körét meghatározzuk vagy másként: pontos határvonalú megnevezési kontextust teremtünk. Mit is jelent ez? A személyek esetében szűk kontextusokban (családi, esetleg nagycsaládi, házközösségi, nemzetiségi, falusi szinten) még remélhető, hogy a család- és keresztnévek segítségével minden ember azonosítható legyen, de minél tágabb környezetben reméljük az egyértelmű azonosítást, annál biztosabban nem fog sikerülni, mert a homonimitás jelensége egyre gyakoribb lesz.⁴

A 'Kovács János' név már nem képes egyértelműen azonosítani egy embert, ha a faluban van több 'Kovács' nevű család, ahol kiosztják a 'János' keresztnévet valamelyik fiúknak. A példa rávilágít a megnevezési kontextus fontosságára. A családon mint névadási kontextuson belül még biztosítható a névtér működése, mindenkinek lehet egyértelmű azonosításra alkalmas nevet adni. Ezt a kontextust tágítva egy darabig még fent lehet tartani a homonimitásmentes állapotot, de a bővülés elérhet egy határt, amin túl már nem lehet biztosítani azt, hogy ugyanaz a név ne mutasson több személyre. Ekkor a névtér használhatatlanná válik, pontosabban részlegesen használhatatlanná válik, hiszen maradhatnak olyan szegmensei, amelyekben még képes jól funkcionálni, de igazából csak úgy lehet visszaállítani a névtér funkcionalitását, ha a szükséges méretűre szűkítjük a névadási kontextus terjedelmét.

A személynévtér példája arra mutat rá, hogy a nevek önmagukban nem alkalmasak azonosításra, csak a kontextus és nevek együttes alkalmazásával remélhetjük a névterek sikeres működését.

Megjegyezzük, hogy a személyek azonosítására az államigazgatás már régóta több megoldást is kidolgozott (személyi igazolvány, útleveél, személyi szám), ezekkel azonban itt nem foglalkozunk.

2.8 Műcímek

A közgyűjtemények, a dokumentumok világában régtől fogva komoly igény volt a nemzetközi azonosító rendszerek létrehozására a különféle dokumentumtípusok esetében. Az effajta

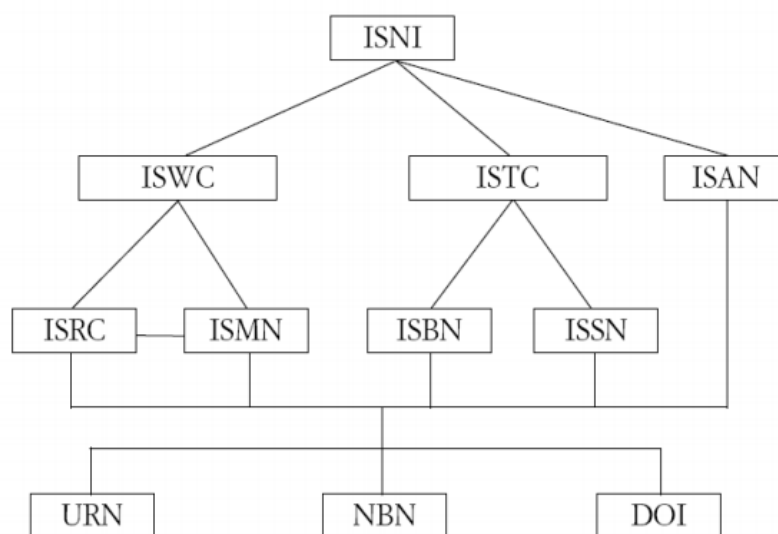
⁴ A szinonimitás itt sem okoz gondot. Egy embernek lehet több neve is, ha azok egyediek, mindegyikükkel azonosítani lehet az adott személyt.

igények létezését, erejét jelzi az a tény, hogy sok nemzetközi szabvány jött létre ezen a területen. A névtérépítkezés szempontjából nézve az alábbi szabványok lehetnek különösen fontosnak.

név- és címazonosítási nemzetközi szabványok

ISO 27729	International Standard Name Identifier	ISNI
ISO 21047	International Standard Text Code	ISTC
ISO 2108	International Standard Book Number	ISBN
ISO 3297	International Standard Serial Number	ISSN
ISO 15706	International Standard Audiovisual Number	ISAN
ISO 15707	International Standard Musical Work Code	ISWC
ISO 3901	International Standard Recording Code	ISRC
ISO 10957	International Standard Music Number	ISMN
ISO 26324	Digital Object Identifier System	DOI

Ezen szabványok a szöveges, az audiovizuális és az auditív dokumentumok kezelésével kapcsolatos elvárásokat rögzítik. A szabványok közti kapcsolatokat szemlélteti az alábbi ábra.



a szabványok rendszere

A szabványok fókuszja irányulhat a dokumentumok, művek létrehozójára, gondozójára, az emberre, illetve a műre mint kreatív alkotás tartalmára vagy a mű tartalma alapján készült – általában sokszorosított – dokumentumra mint kiadványra. Ennek megfelelően a fenti szabványok az alábbi három csoportba sorolhatók.

A szerzőt (létrehozót) azonosítja:

- ISNI: International Standard Name Identifier

A művet azonosítja:

- ISWC: International Standard Musical Work Code
- ISTC: International Standard Textual Work Code
- ISAN: International Standard Audiovisual Number

A kiadványt azonosítja:

- ISMN: International Standard Music Number
- ISBN: International Standard Book Number
- ISSN: International Standard Serial Number
- ISRC: International Standard Recording Code

Ezeket a szabványokat a nemzetközi szabványügyi világszervezet gondozza, és saját (ISO) számokat rendel hozzájuk. A digitális világban vannak olyan szabványok, amelyeket az ISO bocsátott ki, de ettől még nagy nemzetközi elfogadottsággal rendelkeznek. A digitális világ ilyen szabványai közül azok érdekesek számunkra, amelyek a digitális tartalmakat, forrásokat azonosítják.

- URN: Uniform Resource Name
- NBN: National Bibliography Number
- DOI: Digital Object Identifier

A hivatalos szabványosítási tevékenység egyre több területre terjed ki, és várható, hogy olyanok is megjelennek, elterjednek, amelyek érdekesek lehetnek a nemzeti névtérépítések számára is. Jelen helyzetben úgy tűnik, hogy a fenti szabványok lehetnek rövid távon is meghatározók a névterek számára, és majd az idő dönti el, hogy az alábbi listában látható, a hozzájuk hasonló szabványok közül melyikkel kell majd behatóbban foglalkozni.

azonosítással kapcsolatos, további nemzetközi szabványok

ISO 6166	International Securities Identification Number	ISIN
ISO 7812	Issuer identification number	IIN
ISO 13616	International Bank Account Number	IBAN
ISO 15511	International Standard Identifier for Libraries and Related Organizations	ISIL
ISO 17316	International Standard Link Identifier	ISLI
ISO 17442	Legal Entity Identifier	LEI

2.8.1 Könyvek, folyóiratok

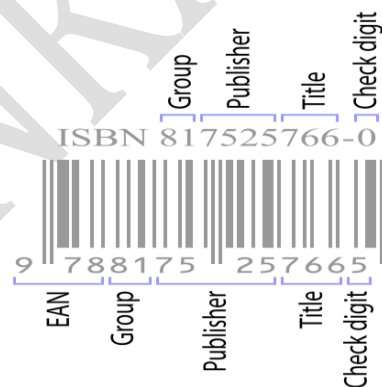
2.8.1.1 ISBN

Az ISBN (International Standard Book Number) a könyvek nemzetközi azonosítója. Előfordul, hogy zenei kiadványokat a könyvkereskedelmi hálózaton keresztül terjesztenek. Ha egy kiadó a könyvpiacra szánt kottát jelentet meg, az ISMN mellett ISBN azonosítóval is elláthatja. Olykor nehéz eldönteni, hogy egy kiadvány (pl. egy daloskönyv, egyházi énekeskönyv vagy egy album terjedelmes szöveges és képi anyaggal) vajon zenei kiadvány-e, "közönséges" könyv, vagy mindkettő. Ilyen esetben mind ISMN-nel, mind ISBN-nel el lehet látni a kiadványt. Ha a kiadó csak egyfajta azonosítót kíván alkalmazni, a választás az ő megítélésére van bízva, de kétség esetén az ISMN-t előnyben lehet részesíteni.

Az ISBN alkalmazásának szabályait az ISBN Útmutató tartalmazza. Az ISBN rendszer irányítását a Nemzetközi ISBN Ügynökség és a nemzeti ügynökségek végzik. Magyarországon a Magyar ISBN és ISMN Iroda. Amikor egy kiadványt ISMN-nel és ISBN-nel is ellátnak, mindkét azonosítót fel kell rajta tüntetni, és egyértelműen meg kell különböztetni őket egymástól.

Az ISBN vagy ISBN-szám (International Standard Book Number) 13 jegyű (a 2007. január 1. előtt kiadott könyveknél 10 jegyű) azonosítószám, a könyvek és egyéb monografikus jellegű művek nyilvántartására szolgáló nemzetközi szabványos számrendszerhez tartozó kód.

ISBN 0-306-40615-2



az ISBN vonalkódja

Az ISBN ötletét W. H. Smith vetette fel először 1965-ben, amikor könyvkereskedését egy számítógépekkel felszerelt raktárépületbe költöztette. 1966-ban készítették elő a Standard Book Numbering (SBN) elnevezésű rendszert, mely így 1967-ben bevezetésre kerülhetett az Egyesült Királyságban.

Az ISO tárgyalásokat kezdett 1968-ban Londonban Dánia, Franciaország, Németország, Írország, Hollandia, Norvégia, Nagy-Britannia, az Amerikai Egyesült Államok és az UNESCO megfigyelői részvételével a brit SBN rendszer nemzetközi bevezetéséről. 1969-ben a berlini és stockholmi

ülések során született döntés az ISBN 1970-es nemzetközi bevezetéséről. Az ISBN lett az ISO 2108 szabvány.

Ezt a rendszert használta 2007. január 1-jéig közel 150 ország, amikor az ISBN számjegyeinek számát 10-ről 13-ra emelték. Magyarország 1974 óta tagja az ISBN-rendszernek, országkódja 963. Az azonosítót az Országos Széchényi Könyvtárban működő Magyar ISBN Iroda osztja ki.

A következő kiadványok kaphatnak ISBN-t:

- Nyomtatott könyvek, brosúrák;
- Braille kiadványok;
- Térképek;
- Oktatási célú video- és diafilmek, szoftverek;
- Kazettán, CD-n vagy DVD-n kiadott hangoskönyvek;
- Elektronikus kiadványok, akár fizikai hordozón, akár online hozzáféréssel;
- Multimédiás kiadványok, amennyiben a tartalom túlnyomó része szövegalapú.

Az ISBN első számcsoportja a GS1 által kinevezett 978 vagy 979 előtag; a második számcsoport a könyv származási országára vagy nyelvére utal, a harmadik a könyvkiadót jelöli, a negyedik a könyv azonosítására, az ötödik (ez csak egy számjegyből áll) az egész számsor matematikai ellenőrzésre szolgál. Az ISBN-ekhez automatikusan hozzárendelhető a hasonlóan szabványos, gépek által leolvasható vonalkód is. (A 979-es előtag 2008-as bevezetése érintette az ISMN rendszerét, amelyek a 979-0 számcsoportba kerültek.)

A könyvsorozatokat és az időszaki kiadványokat ún. ISSN (International Standard Serial Number) számmal azonosítják. Ez kétszer négy, egymástól kötőjellel elválasztott számcsoportból áll, amelyek az ISBN-nel szemben nem bírnak jelentéssel (nem utalnak országra vagy kiadóra).

Az ISBN – és sorozatok esetén az ISSN – feltüntetése a könyvben kötelező, ezek jellemzően a copyright- vagy a kolofonoldalra kerülnek. A mindig nagybetűkkel írandó ISBN/ISSN mozaikszavakat követő számcsoportokat kötőjelek vagy szóközök választják el egymástól.

Egy könyv mindegyik kiadása és változata (kivéve az utánnyomást) egyedi ISBN azonosítót kap. A szám lehet 13 vagy 10 (2007. január 1. előttiék) számjegyű, és az alábbi részekből áll:

- ha 13 számjegyű ISBN, akkor egy GS1 prefix, 978 vagy 979,
- a származási ország vagy egy nyelvi területet jelölő kód,
- a kiadó,
- a termék száma, és
- egy ellenőrző karakter.

A különböző részek különböző hosszúságúak lehetnek, kötőjellel elválasztva. A jelenleg érvényes, 2006-ban megjelent ISBN útmutató szerint „az ISBN öt számcsoportból áll, melyek közül az elsőnek és az utolsóknak a hosszúsága kötött, a középső háromé változó. A számcsoportokat kötőjel választja el egymástól.”

A második számcsoport 0 vagy 1 az angol anyanyelvű országokban, 2 a francia, 3 a német, 4 a japán, 5 az orosz anyanyelvű országokban stb. Az eredeti SBN-ből hiányzott az országazonosító, de elérve egy 0-t a 9 számjegyű SBN helyes ISBN kóddá alakul át. Az ország mező 5 karakter hosszú lehet; a 99936-s Bhután számára van fenntartva.

A kiadó azonosítóját a helyi ISBN ügynökség állapítja meg. A termékszámot vagy maga a kiadó, vagy szintén az ügynökség adja ki. (Magyarországon jelenleg ez utóbbi a gyakoribb.) Nem kötelező minden kiadó hogy ISBN azonosítót rendeljen kiadványaihoz - mint például Kína -, mert az ISBN-re vonatkozó szabvány (MSZ ISO 2108:2005) alkalmazása önkéntes, de a legtöbb könyvesbolt csak ISBN azonosítóval ellátott kiadványt hajlandó forgalmazni.

A kiadók kiadói azonosítókat (más néven kérelmezőelemeket) kapnak. Minél rövidebb a kérelmezőelem, annál több karakter marad a kiadványok azonosítására. Így megeshet, hogy egy kisebb kiadó például 5 karakteres értelmezőelemet kap, de csak 1 karakteres termékazonosítót. Ha egy kérelmezőelem betelik, újat lehet igényelni. Ebből következően több különböző szám jelölheti ugyanazt a kiadót.

Az ellenőrző karakter

Számos irányelv létezik az ellenőrző karakter használatáról az ipar és a könyvtárak között. A kiadók olykor elfelejtik ellenőrizni az általuk kiadott kiadvány ISBN-jének érvényességét, ezzel nehéz helyzetbe hozva a könyvtárakat, könyvesboltokat és az olvasókat. A legtöbb könyvtár és könyvesbolt a kiadó által helytelenül megadott ISBN-t használja. A Library of Congress egy listát tart nyilván a hibás ISBN-nel megjelölt könyvekről. Néhány könyvrendelő rendszer, mint például az Amazon.com, nem kereshető érvénytelen ISBN-nel.

Számítása a 10 számjegyű ISBN esetén

Hivatalos kézikönyvében a Nemzetközi ISBN Ügynökség a 10 számjegyű ISBN ellenőrző karakterét - ami a kód utolsó karaktere - a számjegyek 10-től 2-ig történő súlyozásával számolja modulo 11. Ez azt jelenti, hogy mind a kilenc számjegyet megszorozzák egy 10 és 2 közötti szorzóval, majd ezek összegét kiegészítik a legkisebb olyan számra, ami maradék nélkül osztható 11-gyel. A kiegészítés lesz az ellenőrző karakter.

A fenti módszerre példa, ha az első kilenc karakter "0-306-40615":

$$\begin{aligned} & 10 \times 0 + 9 \times 3 + 8 \times 0 + 7 \times 6 + 6 \times 4 + 5 \times 0 + 4 \times 6 + 3 \times 1 + 2 \times 5 \\ & = 0 + 27 + 0 + 42 + 24 + 0 + 24 + 3 + 10 \\ & = 130 \\ & 11 \text{ legközelebbi többszöröse: } 12 \times 11 = 132 \\ & 132 - 130 = 2 \end{aligned}$$

Tehát az ellenőrző karakter a 2, és a teljes számsor az ISBN 0-306-40615-2.

A másik módszer szerint a számjegyeket megszorozzák a számjegy pozíciójával a számsorban (ez egy 1-9 közötti érték), majd ezek összegét kiegészítik a legkisebb olyan számra, ami maradék nélkül osztható 11-gyel. A kiegészítés lesz az ellenőrző karakter.

A fenti módszerre példa, ha az első kilenc karakter "0-306-40615":

$$\begin{aligned} & 1 \times 0 + 2 \times 3 + 3 \times 0 + 4 \times 6 + 5 \times 4 + 6 \times 0 + 7 \times 6 + 8 \times 1 + 9 \times 5 \\ & = 0 + 6 + 0 + 24 + 20 + 0 + 42 + 8 + 45 \\ & = 145 \\ & = 13 \times 11 + 2 \end{aligned}$$

Tehát az ellenőrző karakter a 2, és a teljes számsor az ISBN 0-306-40615-2.

Látható, hogy a 10 számjegyű ISBN ellenőrző karakterét a számjegyek skaláris szorzataként számoljuk modulo 11. $(1,2,3,4,5,6,7,8,9) \times (x,x,x,x,x,x,x,x,x) \bmod 11 =$ ellenőrző karakter.

A két leggyakoribb hiba az ISBN használatakor (például gépeléskor vagy íráskor) a hasonló vagy egymás mellett álló számjegyek felcserélése. Mivel a 11 prímszám, az ISBN ellenőrző karakter számítási módszere minden esetben jelzi az említett két hibát. Ha a hiba a kiadónál keletkezik, a könyvet hibás ISBN azonosítóval is kiadhatják.

Számítása a 13 számjegyű ISBN esetén

A Nemzetközi ISBN Ügynökség hivatalos 2005-s kézikönyve tartalmazza a 2007. január 1-jétől használandó 13 számjegyű ISBN azonosítók ellenőrző karakterének számítási módszerét.

Az ellenőrző karakter számításához először mind a 12 számjegyet (13 karakteres ISBN, kivéve az utolsó ellenőrző karakter) eggyel vagy hárommal kell megszorozni. Ezeket összegezve modulo 10, majd a maradékot kivonva 10-ből áll elő az ellenőrző karakter. Ha 10 az eredmény, akkor 0 használandó. Például, a 978-0-306-40615-__ számítása:

$$\begin{aligned} & 9 \times 1 + 7 \times 3 + 8 \times 1 + 0 \times 3 + 3 \times 1 + 0 \times 3 + 6 \times 1 + 4 \times 3 + 0 \times 1 + 6 \times 3 + 1 \times 1 + 5 \times 3 \\ & = 9 + 21 + 8 + 0 + 3 + 0 + 6 + 12 + 0 + 18 + 1 + 15 \\ & = 93 \\ & 93 / 10 = 9 \text{ maradék } 3 \\ & 10 - 3 = 7 \end{aligned}$$

Tehát az ellenőrző karakter a 7, és a teljes számsor az ISBN 978-0-306-40615-7.

A 978-0-356-42615-__ számítása:

$$\begin{aligned} & 9 \times 1 + 7 \times 3 + 8 \times 1 + 0 \times 3 + 3 \times 1 + 5 \times 3 + 6 \times 1 + 4 \times 3 + 2 \times 1 + 6 \times 3 + 1 \times 1 + 5 \times 3 \\ & = 9 + 21 + 8 + 0 + 3 + 15 + 6 + 12 + 2 + 18 + 1 + 15 \\ & = 110 \\ & 110 / 10 = 11 \text{ maradék } 0 \\ & 10 - 0 = 10 \\ & 10 \text{ cseréje } 0\text{-ra} \end{aligned}$$

Tehát az ellenőrző karakter a 0, és a teljes számsor az ISBN 978-0-356-42615-0.

Szemben a 10 számjegyű ISBN-nel, ez a rendszer nem minden esetben képes jelezni a helycserélési hibákat.

2.8.1.2 ISSN

Az ISMN és ISBN mellett egy kiegészítő azonosító számrendszer is működik a folytatódó kiadványok (korábban: időszaki kiadványok) számára: az International Standard Serial Number (ISSN). A folytatódó kiadvány egyik típusa az időszaki kiadvány. Ide tartozik – definíciója szerint – minden olyan kiadvány, tekintet nélkül hordozójára, amely egymást követő részegységekben jelenik meg, általában számozással vagy kronologikus megjelöléssel van ellátva, és előre meg nem határozott időtartamra tervezik a folytatását. Az időszaki kiadványokat meg kell

különböztetni a többkötetes kiadványoktól, amelyek meghatározott számú kötet megjelenésével válnak teljessé.

Az időszaki kiadványok közé tartoznak a periodikumok és a sorozatok. Kotta csak elvétve jelenik meg periodikumként, de sorozatként elég gyakran. Mindkét kategóriában ISSN-nel kell azonosítani az időszaki kiadvány címét (amely a periodikum összes részegységén, illetve a sorozat minden kötetén változatlan marad), és sorozat esetében ISMN-t kell kiutalni a sorozatban megjelenő minden önálló műnek. (Ha a sorozat címe megváltozik, új ISSN megállapítására van szükség.)

A kiadók nem végezhetik maguk az ISSN azonosítók kiutalását. Magyarországon az időszaki kiadványokat megjelentető kiadóknak ISSN-ért a Magyar ISSN Nemzeti Központoz kell fordulniuk. Amikor egy kiadványt ISMN-nel és ISSN-nel is ellátnak, mindkettőt fel kell tüntetni rajta, és egyértelműen meg kell különböztetni őket egymástól.

Az ISSN (International Standard Serial Number) az időszaki kiadványok nemzetközi azonosítója. Az időszaki kiadványok (periodikumok) lehetnek folyóiratok vagy sorozatok. (Sorozatnak minősül az időszaki kiadvány akkor, ha mindig önálló, monografikus művek jelennek meg az egyes részegységekben, és e műveket különböző szerzők írják.) Alkalmazását az ISO 3297 nemzetközi szabvány, illetve annak honosított változata, az MSZ ISO 3297:2000 írja elő.

Az ISSN feladata

Minden ISSN egy adott időszaki kiadványt azonosít. Az ISSN azonosítószámhoz elválaszthatatlanul kapcsolódik az ISSN nemzetközi számítógépes nyilvántartási rendszerének előírásai szerint szabványosított címforma, azaz a kulcs cím. E szoros kapcsolat miatt, ha az időszaki kiadvány címében bekövetkezett változás folytán a kulcs cím megváltozik, a megváltozott címet egy másik ISSN-nel kell azonosítani.

Amennyiben az időszaki kiadvány címe, illetve kulcs címe változatlan marad, miközben megjelenési helye vagy kiadója, közreadója, megjelenési gyakorisága, netán külalakja vagy belső tartalma megváltozik, mindezen változások mellett ugyanaz az ISSN fogja azonosítani a kiadványt.

Saját ISSN-nel rendelkeznek az önálló, periodikusan megjelenő mellékletek, az ágazatok, az alsorozatok stb. Az időszaki kiadvány különféle kiadásváltozatait (másnyelvű kiadásváltozatokat, helyi vagy időbeli kiadásváltozatokat, eltérő hordozón, pl. kazettán, mikrofilmlapon, mágnes- vagy kompaktlemezen, avagy a világhálón megjelenő kiadásváltozatokat) külön ISSN-nel kell azonosítani. Az egyszer már felhasznált ISSN nem alkalmazható újra egy másik címmel összefüggésben!

Az ISSN felépítése

Az ISSN nyolc számjegyből álló numerikus kód, melynek elemei - a könyveket azonosító ISBN-től eltérően - semmiféle jelentést nem hordoznak, csupán azonosító funkciót töltenek be. A nyolcadik számjegy ellenőrző számjegy, mely az előző hétből számítható ki, és 0-tól 10-ig terjedhet. A 10-es számjegyet római szám helyettesíti, azaz: X.

Az ISSN feltüntetése két, egyenként négy számjegyet tartalmazó adatcsoportban történik, melyeket kötőjel kapcsol össze, és az ISSN betűkód előz meg.

Az ISSN hálózat

Az ISSN nemzetközi hálózata jelenleg közel 90 nemzeti központból áll. Minden évben újabb nemzeti központok kezdik meg működésüket a világ különböző pontjain – folyamatosan karbantartott címjegyzékük elérhető a következő címen: <http://www.issn.org/the-centre-and-the-network/members-countries/the-issn...>

A nemzeti központok a saját országukban kiadott időszaki kiadványok ISSN-nel történő azonosításáért és a nemzetközi rendszer előírásainak megfelelő nyilvántartásáért felelősek.

A hálózatot összefogó ISSN Nemzetközi Központ Párizsban működik. E központ végzi a nemzetközi szervezetek, valamint a nemzeti központtal nem rendelkező országok időszaki kiadványainak ISSN-nel való ellátását és nyilvántartásba vételét.

A hálózat célja, hogy a tudomány valamennyi területén világszerte megjelenő időszaki kiadványokról – tekintet nélkül azok megjelenési formájára, fizikai hordozójára – nemzetközi, számítógépes nyilvántartást vezessen. E szervezet tagjai felelősek az ISSN kiutalásáért, nyilvántartásáért és használatának ellenőrzéséért, valamint a világon megjelenő időszaki kiadványokat magába foglaló Regiszter létrehozásáért és karbantartásáért.

Az ISSN regiszter

A párizsi ISSN Nemzetközi Központ a kezelése alatt álló nemzetközi adatbázis alapján közreadja az ISSN regisztert (ISSN Register), amely a világszerte kiutalt valamennyi ISSN azonosító szám hiteles nyilvántartását foglalja magába. Az ISSN regiszter tartalmazza az ISSN hálózat által feldolgozott, karbantartott és továbbított bejelentéseket, melyek tájékoztatnak az időszaki kiadványok azonosításához és leírásához szükséges leglényegesebb bibliográfiai adatokról.

A közel 2 millió rekord 180 országban, 144 nyelven megjelent időszaki kiadványról nyújt információt. A nemzetközi adatbázis évente 60-70 ezer új tétellel bővül, miközben közel 170 000 leírás frissítésére kerül sor, karbantartása folyamatos.

Az ISSN nemzetközi regiszter online elérését az ISSN Portál szolgáltatja. Hozzáférhető továbbá Z39.50 és OAI-PMH protokollal, valamint ISO 2709 csereformátumú nyers adatfájlként. A rekordok formátuma igény szerint lehet ISSN MARC 21, ISSN UNIMARC és MARC XML.

A regiszter 1996-ig mikrofilmlapokon jelent meg, 2004-ig mágnesszalagon, és 2006-ig CD-ROM kiadásban, ISSN Compact címmel. A világhálón ISSN Online néven volt ismert 1998 és 2005 között. 2005 márciusától kezdve az ISSN nemzetközi adatbázis az előfizetők számára felhasználói névvel és jelszóval, vagy IP cím azonosításával érhető el, az ISSN Portálon keresztül.

A nyílt hozzáférésű tudományos forrásokot leíró rekordok 2014 decembere óta leválogatásra kerülnek az ISSN regiszter előfizetési adatbázisából a ROAD program keretében szabadon elérhető adatbázis számára: <http://road.issn.org/en/>.

Szintén szabadon letölthető az ISSN-L tábla, amely a különféle fizikai hordozókon megjelenő kiadványok ISSN azonosítóit teszi visszakereshetővé az ún. „kapcsoló ISSN” („linking ISSN”, azaz ISSN-L) segítségével.

Az ISSN jövője

Az ISSN világrendszer stratégiai célkitűzéseinek meghatározása arra a felismerésre épül, hogy - figyelembe véve a jelenlegi és a prognosztizálható trendeket - az ISSN jövője attól függ, mennyire terjed el használata állandó azonosítóként az Interneten. Az ISSN csak akkor őrizheti meg szerepét és hitelét, ha a hálózat rövid időn belül jelentős mennyiségű elektronikus kiadvány azonosítását tudja felvállalni. E felismerésnek fontos következményei vannak az egész hálózatra nézve - egyrészt az erőforrások biztosításának vonatkozásában, másrészt az új feladatok pontos definiálása, a hálózati központok tevékenységének összehangolása terén.

Az ISSN alkalmazási körének szélesítése gyakorlatilag azt jelenti, hogy a hagyományos értelemben vett, egymást követő részegységekben megjelenő időszaki kiadványokon kívül a világhálón elterjedt integráló típusú bibliográfiai források is azonosíthatók ISSN-nel (meghatározott formai kritériumok megléte mellett). Ez utóbbiak között lehetővé válik az adatbázisok ISSN-nel történő ellátása is: célszerűnek látszik az előfizethető adatbázisok nyilvántartásba vétele, majd a bibliográfiai és cikkadatbázisok regisztrációja, valamint a kutatók számára fontos, egyes tudományterületekre szakosodott adatbázisok számbavétele, végül a rendszeresen frissített cím- és cégadatbázisok leírása is.

Az elektronikus és azon belül a távoli hozzáférésű bibliográfiai források számbavételének folyamata már a 20. században, a 90-es évek közepén elindult, és a helyi lehetőségek, kapacitás és támogatottság függvényében halad azóta előre, egyre inkább gyorsuló ütemben. 2015-re mintegy 155 ezer folytatólagos online forrás rekordját tartalmazza a nemzetközi ISSN adatbázis – köztük 1600 rekordot a Magyar ISSN Nemzeti Központ feldolgozásai nyomán. A MARC rekord 856-os mezőjében megadott távoli elérési címek révén az ISSN adatbázis felhasználói, illetve előfizetői világszerte tudomást szereznek a világhálón publikált időszaki kiadványokról.

2.8.1.3 ISTC

Az International Standard Text Code (ISTC) egyedi azonosító szöveg alapú művek számára. Az ISO 21047:2009 szabványt a TC 46/SC 9 testület dolgozta ki, 2009 márciusában publikálták. A szabvány alkalmazásáért felelős testület az International ISTC Agency. Az ISTC-kódokat, a hozzájuk tartozó metaadatokkal együtt egyetlen adatbázisban rögzítik, és ezt az adatbázist teszik kereshetővé, szüretelhetővé.

Az ISTC-azonosítókat be lehet írni a művek bibliográfiai rekordjaiba, a webáruházak katalógusaiba, bárhová, ahol a mával kapcsolatos információt tárolnak, és ezek után a közös ISTC-kód alapján a szétszórt adatok összekapcsolhatók egymással. Az ISTC-azonosító segíthet az azonos című, de más tartalmú művek megkülönböztetésében, vagy épp olykor a más cím alatt megjelenő, ugyanazon művek összekapcsolásában. További előnyök származhatnak az ISTC-kódok alkalmazásából a származtatott művek összekapcsolásakor. Ha például a lefordított mű bibliográfiai rekordjába beleteszik az eredeti mű ISTC-kódját is, akkor ezek automatikusan összekapcsolódnak, és így további szolgáltatásokat lehet rájuk építeni (például keresések során mindkét irányból rá lehet mutatni a másik műre).

ISTC-kód 16 hexadecimális számból áll, amelyeket négy blokkra tagolva kell kitölteni. Az első blokkban kell megadni a regisztrációs ügynökség kódját, a másodikban a regisztráció évét, a harmadikban a mű egyedi azonosítóját, a negyedikben pedig egy ellenőrző számot kell elhelyezni, amely a ISO 7064 szabványban definiált MOD 16-3 algoritmus szerint végez

vizsgálatot. A harmadik blokkban levő műazonosító az adott éven belül és adott ügynökséghez kötődően garantálja az egyediséget. E két adat adja az azonosító kontextusát.

ISTC 0A9-2002-12B4A105-6

2.8.1.4 DOI



A DOI-t (Digital Object Identifier = digitális objektumazonosító) az (elsősorban online megjelenésű) elektronikus kiadványok azonosítására hozták létre; segítségével az adott tartalom elektronikus úton terjeszthető. Hatékony eszköz a (bármilyen formában megjelenő) digitális tartalom hatékony kezeléséhez, összeköti a vásárlót a tartalomszolgáltatóval, megkönnyíti az elektronikus kereskedelmet és lehetővé teszi a szerzői jogok automatikus kezelését.

Mivel a DOI rendszer tagszervezetei nem földrajzi egységeket, hanem bizonyos dokumentumtípusokat és tevékenységeket fednek le, DOI igénylésekor célszerű a központhoz fordulni. A DOI kompatibilis az ISMN-nel: a szerkezetébe akár az ISMN-t, akár más, a kiadók által használt azonosítót be lehet építeni. A DOI perzisztens azonosító.

<https://doi.org/10.1016/j.jtbi.2013.12.011>

2.8.1.5 NBN

Az URN (Universal Resource Name) egy olyan azonosító, amely az interneten található elektronikus dokumentumok hosszútávú azonosítására szolgál. Ennek egy altípusa az NBN (National Bibliography Number). Az URN azonosító típust azzal a céllal hozták létre, hogy kiküszöböljék az URL-ek használatából fakadó hátrányokat. Mivel az URL-ek gyakran változnak, a hivatkozások folyamatos karbantartása nagy erőfeszítést igényelne. A problémát még tovább növeli, hogy a változást közölnünk kell mindazokkal, akik az általunk közzétett forrásra hivatkoznak. Az URN azonosítók (természetesen az URL-el együtt) ezekre a nehézségekre próbálnak megoldást nyújtani.

urn:nbn:de:bvb:19-146642

2.8.2 Zeneszámok

A zeneszámok azonosítására több szabvány is létezik. Ennek az az oka, hogy a zeneszámokat másként kezelik attól függően, hogy magát a művet vagy annak valamelyik felvételét akarják azonosítani, illetve más rendszert használnak a zeneszámok auditív vagy vizuális megjelenési formáinak, a hanganyagoknak és a kottáknak az azonosítására.

2.8.2.1 ISMN

Az ISMN (International Standard Music Number = Nemzetközi Szabványos Kottaazonosító) numerikus kód, mely bármilyen formában megjelenő kották azonosítására szolgál. Alkalmazását az ISO 10957 nemzetközi szabvány írja elő.

Miután az ISBN-ről bebizonyosodott, hogy széles körben alkalmazható, hasznos eszköz a könyvkereskedelemben, a zeneműkiadók is hasonló szabványos azonosító bevezetését javasolták a kották számára. A tárgyalásokat lassította az afölötti vita, hogy vajon elegendő-e egy egyszerű azonosító szám, vagy inkább bibliográfiai kódra van szükség, amely jelzi például a partitúra és a szólamok közötti kapcsolatot. Amikor a Zenei Könyvtárosok, Archívumok és Dokumentációs Központok Nemzetközi Szövetsége (International Association of Music Librarians, Archives and Documentation Centres = IAML) a Nemzetközi ISBN Ügynökséghez fordult egy tíz számjegyből álló, bibliográfiai kódok nélküli azonosító bevezetésének javaslatával, a tervezet megjelent az ISBN Review-ban, és lelkes fogadtatásra talált. Az ISO TC 46-os munkacsoportja befogadta a tervezetet, és komoly viták után, egy Ottawában rendezett találkozón az európai és észak-amerikai kiadók és szakértők megegyezésre jutottak. A fő érv a bibliográfiai kódra vonatkozó eredeti elképzelés elvetése mellett az ISBN-nel kapcsolatos tapasztalat volt, és az a lehetőség, hogy a 10 jegyű számot az ISBN mintájára be lehet építeni a 13 jegyű nemzetközi vonalkódrendszerbe.

Az ISMN szabványt rekordidő alatt dolgozták ki, és 1993 végén meg is jelent. Az ISO-szabványt a Magyar Szabványügyi Testület magyar szabványként 1997 júliusában adta közre. Az ISMN alkalmazása teljes körű ésszerűsítési lehetőségeket kínál a kottakiadás és-kereskedelem, valamint a zenei könyvtárak számára.

Az ISMN eredetileg 10 jegyű volt (l. a 10. fejezetet), 2008-ban viszont 13 jegyűvé alakult. Az eredetileg az ISBN-nek fenntartott 979 előtaggal kezdődik, melyet a 0 karakter követ; ez utóbbi került a 10 jegyű ISMN M betűje helyére. Ellentétben az ISBN-nel, az ISMN nem tartalmaz nyilvántartásicsoport-azonosítót (elterjedtebb nevén országazonosítót), a zene nemzetközi jellegéből adódóan.

Amikor egy 10 jegyű ISMN-t 13 jegyűvé alakítunk, az ellenőrző szám változatlan marad. (Ezzel szemben, ha egy 10 jegyű ISBN-t alakítunk át 13 jegyűvé, újra kell generálni az ellenőrző számot.)

ISMN 979-0-2600-0043-8



az ISMN vonalkódja

Az ISMN kottakiadványokat azonosít, amelyek kereskedelmi forgalomba kerülnek, bérebe vehetők vagy ingyenesen kerülnek terjesztésre, vagy csak szerzői jogi célokra hozzáférhetők. Az ISMN nem alkalmazható hang és videofelvételekre (kivéve az alábbiakban jelzett ritka eseteket), mivel ezeket más számok vagy kódok azonosítják. A hangfelvételeket a hangfelvételek nemzetközi szabványos kódjával, az ISRC-vel (International Standard Recording Code), a videofelvételeket az ISAN-nal (International Standard Audiovisual Number) kell ellátni.

Egy kiadvány minden önállóan hozzáférhető alkotórészét saját ISMN-nel kell ellátni. ISMN-nel azonosítandó kiadványok lehetnek:

- Partitúrák
- Zseb- vagy kispártitúrák (tanulmánypartitúrák)
- Vokális partitúrák
- Szólamgarnitúrák
- Önállóan kapható (hozzáférhető) egyéni szólamok
- Könnyűzenei fóliánsok
- Antológiák
- Más hordozók (kísérő anyagok), amelyek a zenei kiadvány szerves részét képezik (pl. olyan hangfelvétel, amely egy zenemű egy szólama)
- Nyomatott zeneművel megjelentetett dalszövegek vagy versek (amennyiben önállóan kaphatók/hozzáférhető)
- Nyomatott zeneművel megjelentetett magyarázó szövegek (amennyiben önállóan kaphatók/hozzáférhető)
- Daloskönyvek (nem kötelezően!)
- Zenei kiadványok mikroformában
- Zenei kiadványok Braille-írással
- Elektronikus formában terjesztett kották

A következő kiadványok nem kaphatnak ISMN számot:

- Zenei tárgyú könyvek, kivétel, ha kottázott zenei műveket tartalmaz
- Önálló hang- vagy videofelvételek (beleértve a számítógépes formában terjesztett felvételeket)
- Periodikumok és sorozatok egésze, ellentétben a sorozatok egyes köteteivel

Az ISMN felépítése

A zeneművek nemzetközi azonosító száma (ISMN) a 979-0 előtagból és az azt követő kilenc számjegyből áll. Nyomatásban vagy írásban az azonosítót az ISMN betűkód előzi meg. (Olyan

országokban, ahol nem latin betűket használnak, a helyi írásrendszerben alkalmazott rövidítés is szerepelhet az ISMN mellett.)

Az ISMN négy részből áll, ezek közül kettő változó hosszúságú. Nyomtatásban vagy írásban az ISMN egyes részeit kötőjel vagy szóköz választja el egymástól.

ISBN 979-0-2600-0043-8

A négy rész értelmezése a következő.

Előtag	A 979-0 előtag különbözteti meg az ISMN-t az ISBN-től (a könyvek nemzetközi azonosítójától). A GS1 nemzetközi vonalkódrendszerében a 979 általánosságban az ISBN-t jelöli; ennek a számkapacitásnak az egytizedét, a 979-0-val kezdődő tartományt tartották fenn az ISMN számára.
Kiadói azonosító	Ez az elem az adott zenei kiadvány kiadóját jelöli, 3-7 karakterből állhat. A nagyszámú kiadványt megjelentető kiadók rövid kiadói azonosítókat kapnak; a kevés zeneművet megjelentetőök hosszabbakat.
Kiadványazonosító	A harmadik elem a zenemű egy kiadását vagy annak önállóan terjeszthető, hozzáférhető vagy beszerezhető alkotórészét azonosítja; pl. a teljes partitúrát, a zseb- vagy kispartitúrát, a fúvós szólamok csoportját, az oboa szólamot stb. A kiadványazonosítót rendszerint a kiadó állapítja meg egy meghatározott mű számára a rendelkezésére álló számtartományból (lásd még az 5. Nem közreműködő kiadók szakaszt). A kiadványazonosító hossza az előtte álló kiadói azonosító hosszától függ: ez a két elem együtt nyolc számjegyet tesz ki. Az ISMN-t már a kiadvány tervezése és előkészítése során ki lehet utalni, így belső azonosítóként is szolgálhat.
Ellenőrző számjegy	Az ellenőrző számjegy 10-es modulussal számítható ki az 1-es és a 3-as váltakozó súlyszámokkal balról jobbra haladva. Ez azt jelenti, hogy az ISMN első 12 karakterének mindegyikét – azaz, magának az ellenőrző számjegynek a kivételével minden számjegyét – meg kell szorozni váltakozva 1-gyel és 3-mal balról jobbra haladva, és az így kapott szorzatok és az ellenőrző számjegy összegének maradék nélkül oszthatónak kell lennie 10-zel.

2.8.2.2 ISRC

Az ISMN nem vonatkozik a hang- és videofelvételekre (kivéve az olyan eseteket, amikor egy kottakiadvány részeként jelennek meg). Az International Standard Recording Code, ISRC (nemzetközi szabványos felvétel kód) azonosítja egy mű egy adott felvételét (de gyakran nem a

fizikai egységet), tekintet nélkül arra, hogy milyen környezetben vagy hordozón jelenik meg. Az ISRC rendszer irányítását az IFPI: International Federation of Phonogram and Videogram Producers (Hangfelvétel és Videofelvétel Gyártók Nemzetközi Szövetsége) végzi. Az ISRC rendszernek nincs magyarországi tagszervezete. A CISAC magyarországi tagszervezete az Artisjus.

Az ISRC-kód négy részből áll:

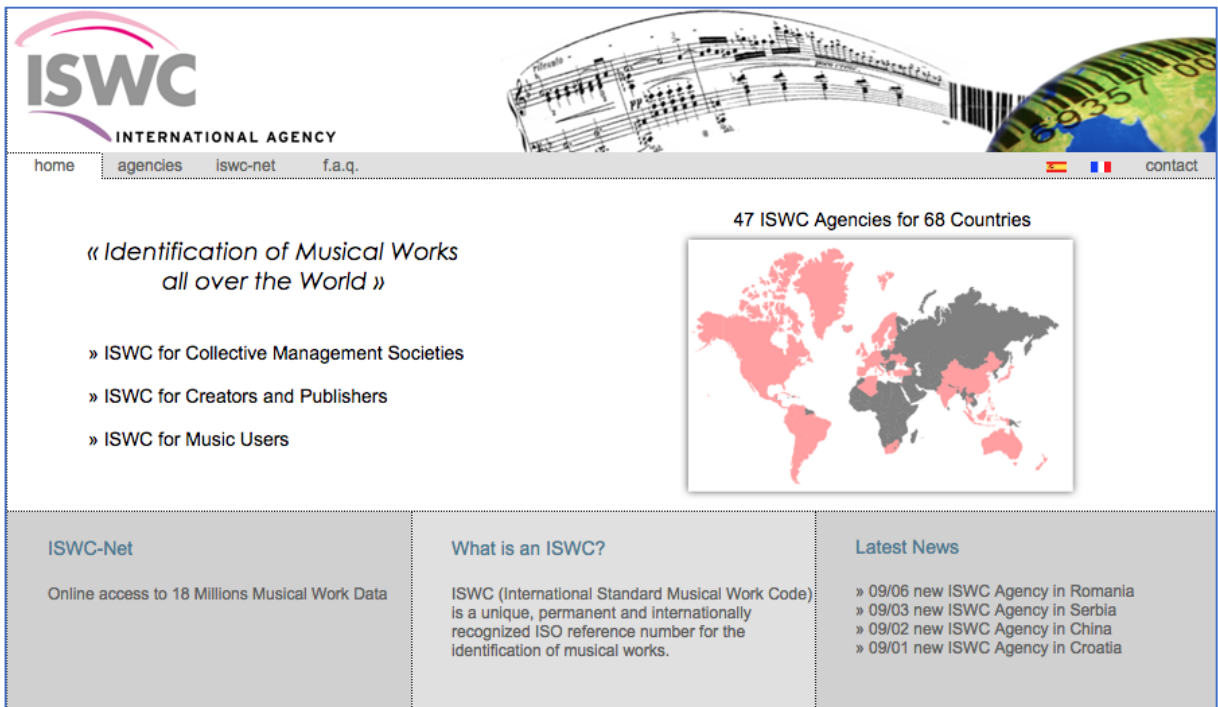
- az első blokk az ország kódja,
- a második az országon belüli regisztrációs szervezet kódja,
- a harmadik a kibocsátás évét jelöli,
- a negyedik a regisztrátor által kiadott kód az adott zeneszámmra.

US-Z03-99-32476

2.8.2.3 ISWC

A könyvekhez hasonlóan a zeneművek is rendelkeznek nemzetközi azonosító számmal, ez az ISWC (International Standard Musical Work Code – nemzetközi szabványos zeneműkód). Az ISWC-kód magát a zeneművet azonosítja, függetlenül annak megjelenési formájától. Alapvető fontosságú a szerzői jogok kezelésében. Minden ISWC egy "T" betűvel kezdődik, ezt követi egy 9 jegyű szám, majd egy ellenőrző számjegy. Az ISWC nemzetközi szinten azonosítja az adott művet. Minden zenemű, függetlenül attól, hogy megjelenik-e vagy sem, illetve hogy újonnan keletkezett vagy korábban hozták létre, kaphat ISWC-t. Ugyanez vonatkozik a zeneművek átdolgozásaira is.

A kód segítségével a külföldi jogvédő társaságok könnyedén beazonosíthatják a magyar műveket, az utánuk befizetett jogdíjakat pedig az Artisjuson keresztül elszámolhatják a magyar jogosultaknak. A zeneszámok Youtube-ra való feltöltésekor is érdemes megadni ezt a kódszámot a mű egyértelmű beazonosításához. Így az ebből a forrásból beérkező jogdíjak biztosan a jogosultakhoz fognak eljutni. Gyakran külföldi szerzői pályázatok, felhasználói szerződések kötelező eleme az ISWC kód.



ISWC
INTERNATIONAL AGENCY

home agencies iswc-net f.a.q. contact

« Identification of Musical Works
all over the World »

- » ISWC for Collective Management Societies
- » ISWC for Creators and Publishers
- » ISWC for Music Users

47 ISWC Agencies for 68 Countries

ISWC-Net
Online access to 18 Millions Musical Work Data

What is an ISWC?
ISWC (International Standard Musical Work Code) is a unique, permanent and internationally recognized ISO reference number for the identification of musical works.

Latest News

- » 09/06 new ISWC Agency in Romania
- » 09/03 new ISWC Agency in Serbia
- » 09/02 new ISWC Agency in China
- » 09/01 new ISWC Agency in Croatia

az ISWC nyitóoldala

Egy ISWC-kód felvételéhez minimum az alábbi információt meg kell adni.

- a zenemű címét,
- a zeneszerzők, a hangszerelők, alkotók nevét, szerepét a mű létrehozásában egy központi foglalkozási szereplista alapján), valamint az alkotók CAE/IPI-számát,
- a mű besorolási kódját (CIS-kód),
- azt a művet, amelyből a mű származik (ha átdolgozásról van szó).

Újra fontos itt jelezni, hogy az ISWC-kód a művet, nem a felvételt azonosítja. A felvételt az ISRC-számmal lehet azonosítani.

ISWC
INTERNATIONAL AGENCY

home agencies iswc-net f.a.q. contact

ISWC Regional Agencies

Agency	Countries	Contact
APRA	Australia, New Zealand, Fiji	showland[a]apra.com.au
CASH	People's Republic of China, Hong Kong, Taiwan	apple.tam[a]cash.org.hk
ECAD	Brasil	gloria_braga[a]ecad.org.br
Latinautor	Dominican Republic, El Salvador, Guatemala, Honduras, Nicaragua	mrevolorio[a]aei-guatemala.org
MIS@Asia	India, Indonesia, Philippines, Singapore, Thailand	raymondan[a]compass.org.sg
Nord-ic	Denmark, Estonia, Finland, Greenland, Iceland, Latvia, Lithuania, Norway, Sweden	johan.hammarstrom[a]stim.se
Song Code	Canada, United States	songcode[a]ascap.com

ISWC Agencies per Country

Country	Agencies
Algeria	ONDA
Argentina	SADAIC
Australia	APRA
Austria	AKM
Belgium	SABAM
Bolivia	SOBODAYCOM
Bosnia and Herzegovina	SNQ
Brasil	ECAD
Canada	Song Code
Chile	SCD
People's Republic of China	MCSC
Colombia	SAYCO
Costa Rica	ACAM
Croatia	HDS-ZAMP
Cuba	ACDAM
Czeck Republic	OSA
Denmark	Nord-ic
Dominican Republic	Latinautor
Ecuador	SAYCE
El Salvador	Latinautor
Estonia	Nord-ic
Fiji	APRA
Finland	Nord-ic
France	SACEM
Georgia	SAS
Germany	GEMA
Greece	AEPI
Greenland	Nord-ic
Guatemala	Latinautor
Honduras	Latinautor
Hong Kong	CASH

ISWC Local Agencies

Agency	Countries	Contact
AACIMH	Honduras	
ACAM	Costa Rica	e.serrano[a]acam.cr
ACDAM	Cuba	olival[a]acdam.cu
ACUM	Israel	odedf[a]acum.org.il
AEI	Guatemala	
AEPI	Greece	
AGADU	Uruguay	
AKM	Austria	waltraud.schmidt[a]aume.at
APA	Paraguay	apa[a]quanta.com.py
APDAYC	Peru	
ARTISJUS	Hungary	avarkonyi[a]artisjus.com
BUMA-STEMRA	Netherland	info[a]bumastemra.nl
CASH	Hong Kong	apple.tam[a]cash.org.hk
GEMA	Germany	jroehr[a]gema.de
HDS-ZAMP	Croatia	Eugen Sugovic
IMRO	Ireland	declan.rudden[a]imro.ie

az ISWC regionális és országos ügynökségei

Az ISWC-szám kiosztását Magyarországon az Artisjus végzi. A kód kiosztásának feltétele a mű bejelentése az Artisjushoz.

Művek adatainak exportálása

Exportálás excel formátumban Exportálás pdf formátumban

Keresés

Összes típusban 5 sor/flap

Művek

Műkód	Műcím	Műbejelentés dátum	Műfaj	Időtartam	ISWC
400007875	FASRYTALÉ	2005.09.15	0-401	3:19	T0070012345
400551314	BELIEVE	2011.11.23	0-401	3:46	T0070012346
4001980309	THE VOICE	2013.06.03	0-403	3:02	T0070012342
4002049827	ONLY TEARDROPS	2013.11.11	0-401	2:10	T0070012347
4002049987	MY NUMBER ONE	2013.11.11	0-401	3:57	T0070012341

zenei művek bejelentése az Artisjuszánál

A művek azonosítása mellett mód van a szerzők azonosítására is. Erre szolgál az IPI-kód (IPI – Interested Parties Information), amellyel a szerzőt, a jogtulajdonost lehet regisztrálni. A művek mellett a személyekre is van Artisjus-azonosító.

Szerzői u
Tel.: (1) 4
ügyfelszo

artisjus
Szerzői
Információs
Rendszer

Művek Forgalmak Értesítő

Artisjus azonosító: 1234567 IPI azonosító: 12345678

Személyi adatok Pénzügyi adatok

a zenei művel kapcsolatba hozható személy IPI-azonosítója

2.8.3 Filmek, audiovizuális tartalmak

2.8.3.1 ISAN

Az International Standard Audiovisual Number (ISAN), mint a nevéből is látszik, az audiovizuális műveket azonosítja. Audiovizuális műnek számít minden olyan alkotás, amely mozgóképekből tevődik össze, hangkíséréssel vagy anélkül. Az ISAN magát a művet azonosítja, nem a fizikai hordozón megjelenő kiadványt vagy az adást. Lehetővé teszi a szerzői jogok birtokosai (szerzők, előadók, gyártók) számára, hogy egyértelműen azonosítsák az audiovizuális művet.

Az ISAN Audiovizuális művek és verzióik azonosítására szolgáló, önkéntes számozási rendszer, amelyet az ISO (Nemzetközi Szabványügyi Szervezet) egyik csoportja (a Műszaki Bizottság 46, 9. albizottság) dolgozott ki. Az ISAN projektet az AGICOA (az Audiovizuális művek Nemzetközi Kollektív Igazgatásának Szövetsége), a CISAC (a Szerzők és Zeneszerzők Szövetségeinek Nemzetközi Szövetsége) és a FIAPF (a Filmgyártók Szövetségeinek Nemzetközi Szövetsége) irányította, mindhárman számos audiovizuális gyártással foglalkozó nemzetközi és nemzeti szervezetet képviselve.

A nemzetközi szabványt, amely az ISAN rendszer alapját képezi 2002-ben adták ki mint ISO 15706-1: 2002 Információ és Dokumentáció - Nemzetközi Szabvány Audiovizuális Szám (ISAN) - 1. rész: Az Audiovizuális mű azonosítója. 2007-ben a ISAN szabványt kiegészítették, hogy verzió azonosító támogatást nyújtson az ISO 15706-2: 2007 Információ és Dokumentáció - Nemzetközi Szabvány Audiovizuális Szám (ISAN) - 2. rész: Verzió Azonosító. Együtt, egy teljes Audiovizuális Tartalomazonosítási rendszert írnak le.

Az ISAN egy olyan önkéntes számozási rendszer, amely nem kapcsolódik semmilyen szerzői jogi nyilvántartáshoz. Az ISAN az ISO 15706 nemzetközi szabványnak felel meg.

Az ISAN változatlan marad a műmegosztás minden szakaszától függetlenül, beleértve minden új verziót.

Az ISAN olyan konkrét leíró információkat (metaadatokat) foglal magában, mint a cím, a rendező, a típus, a időtartam stb. Ez minden műfajra vonatkozik, beleértve a játékfilmeket, a sorozatokat, a dokumentumokat, az élő közvetítéseket, a szórakoztató programokat, a multimédiát stb., és azok sok változatát (szinkronizálás, hangalámondás, közvetítés, Blu-ray), beleértve a szorosan kapcsolódó tartalmakat (pl. előzetesek) vagy szorosan kapcsolódó elemeket (pl. filmzene, poszter).

Bár az ISAN gyökér szegmense ugyanaz marad az Audiovizuális mű minden változatánál, az ISAN verziószegmens lehetővé tesz vékonyabb szemcsés azonosítást verziószinten is, úgymint kiadásokat, nyelvi változatokat és médiafigyelést, mint például a hagyományos filmszalag, digitális mozi, a Blu-ray, DVD, mágnesszalagok, digitális fájlok, digitális streamek (adatfolyamok). Az ISAN kompatibilis a legfontosabb audiovizuális szabványokkal, például MPEG, Windows Media, DVB ATSC, MXF stb.

Az ISAN beépíthető fizikai adathordozókba, digitális streamekbe (adatfolyamokba), digitális fájlokba, valamint vonalkódokba, kiadványokba, csomagolásba, reklámba és audiovizuális megállapodásokba. Az Audiovizuális Tartalmakhoz (művekhez vagy verzióhoz) hozzárendelt ISAN állandó, független a tulajdonjogi változásoktól.

Az ISAN felhasználási módjait úgy alakították ki, hogy azok megfeleljenek az audiovizuális iparág különböző igényeinek. Erőssége ebből az „egyetemességből” fakad.

Az ISAN-szabvány az erős nemzetközi kulturális sokféleség miatt az ISAN Regisztrációs Ügynökségek hálózatán alapul, amelyek a nemzetközi ISAN Rendszer és a helyi ipar közötti kapcsolódási pontként működnek. A regionális ISAN Regisztrációs Ügynökségeket az ISAN-IA nevezi ki, miután támogatást kaptak a helyi ipartól. A regisztrációs ügynökségek alkalmazott ISAN Szolgáltatásokat és szaktudást nyújtanak a helyi nyelven és az ISAN közösségek felé jelentik a helyi elvárásokat.

Az ISAN csak az ISAN Rendszeren keresztül kinevezett Regisztrációs Ügynökségeken keresztül adható ki.

Az ISAN hálózat a Nemzetközi Ügynökség (ISAN-IA), amelyet az AGICOA, a CISAC és a FIAPF alapított, támogatásával üzemel.

Az ISAN-t már több mint 730 000 Audiovizuális Tartalom-hoz rendelték hozzá.

Az ISAN Regisztráció online történik, függetlenül attól, hogy egy vagy több műre vonatkozik. A tömeges ISAN-allokációk egyszerűen végrehajthatók XML-fájl alapú folyamatokon keresztül. Egy ötvennél több ISAN Webes Szolgáltatás csomagos készlet teszi lehetővé az automatikus regisztrációkat és a lekérdezéseket harmadik féltől származó alkalmazásokhoz.

Az ISAN az alkotásokat, azok teljes kereskedelmi ciklusán keresztül azonosítja, mindenféle változatban, különböző nyelveken, vágásokban, formátumokban, médiumokban; mindenkor, minden módon és a terjesztés minden szakaszában.

Az ISAN hasznára van a szerzőknek, a műsorszolgáltatóknak, a gyűjteménytársaságoknak, a vállalatoknak, a gyártóknak, a jogtulajdonosoknak és még sokan másoknak.

Az ISAN előnyei

Minden egyes ISAN célja, hogy egyedülálló nemzetközi azonosítóként szolgáljon minden Audiovizuális mű („AV/mű”), és annak bármely verziója számára. Az Audiovizuális művek és a hozzájuk kapcsolódó verziók ugyanazt a gyökérszegmenst tartalmazzák. Az ISAN leíró információkkal (metaadatokkal) van összekapcsolva, amelyek mindegyike a Központi Adattárban található.

Az ISAN kódok egy egyedi Központi Adatbázis-ból származnak. Minden AV/mű vagy verzió egyetlen egyedi ISAN-al rendelkezik. Minden ISAN csak egy AV/művet vagy egy verziót azonosít. Az ISAN egy permanens és állandó azonosító.

Az ISAN minden típusú Audiovizuális műre és azok mindenféle verziójára vonatkozik, úgymint nyelvi változatok, kiadások, kivonatok, klippek, média változatok, digitális kódolás, valamint az ezekhez szorosan kapcsolódó tartalmak vagy tételek, úgymint filmzenék, feliratozás, dobozott készletek stb.

Az ISAN úgy lett kialakítva a Nemzetközi Szabványügyi Szervezet (ISO) által, hogy megfeleljen a digitális forradalom, tartalomtulajdonosok, közvetítők, a médiavállalatok és a szabványügyi szervezetek által meghatározott követelményeinek.

Az ISAN és a hozzá kapcsolódó leíró információk mindenkor elérhetők online a Felhasználók (Regisztrálók és Olvasók) által. Az API-k és a tömeges folyamatok lehetővé teszik a meglévő folyamatok hatékony integrálását. Az ISAN segíti az ISAN regisztrálók hatékony készletgazdálkodását, különösen a digitális környezetben. Azáltal, hogy megköveteli, hogy az ISAN adásnaplóban szerepeltetését a használat nyomon követését könnyíti meg. Egy teljesen digitális környezetben az ISAN magába a műbe integrálódik.

Kollektív igazgatással történő adminisztráláskor az ISAN használatával könnyebbé válik a jogok nyilvántartása, ami jelentős idő- és költségmegtakarítást eredményez. Ez lehetővé teszi a gyűjteménytársaságok gyorsabb, megbízhatóbb és hatékonyabb jogelosztását.

Az ISAN megkönnyíti a különböző külső adatbázisok közötti elektronikus információcserét. Az ISAN elősegíti a mű használatának jobb nyomon követését a digitális világban, és ezáltal segít a kalózkodás ellen. Az ISAN kódok könnyen alkalmazhatóak, és felhasználhatóak az audiovizuális e-kereskedelmi környezetben.

Az ISAN az ISO tulajdona. Az ISO szerződéses kötelezettséget vállal az ISAN-IA és az ISAN Regisztrációs Ügynökségek ISAN Szabvány költség-visszatérítési alapon való működtetéséhez és az ISAN Szolgáltatásainak ésszerű díjért történő kiszámlázására.

Az ISAN Felhasználási Feltételek lehetővé teszik az ISAN kódok korlátozások nélküli felhasználását kereskedelmi és nem kereskedelmi célokra egyaránt.

Az ISAN azonosító általános szerkezete

Az ISAN-t úgy tervezték, hogy emberek olvassák és információs rendszerek dolgozzák fel, egy 24 bites hexadecimális számként (0-tól 9-ig és A-tól F-ig) vagy 96 bites bináris számként. Az ISAN

szerkezetét úgy alakították ki, hogy kielégítse az audiovizuális ellátási lánc különböző igényeit, és a számozási sémát úgy tervezték meg, hogy milliárd különálló cím azonosítását tegye lehetővé, és így az audiovizuális termelési igényeket legalább egy évszázadra kiszolgálja.



ISAN azonosító

Az ISAN három szegmensre oszlik:

- az első 12 számjegy jelöli a gyökér szegmenst,
- a következő 4 számjegy jelöli az epizódot/ egy sorozat részt,
- az utolsó 8 számjegy a verziószegmenst jelöli.

Az ISAN gyökér- és epizódsgemnsét jelölő első 16 számjegy a művek bármely verziójától függetlenül azonos marad - csak a legutolsó 8 számjegy változik, ahogy a verziók készülnek.

Az ISAN 17. és 26. számjegyei az ellenőrző karakterek, amelyek az ISAN helytelen átírásából eredő hibák ellen védenek. Az ellenőrző karaktereket egy számítógépes *algoritmus* rendeli hozzá automatikusan.

Amikor az ISAN-t olyan műhöz rendelik hozzá, amely nem Epizód, vagy Audiovizuális sorozatmű része, akkor az epizódsgemns 4 nullát (0000) fog tartalmazni.

Amikor az ISAN-t egy Epizód-hoz vagy egy Audiovizuális sorozatmű-höz rendelik hozzá, az epizódsgemns egy sorozathoz tartozó bejegyzett epizód azonosítására szolgál. Ebben az esetben egy epizódsgemns soha nem tartalmaz 4 nullát (0000).

0 értéket (azaz 0000-0000) akkor rendelnek az ISAN Verziósgemns-éhez, ha nincs verzió.

Az ISAN Verziósgemns, ahol az első 4 bit 1111 (azaz az F hexadecimális érték) privát használatra van fenntartva, és soha nem hivatkoznak rá az ISAN Központi Adatbázisában. Az ISAN Verziósgemns-ének ez a tartománya kizárólag belső alkalmazásokhoz van fenntartva (pl. a nyilvános szerkesztéshez készült verzió elkészítése során a belső szerkesztések nyomon követésére), ahogy azt egy ISAN "Verzió Regisztráló" meghatározza, aki az ISAN Verziósgemnsét önállóan hozzárendelheti nem-regisztrált alapon. Az ebből a belső tartományból származó ISAN nem terjeszthető külsőleg, mivel automatikusan más alkalmazások számára érvénytelennek minősülnek.

Nyomatott formátum

A nyomtatott formátumban megjelenített ISAN mindig tartalmazza a két ellenőrző karaktert. Az ISAN pontos átírásához a következő szabályok érvényesülnek:

- az „ISAN” szavat, amelyet szóköz követ, meg kell előznie minden számjegynek;
- az első 16 számjegyet 4 db 4 számjegy hosszúságú egységre kell felosztani, ahol minden egyes egységet szóköz vagy kötőjel választ el egymástól, és amelyeket az ellenőrző karakter követ;
- azután a 8 számjegyű változatot 2 db 4 számjegy hosszúságú egységre kell felosztani, ahol minden egyes egységet szóköz vagy kötőjel választ el egymástól, valamint az ellenőrző karakter követ;

Ezek a szabályok kizárólag arra szolgálnak, hogy vizuális segítséget nyújtsanak az ISAN pontos átírásához. Semmiféle jelentést, egységet vagy karaktert sem rendelnek hozzá semmilyen számjegyhez.. Példa:

ISAN AEF0-1000-6721-0022-X-0000-9034-1

Bináris formátum

Amikor bináris formában jelenítik meg, akkor az ISAN-t 96-bites bináris számként balról jobbra kódolják. A 96 bitet kódolják és előjelmentes egész számként továbbítják. Az ellenőrző számjegyeket figyelmen kívül hagyják.

MSB LSB

RRRRRRRRRRRREEEEEVVVVV VVV

Msb = legjelentősebb bit

Lsb = legkevésbé jelentős bit

XML formátum

Amikor XML dokumentumokban jelenik meg, akkor az ISAN-t a <http://www.isan.org/ISAN/isan.xsd> linken található séma szerint kódolják. Az “ISAN” betűket elhagyják, és a két ellenőrző karakter opcionális az ISAN XML-ábrázolásában. Az <http://www.isan.org/ISAN/isan.xsd> által meghatározott XML séma korlátozásain túl az XML kódolás még a következőképpen kerül korlátozásra:

- ha a check1 attribútum jelen van, akkor az episodeOrPart attribútumnak jelen kell lennie;
- ha a verzió attribútum jelen van, akkor az epizódOrPart attribútumnak jelen kell lennie;
- ha a check2 attribútum jelen van, akkor a check1 attribútumnak jelen kell lennie;
- hacsak az XML-kódolást nem szigorúan gépek között használják, minden attribútumra szükség van.

Nézzünk meg néhány példát!

```
<ISAN root="1881-66C7-3420" episodeOrPart="6541" version="9F3A-0245" />
```

```
<ISAN root="1881-66C7-3420" episodeOrPart="6541" Check1="X" version="9F3A-0245"  
check2="Y" />
```

```
<ISAN root="1881-66C7-3420" episodeOrPart="6541" />
```

```
<ISAN root="1881-66C7-3420" />
```

Megjegyzés: A 4. példa nem teljesen definiálja az ISAN-t, de speciális körülmények között, ahol csak az ISAN gyökérkomponense ismeretes, alkalmazható.

URN formátum

Amikor URN formában jelenik meg, a reprezentáció hasonló az emberi, olvasható formátumhoz, és az "URN: ISAN:" előre definiált, mint a következő példában látható.

```
URN: ISAN: 0000-0000-D07A-0090-Q-0000-0000-X
```

Az ISAN alkalmazása audiovizuális alkotásoknál és verzióknál

Audiovizuális művek és változataik

A rendszer számára az ISAN egy Audiovizuális művet, úgy határoz meg mint olyan "művet, amely kapcsolódó képek sorozatából áll, kísérő hanggal együtt vagy anélkül, amelyet mozgó képként jelenítenek meg eszközök használatával, a kezdeti vagy későbbi rögzítés módjától függetlenül ". Ez semmiképpen sem tekinthető azonban audiovizuális alkotások jogi meghatározásának.

A művek legtöbb típusa sorozatként vagy nem sorozatként regisztrálható. A művek legtöbb típusa az ISAN bevezetésének korai szakaszában használható. A művek legtöbbjének lehetnek verziói (kifejezések és megnyilvánulások), amelyek egy különálló ISAN Verzió-val azonosíthatók

Az AV/művek típusai, amelyekhez az ISAN hozzárendelhető, a következők például:

Kód	Leírás	Meghatározás	Lehet	Lehet
			sorozatszám?	összetett?
HIRDETÉS Reklám /		Művek, amelyek terméket, szolgáltatást, ötletet, személyt vagy céget, beleértve a fogyasztók számára		

	Reklámfilm	szóló promóciós anyagot, színházi szolgáltatásokat, infomercial-t digitális feliratozást reklámoznak. előzetesek (promóciós reklám egy film vagy televíziós produkció számára) NEM szerepelnek ebben a kategóriában, és azokat, mint az előzetes által azonosított mű egy verziójaként kell azonosítani.	Igen	Igen
FF	Játékfilm	Mű, amelyet eredetileg színpadra vagy közvetlenül videóra készült	Nem	Igen
TF	TV film vagy Tévjáték	Televíziós műsorszórásra szánt mű.	Igen*	Igen
DO	Dokumentumfilm	Egy mozgókép, amely tényalapú archívumot hoz létre vagy emberekről, állatokról, zöldegekről és növényekről, eseményekről vagy ipari folyamatokról tudósít.	Igen	Igen
ED	Nevelési	Az ilyen műveket elsősorban oktatási mintsem szórakoztatási célokra szánták..	Igen	Igen
SE	Dramai & sorozat	Komédia Egy fikciós sorozata gyakran ugyanazokkal a főszereplőkkel működik Szériák, vagy kapcsolódó témákról szóló nyílt végű televíziós dráma vagy "szappanoperák", vígjáték, ahol a történet epizódról epizódra folytatódik.	Igen	Igen
SH	Rövidfilm	Fikciós mű, amely rendszerint eredetileg színházi előadásra vagy közvetlenül videóra adtak ki amelynek eredeti hossza kisebb, mint a játékfilmé. Magába foglal rövid animált alkotásokat .	Igen*	Igen
MM	Multimédia	Audiovizuális mű interaktivitással.	Igen	Igen
NE	Hírek	Olyan audiovizuális produkció, amely közérdeklődésre számot tartó témákról számol be. Nem-fikciós programok,riportok és kommentár aktuális eseményekről, amelyeket rendszerint különféle	Igen	Igen

tudósítóktól származó riportokba szerkesztenek

LV	Élő esemény	Minden nem-fikciós mű, amely élő eseményeket rögzít, kivéve olyan sporthoz vagy előadásokhoz kapcsolódóakat, amelyeknek saját típusuk van. Például politikai események széles körű tudósítása, élő beszélgető műsorok, politikai viták, választások, úrbéli közvetítések házasság, temetés stb. élő rögzítése.	Igen	Igen
PF	Performansz	Művek amelyek előadásokat dokumentálnak, ideértve koncerteket, tánc, zene, opera, színpadi produkció, recitálás, bűvész-show, cirkuszi, stand-up komédia, börlszk, vagy más színpadi előadás.	Igen	Igen
SP	Sportesemény, Sporthoz kapcsolódó program	Nem-fikciós tudósítás sporteseményekről.	Igen	Igen
TE	Egyéb televízió szórakozás	Televíziós sugárzásra szánt programok, beleértve reality-show-kat, magazinokat, vetélkedőket, kivéve máshova kategorizáltakat.	Igen	Igen
VC	Zenei videoklip	Rövid audiovizuális nem élő produkció, amely egy adott hangfelvételhez kapcsolódó képeket mutat.	Igen	Igen
CO	Összeállítás	1 - Egyetlen kiemelt mű több kapcsolódó művel egy csomagban. (pl. a hagyományos DVD bónusz funkciókkal) - mindegyik mű a gyűjteményben saját felvétellel rendelkezik. 2 - Egy dobozos készlet, amely két vagy több kiemelt művet tartalmaz egy csomagban.	Nem	Igen (mindig összetett)

Egy ISAN Verzió-t bármely audiovizuális mű adott példányához (vagy verziójához) lehet alkalmazni, amely elemek bizonyos összesítéséből eredve egy ISAN-al már azonosított Audiovizuális mű tartalmát befolyásolja. Változások, mint például kiadások, a feliratok hozzáadása, a szinkronizálás, a hangsáv módosítása, a tartalom rögzítése egy adott médiumon (filmnyomatás, Blu-ray, digitális fájl vagy adatfolyam) a műszaki tulajdonságok (HD / SD, 2D / 3D, kódolási formátum) megváltoztatása mind új ISAN Verzió kiadásához vezethet.

Néhány példa művek módosítására, amelyek audiovizuális alkotások új verzióit eredményezve ISAN Verzió kiadásához vezethetnek:

- Egy adott nyelvi sáv megváltoztatása vagy nyelvi sávok elrendezése (nyelvi változat)
- A feliratozás változásai (feliratozott verziók)
- Változások a konkrét kép vagy hang formátumokban, mint például: széles képernyős vs. „pan & scan”, normál felbontású vs. nagy felbontású, 2D vs 3D, „kiváló minőségű” és „rossz minőségű” (pl D1 szalag átmásolása VHS kazettára, minőségromlással történő átkódolás)
- Átkódolás (pl. váltás AC3-ről MP4 hangkódolásra)
- Az Audiovizuális mű fizikai hordozójának megváltoztatása
- Egy Audiovizuális mű meghatározott célra történő szerkesztése, például: hosszúsági szerkesztések, a kereskedelmi forgalomba hozatalhoz szükséges módosítások, tartalszerkesztési célú változtatások.
- A technikai adatfolyamban bekövetkezett változás, amely befolyásolja a mű tartalmát (pl. egy másik szoftver, amely háttérképeket generál, vagy a függőleges üres intervallum a feliratozáshoz).
- Az Audiovizuális mű formátumának és/vagy időtartamának változása (pl. televíziós közvetítés, légitársaságok számára történő használat vagy hasonló változások egy adott felhasználási mód követelményei szerint)
- A mű digitalizált verziója eredetileg analóg formátumban
- A rendezői vágás;
- Jelentősen rövidített és újratervezett verzió;
- Egy eredetileg fekete-fehérben létrehozott Audiovizuális mű "színesített" verziója, vagy egy eredetileg színesben előállított Audiovizuális mű "fekete-fehér" verziója;
- Az Audiovizuális mű más jelentősen átdolgozott változatai.

Az ISO 15706-2 felhatalmazza továbbá az ISAN-Verzió alkalmazását olyan más tartalmakra is, amelyek az Audiovizuális művek verzióiból származnak, vagy ahhoz szorosan kapcsolódnak, ha ezt a tartalmat tömeges terjesztésre szánják, például műsorszórás céljából, és amikor szükséges a tartalom azonosítása egy létező ISAN vagy ISAN Verzió alkalmazásának szélesebb körű kontextusában. Ilyen kapcsolódó tartalom lehet:

- a sugárzott Audiovizuális mű-ből nyert hangsáv
- leíró hangsáv
- feliratozó sáv,
- több rendszeres hangsáv
- egy Audiovizuális mű hangsávjából kivont csengőhang
- egy kivonat ujjlenyomata vagy a teljes Audiovizuális mű
- egy kivonat audió ujjlenyomata vagy teljes hangsávja.
- az Audiovizuális mű plakátja
- az Audiovizuális művek dobozos készlete

- stb ...

A következő módosítások nem minősülnek különálló Audiovizuális mű-nek vagy ISAN Verzió-nak:

- változás az Audiovizuális mű jogaiban vagy tulajdonosában
- másolás azonos formátumú adathordozók között (pl. szalag-szalag másolat)
- változás a művet leíró metaadatokban
- az Audiovizuális mű használatával kapcsolatos árak, vagy díjak módosítása

Az ISAN Művek és a Verziók közötti kapcsolat

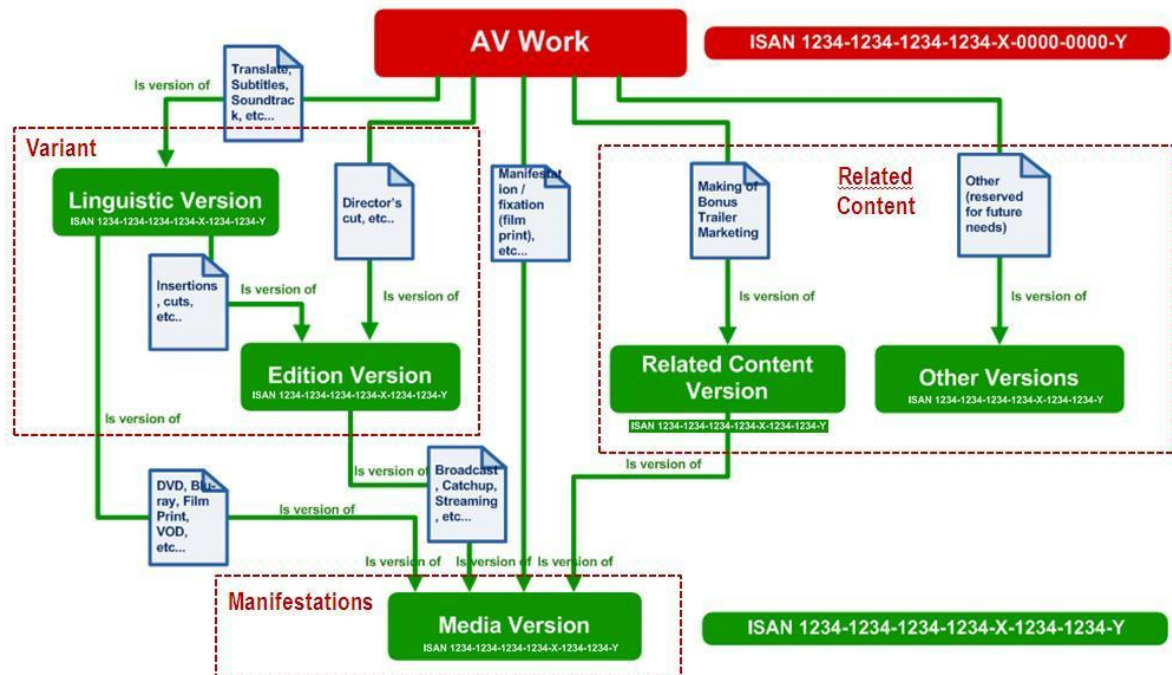
Az ISAN azonosító egy 24 számjegyű szám, amely két fő részből áll:

- az első 16 számjegyet az ISO 15706-1 szabvány határozza meg Audiovizuális mű azonosítójaként;
- a következő (és az utolsó) 8 számjegyet az ISO 15706-2 mint Verzió Azonosító kiterjesztést határozza meg.

A mű és a verzió közötti kapcsolat natív módon szerepel az ISAN Azonosítóban; tulajdonképpen minden változat ugyanazon az első 16 számjegyen osztozik csak az ISAN utolsó 8 számjegye különbözik egyik verzióról a másikkra.

Az Audiovizuális művek verziói közötti különbségeket, valamint az azonos művek verziói közötti kapcsolatot (ha vannak) az egyes verziók metaadataiban definiálják.

Ugyanazon Audiovizuális művek különböző verziói összekapcsolhatók az FRBR alapú modellekkel, ill. más hierarchiák engedélyezéséhez, ahogy az például az alábbi sémában látható.



audiovizuális művek verzióinak összekapcsolása FRBR-alapon

Audiovizuális sorozatművek és epizódok

Az ISAN-rendszer céljaira az Audiovizuális sorozatművek és az epizódok tág értelemben kerülnek meghatározásra. Következésképpen az epizódsgémsben semmilyen nyilvánvaló szekvenciális allokációk sem jelzik azt a sorrendet, amelyben az epizódokat előállították vagy továbbították. Egy Audiovizuális sorozatművet különálló epizódokban vagy olyan részekben állítanak elő amelyek kapcsolatban állnak egymással. Általában egy közös cím van az egész sorozat számára. Példák Audiovizuális sorozatművekre:

- minden olyan televíziós sorozat, műfaji megkötés nélkül, amelyet rendszeres időközönként sugárzott epizódokban állítanak elő;
- napi / heti talk show, hírközlés, sport esemény, stb ...;
- napi vagy rendszeres hírműsorok;
- Ezek NEM Audiovizuális sorozatművek:
- egy film folytatása (pl „Rocky IV”);
- egy Audiovizuális mű részeinek különálló sugárzása;
- évenkénti esemény éves tv-közvetítése (például egy szilveszteri műsor, vagy az Oscar Díj közvetítése).

Összetett művek és összetevőik

Egy összetett mű, olyan audiovizuális alkotás, amely teljesen más Audiovizuális műveket vagy Audiovizuális művekből származó kivonatot tartalmaz, és amelyek mindegyike lényegtelen az egész összetett alkotáshoz képest.

Egy összetett Audiovizuális mű saját ISAN-al rendelkezik, függetlenül bármely olyan ISAN-tól, amely az összetevői számára lett kibocsátva.

Példák az összetett művekre:

- Egy olyan film, amely animációs rajzfilmeket vagy más filmklippeket tartalmaz;
- Egy olyan televíziós műsor, amely korábban előállított filmekből vett részleteket tartalmaz;
- Egy televíziós epizód, amely más tévésorozat epizódjait tartalmazza.

A gyűjtemények mindig összetett művek.

Példák NEM összetett művekre:

- Egy bizonyos jelenet egy játékfilmen belül, kivéve, ha az adott jelenetet önálló terjesztésre szánják;
- Egy képmontázs.

Összetett művek esetén a regisztrálónak meg kell adnia, az összetett művet alkotó minden egyes Audiovizuális mű rendelkezésre álló ISAN számát.

Az ISAN AV művekhez történő hozzárendelése

Az ISAN-t permanensen kell hozzárendelni AV/mű mesterpéldányához annak érdekében, hogy az minden ilyen AV /mű egész életciklusa alatt fennmaradjon.

Digitális formátumú audiovizuális tartalmak esetén, az ISAN magába a műbe kódolható. Sok digitális formátuban alkalmazott audiovizuális szabvány ISAN kompatibilis, úgymint MPEG, Windows Media Format, DVB, ATSC, Blu-ray (AAC3) UltraViolet, Digital Cinema (DCI), MXF, BXF, Electronic Program útmutatók (Open EPG, TV Anytime), stb

Analóg formátumú audiovizuális tartalmak esetében, ajánlott, hogy az ISAN:

- Legyen hozzárendelve a mesterpéldányhoz vagy permanensen hozzákapcsolva a mesterpéldány tárolásával megbízott intézmény nyilvántartáshoz.
- Legyen láthatóvá téve az AV/mű fő- és /vagy végfelsorolásában.
- Javasolt továbbá, hogy minden egyes AV/művel együtt forgalmazott valamennyi kereskedelmi és műszaki dokumentáció tartalmazza az ISAN -t.

Amikor az ISAN ki van nyomtatva vagy másképpen ábrázolva, akár egy Audiovizuális Tartalomon, címkéken, műszaki és kereskedelmi dokumentáción, és a csomagoláson vagy máshol, akkor azt az „ISAN” betűknek, kell megelőzniük, hogy világosan megkülönböztethető legyen minden más termék számtól. Még ha nem is római számokat használnak, az „ISAN” szónak latin betűkkel kell szerepelniük.

Az ISAN pontos átírását elősegítendő, a következő szabályokat kell alkalmazni:

- az „ISAN” szónak, plusz egy üres helynek meg kell előznie minden számjegyet;
- a 16-jegyű ISAN-t fel kell osztani 4 egységnyi 4 számjegyre ahol minden egyes egység szóközzel vagy kötőjellel van elválasztva, amit az ellenőrző karakter követ;
- majd a 8 számjegyű verzió-hosszabbítást kell felosztani 2 egységnyi 4 számjegyre ahol minden egység szóközzel vagy kötőjellel van elválasztva, amit az ellenőrző karakter követ;

Ezek a szabályok kizárólag vizuálisan támogatják az ISAN pontos átírását; semmilyen jelentés sincs bármelyik számjegyre, egységre, vagy a karakterekhez rendelve. Az ISAN-al kompatibilis jelentős szabványokra és alkalmazásokra vonatkozó utalások megtalálhatók itt: www.isan.org

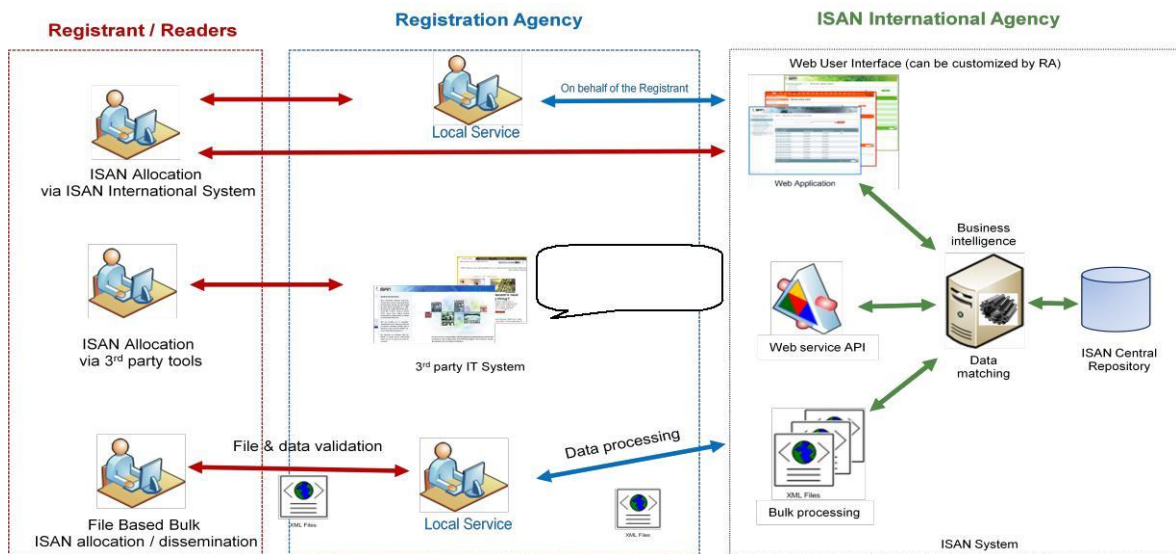
Az ISAN rendszerhez való hozzáférés

Az ISAN-t és a kapcsolódó metaadatokat a ISAN Központi Adatbázis tárolja, amely éjjel-nappal online elérhető, regisztrációt, konzultációt és kereséseket tesz lehetővé. Az ISAN rendszer, az infrastruktúra és az alkalmazások a svájci Genf egy ultramodern létesítményében üzemel, és nyújt biztonságos környezetben szolgáltatásokat, úgymint 24x7x365 ellenőrzést, teljesen redundáns biztonságos internet-kapcsolatot és teljes heti ill. napi back-up-ot. Az ISAN alkalmazás üzemeltetését és karbantartását az ISAN-IA mérnökei és külső szolgáltatók végzik.

Az ISAN adatbázis a regisztráláshoz és az interneten keresztüli lekérdezésekhez/keresésekhez a következőket kínálja.

- Webszolgáltatások, amely lehetővé teszik harmadik fél (mint például műsorszolgáltatók, stúdiók, digitális platformok, stb), informatikai rendszerének felhasználói felületének összekapcsolását a ISAN adatbázissal.
- Az XML fájlon keresztül történő tömeges feldolgozás nagy katalógusok számára is hatékony regisztrációt tesz lehetővé, drága műszaki beruházások, mint pl. az API alkalmazása, nélkül..

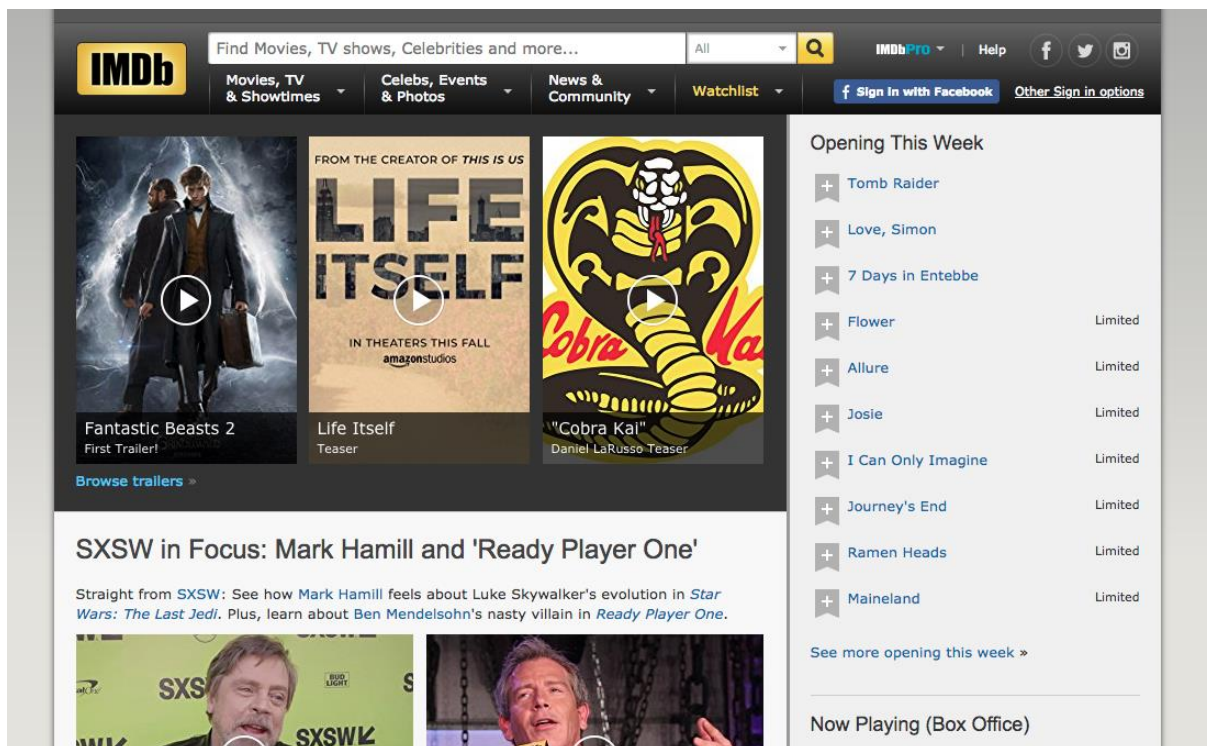
- AZ ISAN Adatbázis helyi másolatai (Oracle mentés) a Támogatói Olvasók rendelkezésre áll.



az ISAN-kódok regisztrációs folyamata

2.8.3.2 IMDb

Az Internet Movie Database (IMDb), melynek tulajdonosa az Amazon.com, egy online információs adatbázis filmszínészekről, filmekről, filmsorozatokról, egyéb televíziós műsorokról, reklámokról és videojátékokról. A regisztrált felhasználók értékelhetik a filmeket és szerkeszthetik a cikkeket. Az IMDb a legismertebb és legnagyobb méretű filmes internetes adatbázis.



az IMDb kezdőoldala ma

Az IMDb hatalmas mennyiségű információt tesz elérhetővé az alkotásokról, többek között alapvető adatokat, mint a színészek és rendezők neve, a stáb tagjai, a történet, kritikák, valamint kevésbé ismert adatokat, mint a háttérinformációk, baklövések, zenei kiadványok listája, képarány és alternatív verziók. A színészeknek, rendezőknek, íróknak és egyéb résztvevőknek is van saját adatbázis információjuk, felsorolva a filmeket és adásokat, melyekben részt vettek, továbbá a legtöbb esetben életrajz is olvasható róluk.

Az IMDb több, mint egy film és videojáték adatbázis, mivel napi film és TV híreket közöl, valamint kiemelt információkat tesz elérhetővé nagyobb filmes eseményekről, például az Oscar-díj átadásról. Az IMDb egy aktív közösséggel is rendelkezik. Minden filmhez tartozik egy szál a lap alján, továbbá általános vitalapok is vannak. Kibővítették a szolgáltatásokat a hivatásos szakemberek számára, így létrehozva az IMDbPro adatbázist, ahol a filmiparban szereplő személyek elérhetőségét, esemény naptárat és sok más szolgáltatás érhető el. Az IMDbPro nem a nyilvánosság számára készült, és így a tartalom sem ingyenes.

IMDbPro Search for a person, title or company **Search** Notifications | Help Deirdre Lovejoy

Home My Page My Tools Casting Notices People Titles Companies News

Deirdre Lovejoy

Actress

Overview Personal Details Media Clients

KNOWN FOR [The Wire](#) (2002–2006) | [The Talented Mr. Ripley](#) (1999) | [Step Up](#) (2006)

GUILDS AEA | SAG-AFTRA

WORK HISTORY Internet | Commercial | Theater | ... [More](#)

JOB TYPES Directing | Acting | Writing

VANITY URL <http://www.imdb.me/deirdrelovejoy>

STARmeter
5,451
9 this week

Contacts (3)
Show all

DIRECT CONTACT
Deirdre Lovejoy
www.deirdrelovejoy.com
ddlovejoysocial@gmail.com

TALENT AGENT THEATRICAL
Abrams Artists Agency
Justin Baxter

TALENT AGENT
Abrams Artists
Richard Fisher NY, Justin Baxter...

Filmography [Filter](#)

FILMS IN PRODUCTION (2 TITLES)	YEAR	BUDGET	STATUS
Billy Lynn's Long Halftime Walk – Denise Lynn	2016		Post-Prod
Thirsty – Doris Townsend	2015		Post-Prod

PAST FILM & VIDEO (14 TITLES)	YEAR	BUDGET	U.S. GROSS
Lionhead – Detective Lundgren	2013	\$2MM	
Killing Vivian (Short) – Toaster	2013		
Bad Teacher – Sasha's Mother	2011	\$20MM	\$100MM
The Starfather – Detective Teller	2010		\$50MM

People Also Viewed

- [Lance Reddick](#)
- [Dominic West](#)
- [Sonja Sohn](#)
- [John Doman](#) (1)

az IMDbPro felülete

Mindenki, aki egy e-mail címmel és egy cookie-kat támogató böngészővel rendelkezik, regisztrálhat az IMDb-re, majd feltölthet információkat, véleményezheti a filmeket, rangsorolhatja őket. A lekérdezések eredménye letölthető (tömörített) szöveges fájlként és a tartalmuk kinyerhető (jellemzően parancssoros eszközökkel). Az ellenőrzés hiányosságai miatt az adatbázisba téves információk is bekerülhetnek (például amatőr, botcsinálta vagy akár teljesen önkényes címfordítások), így a tartalom mindig fenntartással kezelendő.

1989-ben Col Needham és mások jelen voltak a rec.arts.movies Usenet hírcsoportban, ahol filmekről beszélgettek és információkat osztottak meg egymással. GYIK (Gyakran ismételt kérdések) listákat hoztak létre színészekről, színésznőkről és rendezőkről, adatokat és bibliográfiát közöltek a filmkészítőkről, akik elhunytak. 1990 végén közel 10 000 filmről és televíziós sorozatról volt már GYIK. 1990. október 17-én Needham Unix héjprogramokat küldött be, amik képesek voltak keresni a négy GYIK listában, és így megszületett az IMDb. Abban az időben azonban rec.arts.movies filmadatbázisként volt ismert.

az IMDb arculata 25 éve

1993-ra az adatbázist kiegészítették háttérinformációkkal, bibliográfiákkal és tartalmi összefoglalókkal, és egy központosított e-mail-felülettel a lekérdezések egyszerűsítésére. Még ebben az évben áttették World Wide Webre. Az adatbázis a walesi Cardiff Universityn található szervereken fut. Rob Hartill volt az eredeti webes felület alkotója. 1994-ben az e-mail-felületet továbbfejlesztették, hogy képes legyen fogadni információkat, így már nem kellett a lista karbantartónak elküldeni a változásokat. Az évek során tartalom tükrök hálózata alakult ki a világ minden táján – adományozott sáv szélességgel.

Minden közreműködő, aki tartalmat hoz létre az adatbázisban, szerzői jogokkal rendelkezik az általa létrehozott tartalomra, de teljes jogot ad másolásra, módosításra, valamint hozzáférést biztosít magának az IMDb-nek is. Az IMDb nem teszi lehetővé mások számára írásos engedély nélkül, hogy felhasználják a film összefoglalókat vagy a színészek életrajzait. A reklámszűrő szoftverek használata az oldalon kifejezetten tilos. A filmográfiák csak egy kis részét lehet idézni, és azokat is csak nem kereskedelmi weboldalakon. Az utóbbi adatfelhasználási megkötés nem érvényesíthető az Amerikai Legfelsőbb Bíróság ítélete értelmében (Feist v. Rural), mivel a tények gyűjtése nem tartozik a szerzői jog hatálya alá.

Az IMDb tipikus crowdsourcing jelenség, így magán viseli azok szabályozatlanságát, kontrollálhatatlanságát. Ennek ellenére ma már megkerülhetetlen tényezőnek számít a filmek, mozgóképek világában, és egy audiovizuális nemzeti névtér számára természetesen adódó nemzetközi kapcsolódási pontot jelent.

2.9 Személyek azonosító rendszerei

A személyek azonosítása minden területen lehet valamiért fontos, de a kreatív szakmákban sokszor szükség lehet arra, hogy megtaláljunk kutatókat, tudósokat, művészeket, írókat stb., akik valamilyen tevékenységük révén bekerülnek a nyilvánosság valamelyik szegmensébe. Mivel régóta tudjuk, hogy a személynevek önmagukban nem alkalmasak arra, hogy rajtuk keresztül megtaláljuk a minket érdeklő embereket, ezért régtől fogva komoly igény mutatkozott a személyek azonosítására képes rendszer bevezetése iránt.

2.9.1 Tudományos kutatók és cikkek azonosító rendszerei

A tudomány nemzetközi jellege miatt nem meglepő, hogy az elmúlt évtizedekben a tudományos világban több kezdeményezés is elindult a kutatók egyértelmű azonosítása érdekében. Ezt a célt nyilván csak egyetlen nemzetközi rendszerrel lehet csak elérni, de azt ma még nem lehet látni, hogy a jelenleg versengő rendszerek közül melyik lesz majd a „befutó”.

2.9.1.1 MTMT

A Magyar Tudományos Művek Tára (MTMT) a magyarországi szerzők hiteles, tudományos publikációs adattára. A publikációk és hivatkozások adatait ide fel kell tölteni, melyek hitelességét könyvtárosok ellenőrzik. Ezek alapján történik a publikációs teljesítések igazolása a fokozatszerzések, habilitáció, pályázatok esetén. Az MTMT létrehozásának célja az volt, hogy legyen egy – többféle célra is hasznosítható – nemzeti tudományos bibliográfiai adatbázis Magyarországon. A Magyar Tudományos Akadémia, a Magyar Akkreditációs Bizottság, a Magyar Rektori Konferencia, az OTKA és az Országos Doktori Tanács 2008 második felében kezdte meg az egyeztető tárgyalásokat. Az öt alapító intézmény, továbbá a koncepciót támogató OKM és NKTH egyetértett abban, hogy a hazai tudományos kutatás hiteles nyilvántartását és bemutatását kell megvalósítani az MTMT, illetve a rendszerhez csatlakozó szervezetek segítségével.

Az MTMT használatából származó előnyök a következők.

- Országosan egységes rendszerben vehet részt az intézmény;
- Fejlesztése és üzemeltetése lényegesen olcsóbb, mint a helyi fejlesztés és működtetés;
- Van kipróbált előzménye, a KPA, amit sok kutató ismer. Használták az ODT, az OTKA, az MTA igényeinek megfelelően;
- Az adatok sok célra felhasználhatók;
- Alkalmas belső igények kielégítésére (habilitáció, belső pályázatok, doktori képzés, statisztika, vezetői információs rendszer (VIR, AVIR) egyes adattípusaihoz);
- A külső minőségbiztosítás beépül a rendszerbe;
- Széles szakértői bázison működik, képzési lehetőséget biztosít;
- Nagyobb szervezetek adatigényeivel kompatibilis;
- Non-profit szervezet működteti.

Az MTMT használata előnyös a kutatók számára is. Miért jó a kutatóknak?

- Személyes tudományos bibliográfia összeállítása egyszerű;
- Kis munkaidővel karbantartható;
- Segíti pályázatok és jelentések készítését;
- Széles körben ismert formátum;
- Személyes honlapra beilleszthető;
- MTA, ODT, MAB, OTKA elfogadja, preferálja.

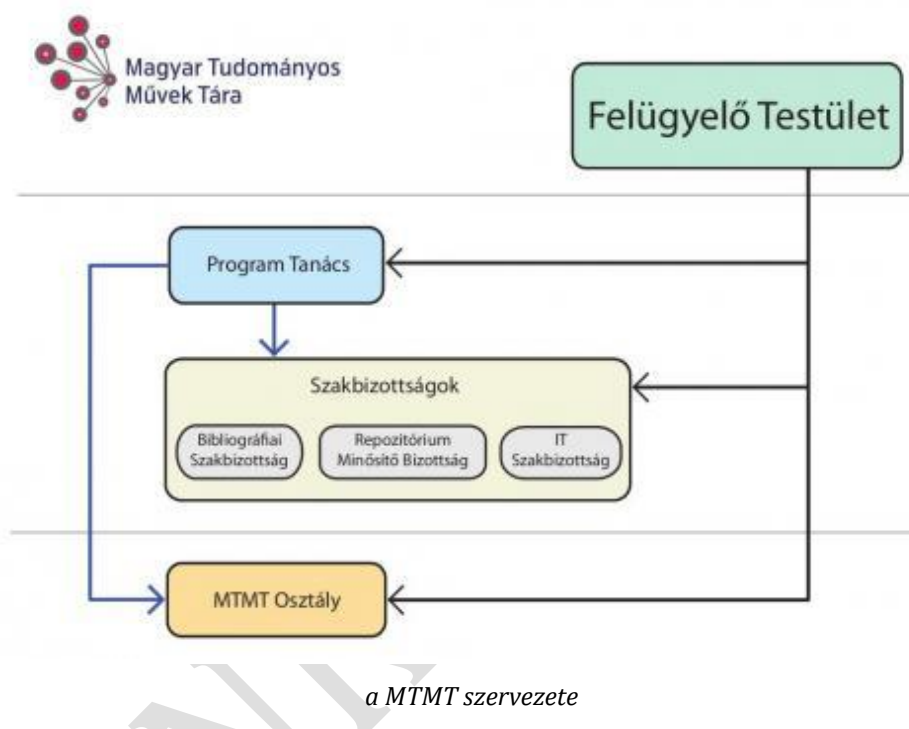
Miért jó a könyvtáraknak, könyvtárosoknak?

- Kiegészítő könyvtári funkciókat nyújt (folyóiratcikk nyilvántartás, idézetek nyilvántartása);
- Egységes folyóirat publikációs és idézési nyilvántartás;

- Ügyfélszolgálat (probléma megoldás, szoftverhasználati segítség).

Az együttműködés jogi alapja és szervezeti kerete

Az együttműködés a rendszerhez önkéntesen csatlakozó intézmények és az MTMT-t működtető MTA Könyvtára (a továbbiakban: MTAK) közötti szerződés alapján valósul meg. A felhasználók összességét az MTMT Program Tanács (PT) képviseli. A Felügyelő Testület (FT) ellenőrzi a tevékenységet, és dönt a hatáskörébe tartozó kérdésben.



Az MTMT operatív működését az MTA Kutatásszervezési Intézete (KSZI) szervezeti egységként létrehozott, majd annak jogutód nélküli megszűnése után az MTAK szervezeti keretein belül működő MTMT Osztály (korábban Koordinációs Iroda) biztosítja. Az Osztály folyamatos kapcsolatot tart azokkal a könyvtárosokkal és/vagy más munkatársakkal, akik az egyes intézmények adatbázisát karbantartják.

Az MTMT szolgáltatásai.

- Modern informatikai eszközökkel hiteles tudományos eredmény-nyilvántartást biztosít, és bemutatja az eredményeket, magyar és részben angol nyelven.
- Nyilvános, hiteles, átvethető és hivatkozható publikációs listákat ad. Ezek felhasználhatók az intézményi akkreditációban, a doktori iskolákkal kapcsolatos követelmények teljesítése során, MTA, OTKA és NKTH pályázatoknál.
- Segíti az akadémikus tagajánlásokat, az "MTA doktora" pályázatok beadását és azok értékelését.
- Bibliográfiai adatok révén átmenetet biztosít a teljes szövegű közleménytárhoz (repozitórium).

- Az adatot szolgáltató intézmények tudományos teljesítményéről összesített képet ad. Ez felhasználható az adott intézmények önértékeléséhez és/vagy külső értékeléséhez.
- A rendszer automatikus adatfeltöltést és megfelelő előkészítéssel adatkonverziót is végez, meglévő adattárakból és könyvtári rendszerekből.

A szolgáltatások korszerűsítésén az MTMT Osztály a források figyelembe vételével folyamatosan dolgozik.

kutatói oldal az MTMT oldalon

Az MTMT működése

Az MTMT adatbázist a csatlakozott intézmények közösen használják, abban elkülöníthető módon kezelhetik saját intézményük adatait. Az együttműködő partnerek számára a szoftver ingyenes, és egyedileg nem módosítható. Az esetleges helyi igények kielégítésére kiegészítő funkciók fejleszthetők, de azok nem veszélyeztethetik az adatbázis biztonságos működtetését, és külön költségként merülnek fel. A szoftver karbantartását, frissítését az MTMT központi egysége végzi. Szakmai kérdésekben a résztvevők segítik egymást, együtt alkotják az MTMT szakmai bázisát. A központi egység végzi a koordinációt. Feladatkörébe tartozik a tanácsadás és a képzés is. Az együttműködés során felmerülő vitás kérdések megválaszolása a Felügyelő Testület kompetenciájába tartozik. Az együttműködő partnereket egyedileg érintő költségkalkulációt az MTMT Osztály készíti el, és terjeszti be a Program Tanácsnak. A csatlakozott intézmények a költségvetés ismeretében nyilatkoznak az elfogadásról, illetve esetleges szerződésbontási szándékukról. A jóváhagyott költségvetést a Kutatásszervezési Intézet elkülönítve kezeli, saját költségvetésén belül.

Az MTMT adatrendszerének működése során a kutatók mind saját – számszerű – azonosítót kapnak, amelyet a rendszer minden publikációjukhoz hozzákapcsol attól függetlenül, hogy milyen név alatt jelent meg a publikáció. Az MTMT-ben a személy tehát egyértelműen azonosítva van, és minden névváltozatot is számon tart a rendszer, de ezek a névváltozatok már nem rendelkeznek saját azonosítóval, csak a publikált dokumentum metaadatai közt szerepelnek.

Az MTMT-ben a publikációknak is egyedi azonosítója van, de az ebben rejlő potenciált a rendszer egyelőre nem használja ki.

2.9.1.2 ODT

Az Országos Doktori Tanács (ODT) a felsőoktatási intézmények doktori tanácsai elnökeiből álló testület. Az ODT számára jogszabályokban előírt feladatai közt szerepel az az elvárás, hogy „működteti a fokozatszerzési eljárások és értekezések nyilvántartását”. E cél megvalósítása érdekében az ODT létrehozott egy nyilvántartó rendszert, amelyen belül a tudományos fokozatszerzési folyamat összes fontos résztvevőjét egyértelműen azonosítókkal látták, látják el.

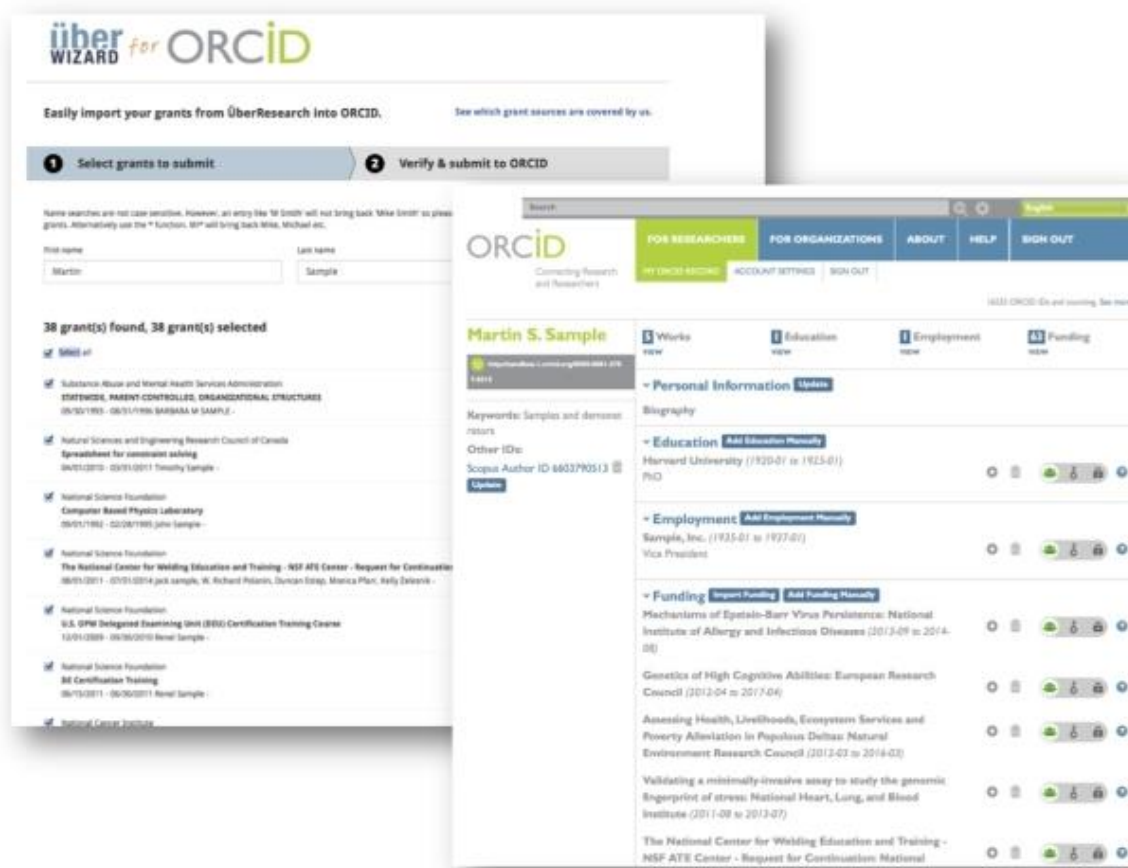
Személyes adatok	
	
név	Szakadát István
születési év	1958
intézmény neve	Eötvös Loránd Tudományegyetem Pécsi Tudományegyetem Budapesti Corvinus Egyetem
doktori iskola	ELTE Szociológia Doktori Iskola (oktató) PTE Nyelvtudományi Doktori Iskola (oktató) BCE Társadalmi Kommunikáció Doktori Iskola (témavezető)
doktori képzéssel kapcsolatos munkájának megoszlása	ELTE Szociológia Doktori Iskola 50% BCE Társadalmi Kommunikáció Doktori Iskola 50%
adott-e már oktatóként valamely doktori iskolát működtető intézménynek akkreditációs nyilatkozatot?	Budapesti Műszaki és Gazdaságtudományi Egyetem
Elérhetőségek	
drótpostacím	i@syi.hu
telefonszám	+36 1 458-0633
	saját honlap
	saját honlap (angol)
Fokozat, cím	
tudományos fokozat, cím	PhD
fokozat megszerzésének éve	1999
fokozat tudományága	szociológiai tudományok
fokozatot kiadó intézmény neve	MTA
Jelenlegi munkahelyek	
1997 -	Budapesti Műszaki és Gazdaságtudományi Egyetem egyetemi oktató

személyi adatlap az ODT oldalán

2.9.1.3 ORCID

A tudományos publikációk szerzőinek automatikus azonosítása hozták létre az egyre szélesebb körben elfogadott ORCID (open researcher and contributor ID) szerzői azonosítási rendszert. A rendszer terjedését nagyban segíti az a tény, hogy az ORCID-ban rögzített személyes adatokat egyre több folyóirat, kiadó automatikusan áttemeli saját rendszerébe (például egy kézirat beküldésekor). Az ORCID-ban a kutatók publikációk adatai is tárolhatók, és ezek szinkronizálhatók más adatbázisok (pl. Scopus, ResearcherID, Crossref) listáival.

Az ORCID a kutatók azonosítását egy 16 számjegyből álló egyedi azonosítóval oldja meg. A számok egy központi regiszterben tárolódnak, ahonnan sokféle irányban lehet elvinni az egyértelműen beazonosítható kutatók adatait.



keresési eredmények és szerzői profil az ORCID oldalon

Az ORCID összefésüli a ResearcherID, a Scopus Author ID-nk vagy a CrossRef által indexelt cikkeket adatait, s onnantól automatikusan lehet nyomon követni a kutató publikálási tevékenységének és cikkeinek hatását. Az ORCID-felhasználók száma 2013 óta folyamatosan nő, 2018 tavaszán az élő ORCID-azonosítóval rendelkező kutatók száma meghaladta a 4,5 milliót (<https://orcid.org/statistics>). Az ORCID 37 műtípust támogat, a folyóiratcikktől kezdve a táncelőadásig bezárólag, a tudományos és művészeti produktumok széles skáláját képes kezelni. Az ORCID kapcsolatot kínál más kutatói szolgáltatáshoz. Össze lehet kötni az ORCID profilt olyan weboldallal, mint a Web of Science (<http://wokinfo.com/researcherid/integration/>), Figshare (<https://figshare.com/>) vagy az Impactstory (<http://impactstory.org/>). Az összekapcsolások eredményeként könnyedén áramolhatnak az információk minden irányban, így nem kell ugyanazt az információt több helyre is feltölteni. Nap mint nap új szolgáltatások kapcsolódnak az ORCID-hez, így egyre növekvő számú platform, repozitórium, pályázati alap között áramlik az információ. Folyóiratok, alapítványok és intézetek mozdulnak el az ORCID irányába. Több mint ezer folyóirat – többek között a PLOS, Nature és Elsevier által kiadottak – tette lehetővé a szerzőik számára, hogy az ORCID használatával kezeljék az információkat a

kézirat feltöltő rendszerekben. Az ORCID képes arra, hogy a szerző megjelent publikációit összegyűjtse és a cikkekre vonatkozó metrikák aggregálásával szerzői szintű tudományometriai adatokat mutasson. Az ORCID szolgáltatása nyílt forráskódú, szabadon lehet az API-ját használni, lehet az adataiban bányászni.

2.9.1.4 ResearcherID

A ResearcherID a Thomson Reuters Web of Science adatbázisával összekapcsolt, szerzői azonosítót biztosító szolgáltatás. Létrehozható benne szerzői adatbázis, szinkronizálható az ORCID adatbázissal. A ResearcherID szerzői profil létrehozása és használata külön felületen, a <http://www.researcherid.com/> címen történik. Az oldalon a regisztrációt követően lehetőség van egyéni publikációs listák és hivatkozási riportok létrehozására, valamint a szerzőnek a Web of Science adatbázisban szereplő publikációinak magához rendelésére.

RESEARCHERID THOMSON REUTERS

Home Login Search Interactive Map EndNote Publons >

Szakadát, István Get A Badge ResearcherID Labs

ResearcherID: H-3517-2012
Other Names: Syl
URL: <http://www.researcherid.com/rid/H-3517-2012>
Subject: Mathematical Methods In Social Sciences; Sociology
ORCID: <http://orcid.org/0000-0002-1960-191X>
My URLs: <http://syl.hu>

My Institutions (more details)
Primary Institution: Budapest University of Technology and Economics, BME
Sub-org/Dept: Department of Sociology and Communication
Role: Researcher (Academic)

My Publications
My Publications (8)
View Publications >
Citation Metrics

My Publications: View
This list contains papers that I have authored.

8 publication(s) Page 1 of 1 Go Sort by: Publication Year Results per page: 10

- Title: Hidden Patterns of Reciprocity
Author(s): Syl,
Source: Journal of Theoretical Biology Volume: 345 Pages: 92-98 Published: 2014
added 19-Feb-14
- Title: MEO ontology infrastructure
Author(s): Szakadát, István; Szóts, Miklós; Gyepesi, György; et al.
Source: Advances in Information Systems Development, Vol 1 Pages: 369-377 Published: 2007
Times Cited: 0
DOI: 10.1007/978-0-387-70761-7_32
added 25-Sep-12
- Title: Dynamics of information access on the web
Author(s): Dezso, Z.; Almaas, E.; Lukacs, A.; et al.
added 25-Sep-12



kutatói adatlap a ResearcherID oldalon

A ResearcherID felületére belépve a My Researcher Profile menüpontban érhetőek el a szerző adatai és publikációs listája. Az oldal felső részében megtalálható a szerző ResearcherID azonosítója és a profil URL címe, amely a szerző adatlapjára vezető linkként más weboldalakon is feltüntethető. Az oldal ORCID mezője lehetőséget ad a szerző ResearcherID és ORCID szerzői profiljának összekapcsolására. A szerző a My Researcher Profile oldal My Publications nevű részében állíthatja össze publikációs listáját. A Web of Science felületén megtalálható saját publikációk az adatbázisban történő keresést követően a közlemények szerzőhöz rendelésével vehetők fel a listára, de az adatbázis által nem indexelt publikációk is hozzáadhatók. A Web of Science heti rendszerességgel frissíti a ResearcherID azonosítókhoz rendelt adatbázisban

szereplő közleményeket a My Publications listák alapján, azoknál a szerzőknél, akik nyilvánossá tették profiljukat. A saját publikációkat regisztráló lista gyakori frissítése és pontos vezetése tehát fontos ahhoz, hogy a Web of Science felületén az azonosító alapján egyértelműen megtalálhatóak legyenek az adott szerző adatbázisban szereplő közleményei.

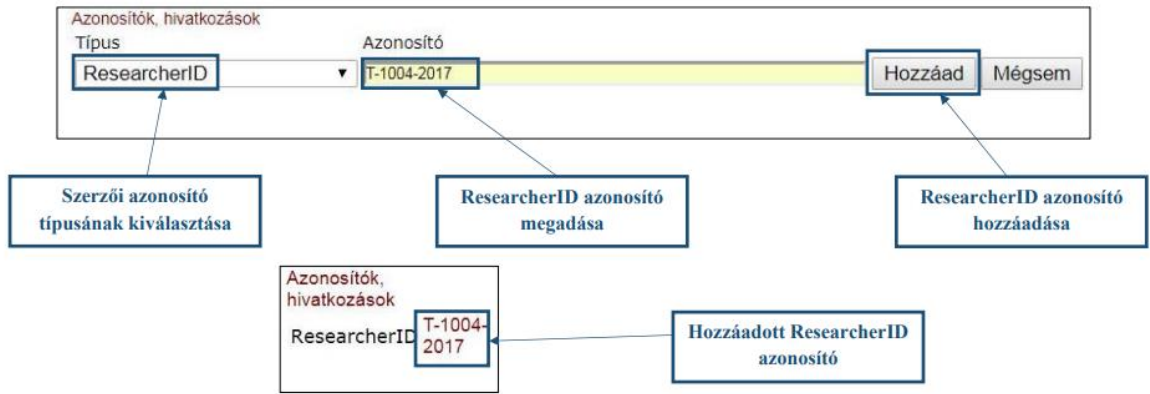
Az Add vagy Add Publications gombra kattintást követően megjelenő új oldalon a Web of Science adatbázisból történő adatáttemeléshez és a közlemények szerzőhöz rendeléséhez juthatunk. További lehetőségekként az EndNote bibliográfiakezelő program webes felületéről történő adatimportálás, illetve saját RIS formátumú fájlként publikációs lista feltöltése érhető el. A Web of Science adatbázisban való keresés és az adatok importálása, azonban csak az adatbázis hozzáférési jogosultságával rendelkező felhasználóknak érhető el.

publikáció hozzáadása a My Researcher Profile oldalon

A Search Web of Science Core Collection opció kiválasztása után megjelenő oldalon a szerző vezetéknevének és a keresztnév kezdőbetűjének megadása kötelező ahhoz, hogy a Search gombbal közleménykeresést lehessen indítani az adatbázisban. A pontosabb eredményhez azonban érdemes további kritériumokat is beállítani.

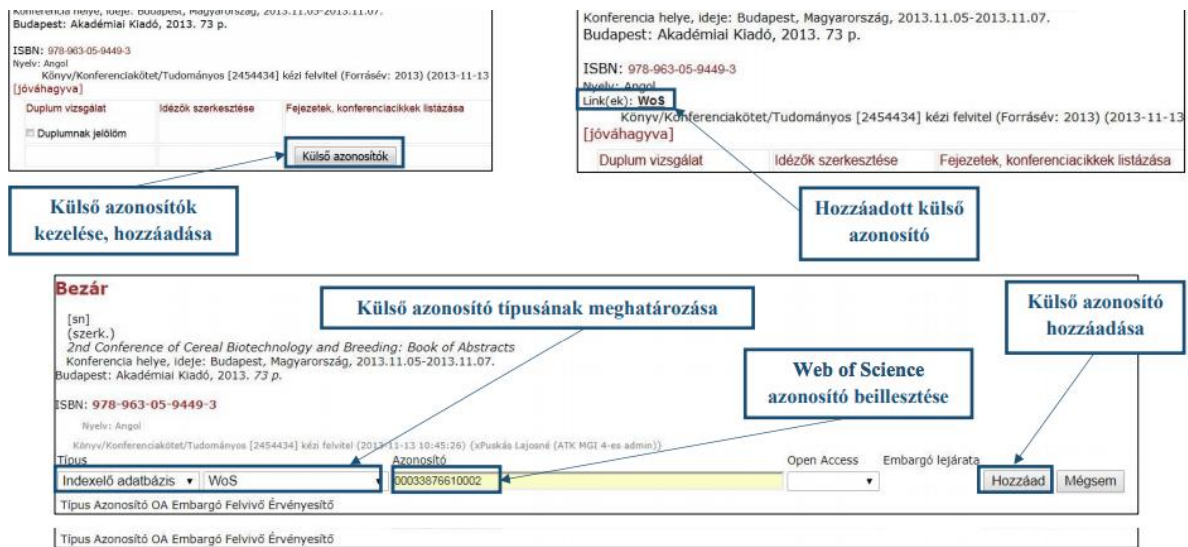
A találati listán a szerző által írt közlemények egyenként vagy csoportosan is kijelölhetők, majd az Add feliratra kattintva hozzáadhatók a szerző ResearcherID profiljának My Publications listájához. A közleményekkel együtt az idézettségi adatok is áttemelésre kerülnek. A saját közlemények egyértelmű beazonosításához érdemes a címre kattintva a részletes adatokat is előhívni.

ResearcherID azonosítót csatolni lehet az MTMT-ben a szerzői adatlaphoz.



ResearcherID azonosító csatolása szerzői adatlaphoz MTMT-ben

Az MTMT felületén a szerző a személyi adatlapjának Azonosítók, hivatkozások gombjára kattintva tud a profiljához szerzői azonosítókat csatolni. A linkként működő ResearcherID hozzáadásával az MTMT felületéről egy kattintással elérhetővé válnak a szerző Web of Science adatbázisban megtalálható publikációi és azok idézettségi adatai. Az azonosító hozzáadásához az Azonosítók, hivatkozások felírra kattintva megjelenő Típus mező legördíthető menüjéből a ResearcherID lehetőséget kell választani és az Azonosító mezőbe be kell másolni a ResearcherID azonosítót. Ezt követően a Hozzáad gombra kattintva az azonosító megjelenik a szerző személyi adatlapján.



ResearcherID azonosító szerkesztése az MTMT rendszerében

Az MTMT rendszerében a lehetőség van WoS azonosítók feltöltésére vagy szerkesztésére. A Web of Science azonosító hozzáadásához a Típus mezők legördíthető menüiből az Indexelő adatbázis és a WoS lehetőségeket kell választani. Ezt követően az Azonosító mezőbe beilleszthető az azonosító számsor, majd a Hozzáad gombbal csatolható a közleményhez. A Bezár gombra kattintva az azonosító megjelenik a közlemény alatt „Link(ek): WOS” formában, amire kattintva

elérhető a publikáció Web of Science rekordja. Ha a Web of Science adatbázisból történő importálással került be az MTMT rendszerébe a publikáció, akkor az azonosító automatikusan áttemelésre került.

2.9.1.5 Web of Science

A Web of Science nem szerzőket, hanem tudományos műveket azonosító szolgáltatás. A Web of Science rendszeresen frissülő hivatkozáskereső bibliográfiai adatbázis, amely számos tudományterület rangos szakfolyóiratait, szakkönyveit és konferenciaanyagait szemlézi. Lehetővé teszi ezáltal a publikációkra történt hivatkozások megtalálását, valamint a kutatási trendeknek és az intézmények tudományos életben betöltött pozíciójának a vizsgálatát.

Az adatbázis felületén egy szerző publikációi és azok adatai a Basic Search felületen szerző szerinti kereséssel érhetőek el.

The screenshot shows a Web of Science article page for the paper "Spatial Variability in Litterfall, Litter Standing Crop and Litter Quality in a Tropical Rain Forest Region". The page is divided into several sections:

- Article Title and Authors:** Spatial Variability in Litterfall, Litter Standing Crop and Litter Quality in a Tropical Rain Forest Region. By: Piment, SA; Piment, Scott A; Cargiles, DA; Gorglen, Robert A; Siro, LP; Siro, Luke P; Valdez-Ramirez, V; Valdez-Ramirez, Vanessa; Williams, SE; Williams, Stephen E.
- Abstract:** Understanding the spatial variability in plant litter processes is essential for accurate comprehension of biogeochemical cycles and ecosystem function. We assessed spatial patterns in litter processes from local to regional scales, at sites throughout the wet tropical rain forests of northern Australia. We aimed to determine the controls on: (i) climate, soil plant community composition on annual litter standing crop, annual litterfall rate and in situ leaf litter decomposition. The level of spatial variance in these components, and leaf litter N, P, Ca, lignin, alpha-cellulose and total phenolics, was determined from within the scale of subregion, to site (1 km transects) to localplot (similar to 30 m²). Overall, standing crop was modeled with litterfall and its chemical composition, in situ decomposition, soil N₂ and topography (R² = 0.69, 36 plots). Litterfall was most closely aligned with plant species richness and stem density (negative correlation); leaf decomposition with leaf P and lignin, soil N₂ and dry season moisture (R² = 0.20, 48 plots). The predominant scale of variability in litterfall rates was local (plot), while litter standing crop and alpha-cellulose variability was more evenly distributed across spatial scales. Litter decomposition, N, P and phenolics were more aligned with subregional differences. Leaf litter C, lignin and Ca varied most at the site level, suggesting more local controls. We show that variability in litter quality and decomposition are more easily accounted for spatially than litterfall rates, which vary widely over short distances possibly in response to idiosyncratic patterns of disturbance.
- Keywords:** Australia; climate; decomposition; litter quality; litter standing crop; litterfall; spatial variability; tropical rain forest.
- Document Information:** Document Type: Article; Language: English; Accession Number: WOS:000338766100002; ISSN: 0006-3606; eISSN: 1744-7429.
- Journal Information:** Impact Factor: Journal Citation Reports.
- Citation Network:** 7 Times Cited, 61 Cited References, View Related Records, Create Citation Alert.
- All Times Cited Counts:** 8 in All Databases, 7 in Web of Science Core Collection, 6 in BIOSIS Citation Index, 1 in Chinese Science Citation Database, 0 in Data Citation Index, 0 in Russian Science Citation Index, 0 in ScELO Citation Index.
- Usage Count:** Last 180 Days: 3, Since 2013-11: Learn more.
- Most Recent Citation:** Fu, Changjun. Seasonal Dynamics of Litterfall in a Sub-Alpine Spruce-Fir Forest on the Eastern Tibetan Plateau. *Alpine Botany*.

publikáció és publikáció azonosító a Web of Science oldalon

A Web of Science adatbázisban minden publikáció rendelkezik egyedi azonosítószámmal, amely az MTMT rendszerében úgynevezett külső azonosítóként használható. Az MTMT-ben szereplő közleményekhez csatolható külső azonosítók linkként működnek, így általuk lehetővé válik az adott mű adatainak vagy teljes szövegének elérése, ellenőrzése.

2.9.1.6 Google Scholar

A Google Tudós (Google Scholar) a Google által kifejlesztett és fenntartott tudományos keresője, amely különféle témájú és formájú tudományos publikációk között keres. 2004. november 18-án indult, mára szinte minden online elérhető referált folyóiratban keres. 2006. február 20-ától a magyar könyvtárakban is keres a MOKKA-n keresztül.

A Google Scholar indexelő algoritmusa egy-egy publikáció fontosságának meghatározásakor azt veszi elsősorban figyelembe, hogy más cikkek milyen gyakran hivatkoznak rá. A keresések a publikációk teljes szövegében történnek, szűkíthetők szerző, év vagy megjelenési hely alapján, a

találatoknál a kereső megjeleníti a cím és az esetleges online elérhetőség mellett azt is, hogy milyen más tudományos munkák hivatkoznak rá, és milyen könyvtárakban található meg. Utóbbihoz a WorldCat adatbázist, és egy saját, Library Links nevű szolgáltatást használ.

A Google Scholar tevékenységével szemben megfogalmazódott az a kritika, hogy nem csak nem nyújt, de nem is akar teljes körű szolgáltatást nyújtani, mert csak a neten elérhető adatokra támaszkodik, márpedig a weben eddig a tudományos szakirodalomnak csak kisebb része jelent meg. Természetesen a hagyományos világ digitalizálásával, a retrospektív feldolgozó munkák befejeztével ez a kritika érvényét vesztheti.

a Google Scholar keresőfelülete

A Google Tudós esetén lehetséges szerzőként regisztrálni, ami után a szerző közleményeit és ennek hivatkozásait automatikusan a szerzői adatbázishoz rendeli. A gépi kereséseknek köszönhetően ez az összerendelés nem mindig pontos, de van lehetőség az adatok kézi szerkesztésére. Automatikus értesítést lehet kérni a cikkek új idézeteiről.

A Google Scholar többfajta mérőszámot is számol a kutatók tudományos hatásának mérésére.

2.9.1.7 Scopus

A 2004 novemberében indított Scopus szolgáltatás a lektorált szakirodalom legnagyobb absztrakt és citátum adatbázisa, amely intelligens eszközökkel kiegészítve lehetővé teszi a kutatások nyomon követését, elemzését és megjelenítését. A több mint 5000 nemzetközi

kiadótól származó, több mint 19000 kiadvány mellett a SciVerse Scopus gyors, könnyű és átfogóan használható eszközt nyújt a kutatóknak, amely jól megfelel a kutatási igényeknek a természettudományok, a műszaki tudományok, az orvostudomány, a társadalomtudományok, és az utóbbi időben már a humán tudományok területein is.

Scopus Preview [Author search](#) Sources Help ▼ [Register >](#) [Login ▼](#)

83,944 author results [About Scopus Author Identifier >](#)

Author last name "smith" [Edit](#)

Refine results

[Limit to](#) [Exclude](#)

Source title ▼

Affiliation ▼

City ▼

Country/territory ▼

Subject area ▼

[Limit to](#) [Exclude](#)

Sort on: [Document count \(high-low\)](#) ▼

All ▼ [Set document feed](#) [Request to merge authors](#)

	Author	Documents	Subject area	Affiliation	City	Country/Territory
<input type="checkbox"/> 1	Davey Smith, George Smith, George D. Smith, George Davey Davey-Smith, George	1654	Medicine ; Biochemistry, Genetics and Molecular Biology ; Agricultural and Biological Sciences; ...	University of Bristol	Bristol	United Kingdom
	View last title ▼					
<input type="checkbox"/> 2	Smith, James G. Smith, Jim Smith, J. Smith, J. G.	1301	Physics and Astronomy ; Mathematics ; Materials Science; ...	University of Colorado at Boulder	Boulder	United States
	View last title ▼					
<input type="checkbox"/> 3	Smith, Richard D. Smith, Richard O. Smith, R. D. Smith, R.	1105	Biochemistry, Genetics and Molecular Biology ; Chemistry ; Medicine; ...	Pacific Northwest National Laboratory	Richland	United States

keresési eredmények a Scopus oldalán

A SciVerse Scopus fókuszában 2012 májusában több mint 19,500 kiadvány állt. Ezek közül 18,500 lektorált folyóirat (ezen belül 1,800 Open Access folyóirat), 400 szakmai kiadvány, 340 könyvsorozat volt. Az adatbázisban a 47 millió rekord közül 26 millió rekord 1996-ig visszamenő referenciákkal (melyekből 78% tartalmaz referenciát) rendelkezik, 21 millió rekord 1996 előttről, egészen 1823-ig visszamenően dolgozza fel az adatokat.

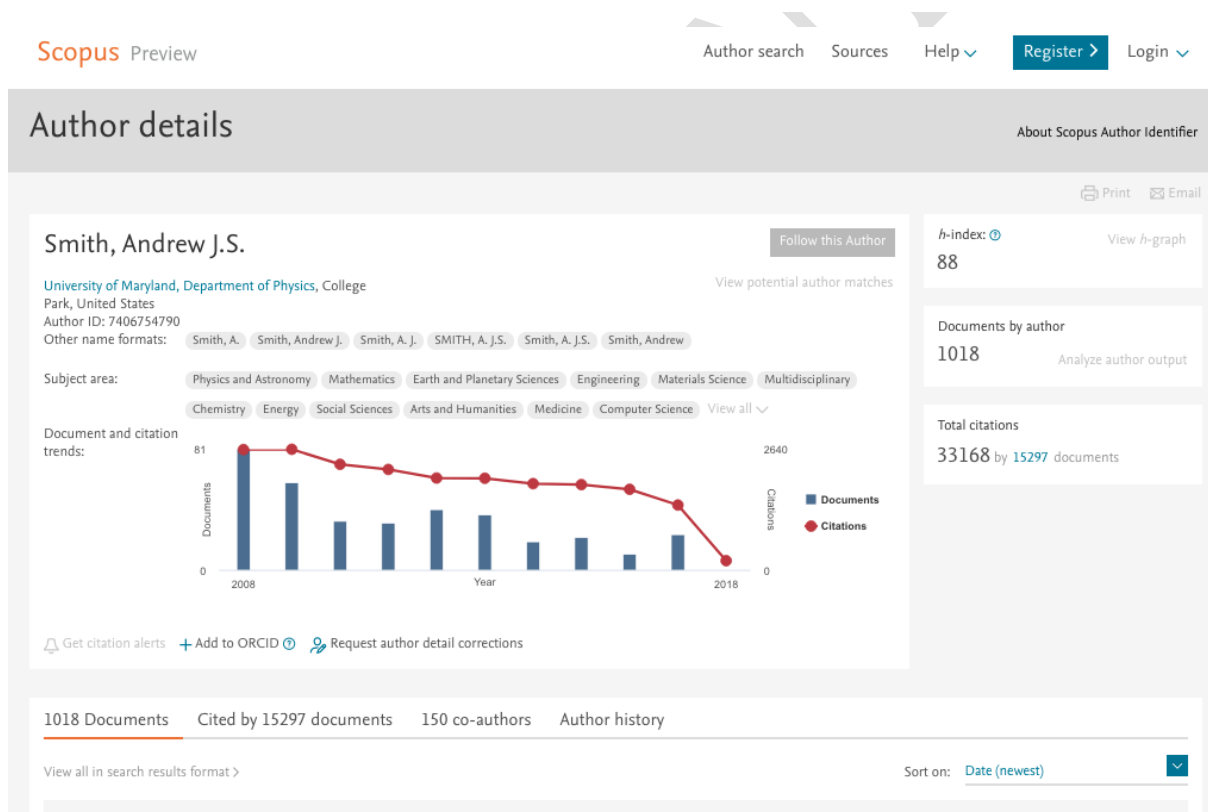
A folyóiratok megoszlása tudományterületek szerint:

- kémia, fizika, matematika és mérnöki tudományok: 4500 folyóirat;
- élet- és orvostudományok: 5900 folyóirat (benne a teljes Medline adatbázis);
- társadalomtudományok, pszichológia és közgazdaságtudomány: 2700 folyóirat;
- biológia, agrártudományok és környezetvédelem: 2500 folyóirat;
- általános tudományok: 50 folyóirat.

A Scopus adatbázis folyóiratainak 52%-a európai kiadású.

A Scopus szolgáltatásai között több olyan van, amely valamiféle névtér-kezdeményezésnek is tekinthető.

- Linkeket biztosít a teljes szövegű cikkekhez és egyéb könyvtári forrásokhoz.
- Mindenkinek biztosít egy azonosítót (Author Identifier), amelynek révén automatikusan hozzá lehet rendeni a szerző publikált kutatásaihoz a h-indexet (Hirsch indexet) a tudományos kutatás hatásának mérésére.
- Az idézet-nyomkövetővel (Citation Tracker) egyszerűen lehet keresni, ellenőrizni, követni a cikkek hivatkozásait valós időben.
- A rendszer biztosít egy intézményi azonosítót is (Affiliation Identifier), amely révén automatikusan azonosítani lehet, és az intézményhez lehet rendelni az összes kutatási eredményt.
- A Journal Analyser segítségével könnyen képet lehet kapni a folyóiratok teljesítményéről.
- Kiépített kapcsolat van a SciVerse ScienceDirect, Reaxys, a ProQuest CSA Illumina irányába.
- Az adatokat könnyű exportálni a bibliográfia-kezelő alkalmazásokkal, mint például a RefWorks, EndNote and BibTex.



személyi adatlap a Scopus oldalán

A Scopus rendszerét szerzői adatbázisként is lehet használni, ahol a cikkek a Google Scholarhoz hasonlóan automatikusan a szerzőhöz rendelődnek hivatkozásaikkal együtt. Előfordulhat téves vagy hiányzó közleményhozzárendelés, illetve az is, hogy ugyanazon szerzőt több szerzőként tartja nyilván a rendszer - ilyen esetekben korrekciót lehet kérni.

Az innen exportált cikkek és hivatkozások adatai az MTMT-be importálhatók. Szinkronizálható az ORCID adatbázissal. Automatikus értesítést lehet kérni a szerzői cikkek új idézeteiről, a cikkek adatbázisba kerüléséről.

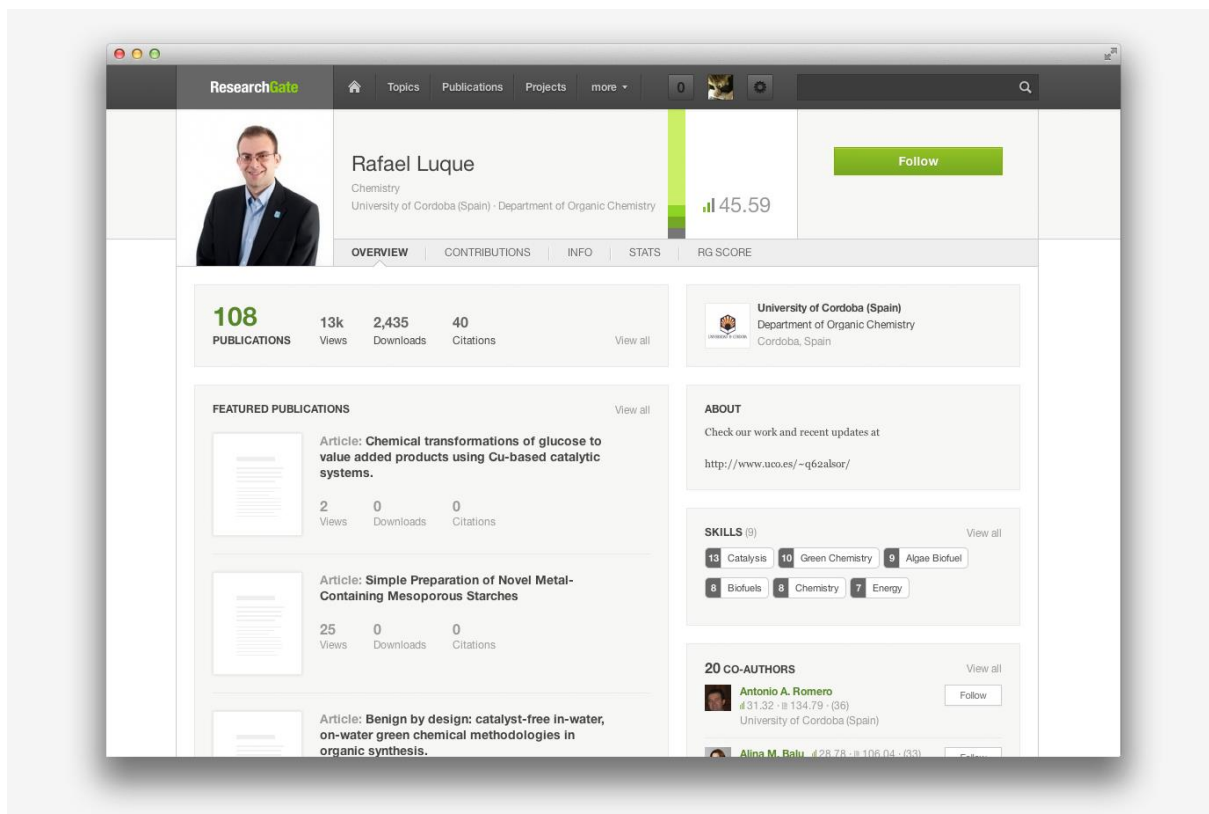
2.9.1.8 ResearchGate

A ResearchGate egy díjmentes közösség háló alapú internetes oldal, amely a különféle tudományágzatokban tevékenykedő kutatók számára lett kifejlesztve. Az oldal szabad hozzáférést kínál a különböző szakfolyóiratok cikkeihez, melyek egy több mint 30 millió bejegyzést tartalmazó, és folyamatosan bővülő adatbázisban található.



a ResearchGate oldala

A keresés kezdetekor a felhasználó keresőbe beírta, például „orvosi folyóiratcikkek” kulcsszó megadása után, máris megtekintheti a hasonló témájú, rendelkezésre álló releváns, időrendi sorrendben megjelenő tanulmányokat. Ugyanezzel a funkcióval kereshetők azok a kollégák a világ minden tájáról, akik az adott tudományágban dolgoznak. A ResearchGate 2009 óta a tagoknak nyílt hozzáférést biztosít a szabadabb publikáláshoz. A feltöltött dokumentumok bekerülnek a belső források közé, melyeket a tagok a szabadon letölthetnek, megtekinthetnek. Szolgáltatásai közé tartozik továbbá a fájlmegosztás, fórumok, módszertani beszélgetések, vitacsoportok stb. Ezen funkciók használatának kiinduló feltétele egy személyes kutatóprofil létrehozása. A platform tartalmaz egy úgynevezett állásbörzét, mely a világszerte kínáló tudományos munkakörű állásokat gyűjti össze. A munkakeresés során megadhatók választási szempontok, mint kulcsszó, pozíció, munkaterület, ország stb.



személyi adatlap a ResearchGate oldalán

Sok tudományos szervezet, mint az International Academy of Life Sciences (IALS), az European Science Foundation, a ResearchGate-et használják közös találkozási pontként, ahol a tagok és résztvevők megoszthatják az adatokat, lefolytathatják a szakmai vitákat, megbeszéléseket. A platform ezért kifejlesztett egy úgynevezett subcommunities-t is, melyen belül csak az adott intézmény tagjai kommunikálhatnak. Számos nemzetközi egyetem professzorai garantálják az Advisory Board (tanácsadó szolgáltatás) minőségét és függetlenségét. A ResearchGate, melynek központja Bostonban és Berlinben található, 2011. májusi megalapítása óta a világ 192 országában van jelen több mint 1400000 taggal.

2.9.1.9 Academia.edu

Academia.edu egy amerikai nonprofit közösségi háló tudósok számára, amely prémiumszolgáltatásokat is nyújt fizetés ellenében. A platformon lehetőség nyílik a publikációk megosztására, az impact faktor monitorozására, és az adott tudományterületen zajló kutatások nyomon követésére. A honlap 2008 szeptemberében indult, és 2017 decemberére 36 milliós egyedi látogatót ért el, és addigra a felhasználók több mint 20 millió anyagot töltöttek fel. Az Academia.edu alapítója Richard Price.



István Szakadát

Budapest University of Technology and Economics, Department of Sociology and Commu... Sociology +1

István Szakadát graduated as mathematical engineer in 1982, and as a sociologist in 1985. He got his PhD in Sociology in 1999. From 1987 he is working for the Technical University of Budapest as a teacher at the Department of Sociology. During his academic career he w... [more](#)

159 Followers | 87 Following | 6 Co-authors | 2,037 Total Views

[+ FOLLOW](#)

PAPERS



A HunNER korpusz

by [Eszter Simon](#) és [István Szakadát](#)

IV. Magyar Számítógépes Nyelvészeti Konferencia, 2006

[Download](#) 15 Views



Hunpars: mondattani elemz alkalmazás

by [Anna Babarczy](#) és [István Szakadát](#)

salmon.nyttud.hu

Kivonat: A Hunpars-projekt célja egy nyílt forráskódú elemz alkalmazás létrehozása, amely automat... [more](#)

[Download](#) 7 Views

személyi adatlap az Academia.edu oldalon

A honlapon a felhasználók létrehozhatják a saját személyes oldalukat, feltölthetik a publikációikat, megadhatják saját kutatási, illetve érdeklődési területeiket, és böngészhetnek a hasonló érdeklődésű személyek profiljaiban. Az oldalnak 2017 decemberében 58 millió felhasználója volt világszerte.

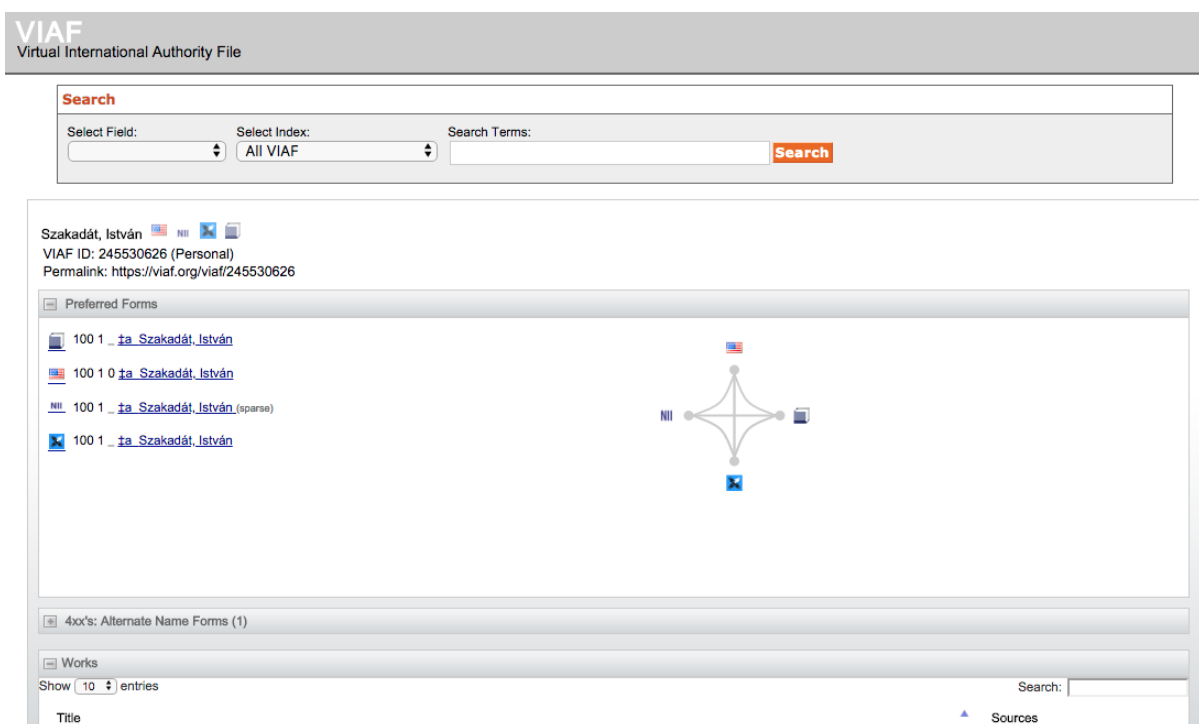
2.9.2 További azonosító rendszerek személyekre, helyekre

A személyek, testületek, helyek azonosítására természetesen nem csak a kutatók esetében van szükség, más szakmai területeken is alkalmaznak ilyen rendszereket, sőt, a közgyűjtemények világában már jóval előbb belekezdtek az azonosító rendszerek kiépítésébe.

2.9.2.1 VIAF

A VIAF (Virtual International Authority File, Nemzetközi Virtuális Besorolási Állomány) egy, az OCLC (Online Computer Library Center, Online Számítógépes Könyvtári Központ) által üzemeltetett, nyilvános katalógus rendszer. Célja a különböző nyelvterületek eltérő katalogizálási szabályai alapján kialakított névanyagok egész világon elérhető virtuális katalógusba rendezése.

A VIAF az országoként, nyelvenként, hagyományokként eltérő névvariánsokat, és személyekhez tartozó adattartalmakat kapcsolja össze.



szerzői oldal a VIAF katalógusában

A VIAF a személyeken túl további entitásokat (pl. testületeket, földrajzi helyeket, műveket, kiadványokat) is felvesz az azonosítottak közé, és az azonosítókat nem különíti el az entitástípusok szerint. A VIAF előnye az, hogy nemzetközi azonosító rendszer, és komoly nemzeti közgyűjtemények csatlakoztak hozzá, amelyek folyamatosan küldik be saját authority-rekordjaikat. A VIAF weboldalán lehetséges a teljes állományban is keresni, de mód van arra is, hogy a csatlakozott intézmények, illetve az entitástípusok szerint lehessen szűrni.

Az IFLA-n belül már 1978-ban felvetődött egy közös, nemzetközi besorolási állomány létrehozásának az igénye, a megvalósításra azonban még 2003-ig várni kellett. Ekkor indította el a VIAF-ot a Német Nemzeti Könyvtár (Die Deutsche Bibliothek), a Kongresszusi Könyvtár (Library of Congress) és az OCLC (Online Computer Library Center). A VIAF-ra indulásakor a katalogizálás segédeszközeként tekintettek, ám az elindulása után más funkcióra is használni kezdték. Az OCLC felmérése szerint a Linked Data alapú adatsere-folyamatokban a VIAF a leggyakrabban használt adatforrás. Az új szerepből következnek az új elvárások, amelyek értelmében a VIAF-nak – LD-szolgáltatóként – hiteles és megbízható adatokat kell tartalmaznia.

A VIAF-azonosító egy szám, és a különböző entitásokhoz rendelt számokat sorban osztják ki, nem különítik el egymástól őket. Az alábbi VIAF-szám Budapest azonosítója.

VIAF: 154759119

A VIAF – a partnerintézmények által beküldött adatoknak köszönhetően – folyamatos frissül, azonban ez egyfajta függést is jelent, ráadásul a könyvtári forrásadatbázisokkal való folyamatos kapcsolat veszélyezteti a teljes adatállomány stabilitását, perzisztenciáját, ami felveti ellenőrző mechanizmusok beiktatásának szükségességét.

Az Országos Széchényi Könyvtár a MEK-ben 2013. július 16-tól folyamatosan felveszi az egyes művek szerzőinek a VIAF azonosítóját.

CÉDULA	CÍMKÉS	HUNMARC	MARC21	XML
CÉDULA FORMÁTUM				
<p>MEK-13611 Parlamenti képviselői helyek megoszlásának becslése közvéleménykutatási adatok alapján / Mészáros József, Szakadát István tanulmány(ok) ; magyar (Tárki társadalompolitikai tanulmányok ; 6.) eredeti kiadvány: Parlamenti képviselői helyek megoszlásának becslése közvéleménykutatási adatok alapján / Mészáros József, Szakadát István Budapest : Tárki, 1998 (Tárki társadalompolitikai tanulmányok, ISSN 1418-0839 ; 6.) ISBN 963 7869 12 3 OSZK: http://nektar.oszk.hu/hu/manifestation/327913 <i>Magyar belpolitika, Kormányzat, parlament, Politológia parlamenti választás, választási rendszer, közvélemény-kutatás, társadalomstatisztika, prognosztika, Magyarország, 1990-es évek</i> MEK-be került: 2014-12-17 URL: http://mek.oszk.hu/13600/13611 Jogi közlemény: Jogvédett.</p>				

a szerzők neve mögött a VIAF kék azonosító ikonja a MEK katalóguscéduláján

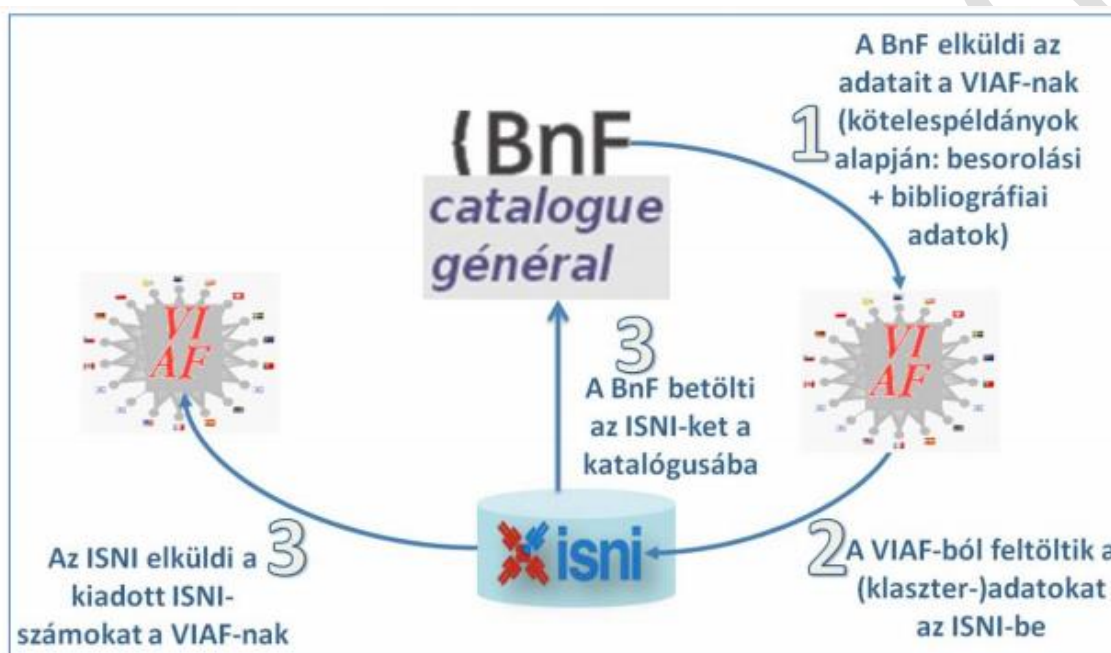
2.9.2.2 ISNI

A könyvtári világban már ismert és elterjedt VIAF-azonosító mellett létezik egy másik nemzetközi szabványos megoldás is, az ISNI-kód (nemzetközi szabványos névazonosító – ISO 27729:2012 Information and documentation – International standard name identifier). Az ISNI célja hatékony eszközt biztosítani az alkotók, a produkciós iparban és a tartalomszolgáltatásban szereplők pontos, egyértelmű azonosításához. Az ISNI nem ad közvetlen hozzáférést az adott személyekről szóló információkhoz, de kapcsolatot teremt a közismert névváltozatok között, egyértelműsíti, hogy kiről van szó, és ezzel lehetővé teszi a kereshetőséget az olyan rendszerekben, amelyekben a személyekről széles körű információk találhatók.

Az ISNI névazonosító kód két különböző problémakörre is megoldást nyújt: egyrészt a különböző információs rendszerek egymástól eltérő azonosítási szisztémát alkalmazhatnak, amelyeket kívánatos lenne összhangba hozni egymással, másrészt pedig a nyilvánosság által ismert nevek, illetve a nevekkal jelölt személyek egyértelműsítését, beazonosítását is segíti.

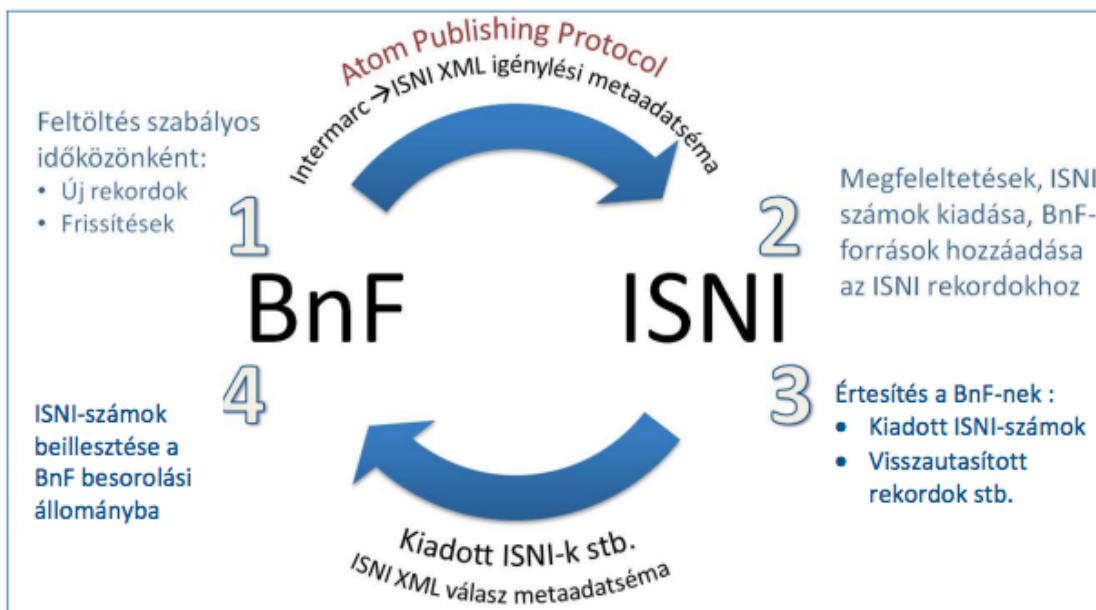
A VIAF működésével kapcsolatos gondok megoldásaként javasolták sokan a nemzetközi szabványos névazonosító, az ISNI használatát. Utóbbi épít a VIAF-ra, ugyanakkor humán erőforrást (egy ún. „Quality Team”) biztosít az állomány karbantartására, a névazonosságokból

fakadó hibák kiküszöbölésére. Az ISNI nemzetközi szabvány, nemzetközi névtérnek tekinthető, ami miatt bevezetése, alkalmazása kérdéseket vet fel közgyűjtemények, könyvtárak munkafolyamatainak újratervezésében. Példaként érdemes áttekintenünk, hogy milyen módon alakított ki kapcsolatot a francia nemzeti könyvtár (BnF) az ISNI-vel annak érdekében, hogy elmozduljon az adatok újrafelhasználását támogató automatizált eljárások irányába, összehangolva az újszerű technológiák alkalmazásából fakadó előnyöket a nemzeti könyvtári feladatokkal és minőségi elvárásokkal (Dancs 2017). A BnF-ben nem a könyvtárosok munkájának teljes kiváltására, sokkal inkább emberek által ellenőrzött automatizmusok, azaz félautomatizmusok implementálására törekednek. Az alábbi ábrán az látható, ahogyan a BnF kapcsolatba kerül a VIAF-fal, ISNI-vel.



BnF-ISNI adatcsere

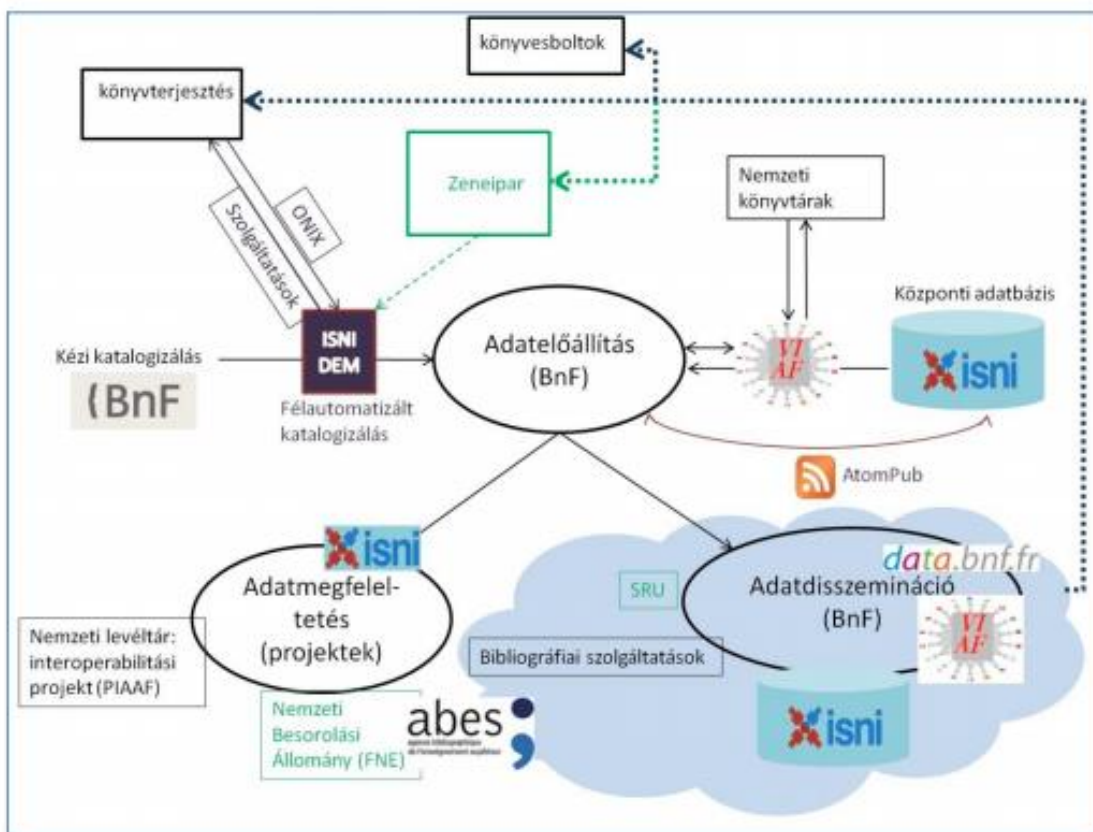
A könyvtáron kívüli adatfolyamok becsatornázása az ISNI-számigénylés megvalósításával történik. Fontos szerepet játszik ezen a ponton a könyvterjesztésben alkalmazott ONIX-szabvány. A félautomatizált katalogizálás azt jelenti, hogy a bibliográfiai leírásokat készítő munkatársak építhetnek a könyvpiari értéklánc más pontjairól (pl. kiadók, könyvterjesztők) érkező rekordokra.



BnF-ISNI közvetlen (automatizált) adatkapcsolat

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

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



adatfolyamok az optimalizálás után

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

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

A VIAF és az ISNI közötti eltérések

Tudvalevő, hogy a VIAF elsődleges felhasználási területe a könyvtárak számára az adataik láthatóságának növelése, míg az ISNI a kiadókkal, könyvterjesztőkkel és jogkezelő szervezetekkel való együttműködés szempontjából jelent elsősorban perspektívát. A két azonosító a közöttük lévő átfedések ellenére mind műszaki értelemben, mind a finanszírozást illetően különbözik. Az eltéréseket az alábbi táblázat foglalja össze.

	VIAF	ISNI
Fenntartás/ Működtetés	az OCLC szolgáltatása, amelynek fejlesztését, működését egy a könyvtárak képviselőiből álló tanácsadó testület (VIAF Council)	nemzetközi szabvány, fenntartása a Nemzetközi ISNI-Ügynökség hatáskörébe tartozik, technikai háttéréről az OCLC gondoskodik

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

		klasztereknek utal ki ISNI-számot, amelyek megbízhatósága elér egy meghatározott szintet
Frissítések	a VIAF havonta frissül	folyamatos frissülés az ügynökségek igényei alapján, illetve az online számkiutalások tükrében, beleértve a Minőségcsoport rekordszerkesztési műveleteit, illetve a végfelhasználói igénylések kezelését
Adatdisszemináció és licenc	a VIAF nyílt kapcsolt adatként terjeszthető ODC-By licenc10 alatt	az ISNI egy hivatkozható nyilvántartás, amely egy API-n keresztül nyilvánosan lekérdezhető; az adatok nyílt licenc alatt felhasználhatók, jelenleg tervezik nyílt kapcsolt adattá fejlesztését
Célközönség	könyvtárak, a kulturális örökség intézményei, oktatás és kutatás	globális és több szektort érint (könyvtárak, levéltárak, kutatás, zene- és filmipar, könyv- és folyóirat-kiadás, jogkezelők stb.)
Példa a felhasználásra	a katalogizálás forrása, a könyvtári entitások összekapcsolása repozitóriumokkal (HUB-funkció), mint például a Wikipedia és a Wikidata	linkek automatizált létrehozása a besorolási rekordokhoz a kiadói és könyvtári adatok összekapcsolásával

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

Az ISNI különös adottsága, ami ugyancsak megkülönbözteti a könyvtári besorolási rekordok nemzetközi szintű azonosítására használt VIAF-tól, hogy különbséget tesz személy és személyiség között, amit a személyes adatok védelmének (privacy) kezelésével valósít meg. Ez közelebről azt jelenti, hogy lehetőség van arra, hogy egy személy, ha érdeke úgy kívánja, több személyiséget (álnevet) használjon anélkül, hogy a személyiségek, illetve a hozzájuk kapcsolódó nevek közötti kapcsolat nyilvános volna. Az ISNI névterét használó könyvtári rendszerekbe tehát a kapcsolati információ alaphoz nem kerül át. A nyilvánosság ugyanakkor – természetesen – a későbbiekben módosítható.

Az ISNI-kód egy négy blokkból álló számsorozat. Lentebb egy magyar politikus ISNI-száma látható.

ISNI: 0000 0001 2141 7284

Az ISNI oldalára ugorva bibliográfiai rekord leírásához hasonló oldalt kapunk, ahol a – különböző forrásokból származó – névváltozatokon túl elérhetők még a személyhez kapcsolható más személyek linkjei, a személyhez köthető művek listája és egyéb megjegyzések.

shortlist	title data	search history	1
results search [or] ISN:0000000121417284 1 hits			
labels	sources data	marc21	

Please help us improve this record

If you have any supplemental information about the identity listed here, please click in this box to go to the contribution form.

Thank you in advance!

ISNI: 0000 0001 2141 7284

Name: Orbán, Viktor
Orbán Viktor (magyar politikus, a Fidesz elnöke)
Victor Orbán
Viktor Orbán
Viktor Orbán (atual primeiro-ministro da Hungria)
Viktor Orbán (Hongaars politicus)
Viktor Orbán (Hungarian politician)
Viktor Orbán (Hungarian politician, chairman of Fidesz)
Viktor Orbán (personnalité politique hongroise)
Viktor Orbán (politico húngaro)
Viktor Orbán (politico ungherese)
Viktor Orbán (ungarischer Politiker)
Viktor Orbán (ungarsk politiker)
Viktor Orbán (wegierski polityk)
Viktors Orbāns
Виктор Орбан
Виктор Орбан (унгарски политичар)
Виктор Орбан
Orbán, Viktor
אורבאן, ויקטור
فيكتور أوربان
(ويكتور اربان (سيناستعداد اهل مجازستان)
ভিক্টর অরবান
오르반 빅토르
オルバーン・ヴィクトル
奧班·維克多

Dates: 1963-

Creation class: Language material
Im
Projected medium
Text

Creation role: author

Related names
Titles
Notes

az ISNI oldala Orbán Viktorról

2.9.2.3 PIM személynévtér

Magyarországon a Petőfi Irodalmi Múzeumban kezdtek el elsőként építeni személynévteret, amelynek azonosítóit ma már a személyek magyar Wikipédia-oldalaira is beillesztik a többi globális személynév-azonosító mellé. A PIM személynévtér számmal azonosítja a személyeket.

<https://opac-nevter.pim.hu/record/-/record/PIM123214>

.A névtér oldalain elérhető a PIM adatbázisaiból a kiválasztott személyhez kapcsolható adatok teljes köre úgy, a személy életrajzi adatai, személyes bibliográfiai adatai, esetleges kitüntetései stb.

Tarlós István

Database: INT; IKN
Unified name form: Tarlós István
Date of birth/death: 1948-
Profession: gazdasági szakmérnök; mélyépítő mérnök; polgármester
Birthplace: Budapest
Date of birth: 1948. V. 26.
Schools/Carreer: III. ker. Önkormányzat

▼ Award:
▼ Data sources:
▼ Wiki:
Works

Link to this record: <http://resolver.pim.hu/auth/PIM123214>

a PIM személynévtér oldala Tarlós Istvánról

2.9.2.4 ULAN

A Getty Intézet 1984-ben kezdett bele a művészek alapadatainak gyűjtésébe, rendszerezésébe, és ennek eredményeként született meg az ULAN, a művésznevek egyesített listája (Union List of Artist Names), amely egy rögzített (korlátozott) szköppal rendelkező személynévtérként is értelmezhető. A szolgáltatás oldalán az olvasható, hogy magukat strukturált szótárként definiálják.⁵ Sokféle adat felvehető a rendszerbe, az alábbi adattípusokat kötelező kitölteni a rendszerben egy új elem felvételekor.

- record type
- name
- name source
- display biography
- nationality
- role
- birth date and death date

Az ULAN gyűjti és egymástól elkülöníti a művészek neveit, de csak a személyeket látja el egyedi azonosítóval, a neveiket már nem.

⁵ „Az ULAN egy strukturált szótár, amely művészek, mecénások, cégek, múzeumok és más kapcsolódó személyek vagy szervezetek neveit tartalmazza a műalkotások, építészeti alkotások és gyűjtemények területén” (<http://www.getty.edu/research/tools/vocabularies/ulan/about.html>).



Click the icon to view the hierarchy.

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

ID: 500339585

Record Type: Person

Mozart, Wolfgang Amadeus (Austrian composer, 1756-1791)

Names:

- Mozart, Wolfgang Amadeus** ([preferred](#),U,index,LC)
- Wolfgang Amadeus Mozart** (U,display)
- Mo□□□□art, Vol'fgang Amadei** (U)
- Mötsaruto** (U)
- Mo□□□□art, Iogann-Krizost Vol'fgang Gotlib** (U)
- Mozart, Johann Chrysostom Wolfgang Amadeus** (U)
- Mozart, W. A.** (U)
- Mozart, Wolfgang Amadeo** (U)
- Mo□□□□art, V. A.** (U)
- Mocartas, V. A.** (U)
- Motsart, Volphnkank Amedaios** (U)
- Mo□□□□art, Volfang Amadeus** (U)
- Mozzart, Apollo** (U)
- Mozart, Joannes Chrysostomus Wolfgangus Amadeus** (U)
- Mozart, Johannes Chrisostomus Wolfgangus Theophilus** (U)
- Mozhate** (U)
- Моцарт, Вольфганг Амадей** (U)
- מוצרט** (U)
- מוצרט, וולפגנג אמדאוס** (U)
- 莫札特** (U)

Nationalities:

Austrian ([preferred](#))

Roles:

composer (music) ([preferred](#))

Gender: male

Birth and Death Places:

Born: Salzburg (Salzburg, Austria) (inhabited place)
Died: Vienna (Vienna state, Austria) (inhabited place)

List/Hierarchical Position:

- Non-Artists
- Mozart, Wolfgang Amadeus (I,U)

Biographies:

(Austrian composer, 1756-1791) [VP Preferred]
(artist, 1756-1791) [GRL]

Additional Names:

Sources and Contributors:

- Моцарт, Вольфганг Амадей [GRL]
..... Library of Congress Authorities database (n.d.) n 80022788
- מוצרט [GRL]
..... Library of Congress Authorities database (n.d.) n 80022788
- מוצרט, וולפגנג אמדאוס [GRL]
..... Library of Congress Authorities database (n.d.) n 80022788
- 莫札特 [GRL]
..... Library of Congress Authorities database (n.d.) n 80022788
- Mocartas, V. A. [GRL]
..... Library of Congress Authorities database (n.d.) n 80022788
- Mo□□□□art, Iogann-Krizost Volfgang Gotlib [GRL]
..... Library of Congress Authorities database (n.d.) n 80022788
- Mo□□□□art, V. A. [GRL]
..... Library of Congress Authorities database (n.d.) n 80022788
- Mo□□□□art, Volfang Amadeus [GRL]
..... Library of Congress Authorities database (n.d.) n

Az ULAN azonosítóit gyakran feltüntetik a Wikipédia művészekről szóló oldalain is, és ebben a funkciójában már – korlátozott szókppal bíró – globális névtérként is tekinthetünk rá.

2.9.2.5 GDN

A Német Nemzeti Könyvtár besorolási adatállománya (Gemeinsame Normdatei) a német nyelvterületen használt rendszer, amely az RDA szabályzat alapján rendel egyedi azonosítókat többféle entitáshoz, így például személyekhez, testületekhez, földrajzi helyekhez, tárgyszavakhoz.

<small>GND</small>	
Link zu diesem Datensatz	http://d-nb.info/gnd/4008684-7
Geografikum	Budapest
Andere Namen	Haupt- und Residenzstadt Budapest Budapest / Tanács (Spitzenorgan) Székesfőváros Budapest Boedapest Ofenpest Ofen-Pest Aquincum (röm.) Budimpešta Budapesta Budapešta Budapeszt Budapešt Budapestum Buda Pest Ofen Pest-Buda
Quelle	B 2006 M Druckorte des 16. bis 19. Jhs., S. 36 Homepage (Stand: 10.01.2017): http://budapest.hu/
Erläuterungen	Definition: Hauptstadt Ungarns, 1.1.1873 durch Zusammenschluß von Buda, Obuda u. Pest gegründet
Zeit	1873-
Land	Ungarn (XA-HU)
Koordinaten	E 019°02'23" / N 047°29'52" E019.039909 / N047.498010
Vorgänger	Buda Pest (Stadt)

Budapest GDN-rekordjának részlete

Ezeket az adatokat szabadon hozzáférhetővé téve más szolgáltatások számára a GDN-azonosítók – valamely entitástípusra szűkítve a fókuszot – működhetnek globális névtérként is. A Wikipédia-oldalakon gyakran beágyazzák a tárgyalt dolog GDN-ID-jét, ami önmagában mutatja, megerősíti azt a tényt, hogy a GDN valóban globális névtérként funkcionál.

2.9.2.6 BNF

A Francia Nemzeti Könyvtár (Bibliothèque nationale de France, BnF) névazonosító rendszere ugyanúgy és ugyanazért szolgálhat globális névtérként, mint a német nemzeti könyvtár (GDN) besorolási állománya, a GDN. A Wikipédia a BNF-azonosítókat is gyakran elhelyezi a megfelelő oldalakon.

Budapest (Hongrie)

Vedette matière nom géographique. S'emploie en tête de vedette ou en subdivision.

La capitale hongroise Budapest est la réunion en 1872 de trois communes : Buda (Ofen en allemand), Óbuda (Aquincum=Altofen, la ville romaine) et Pest (Contra Aquincum ou Transaquincum)

<Employé pour :

- Altofen (Hongrie)
- Aquincum (ville ancienne)
- Buda (Hongrie)
- Bude (Hongrie)
- Contra Aquincum (ville ancienne)
- Óbuda (Hongrie)
- Ofen (Hongrie)
- Ofenpesth (Hongrie)
- Pest (Hongrie)
- Pesth (Hongrie)
- Transaquincum (ville ancienne)

<<Terme(s) générique(s) :

[Hongrie -- Divisions politiques et administratives](#)

Source(s) :

Pesth und Ofen mit ihren Einwohnern, besonders in medicinischer und anthropologischer Hinsicht, 1838 . - L'Hunyadi Janos, source purgative à Bude (Ofen en Hongrie), 1871 . - Az aquincumi amfiteátrumok, 1956
Révai új lexikona, 1998 . - Magyar helységnév-azonosító szótár, 1992 . - GDEL

Domaine(s) : 914

Correspondance(s) exacte(s) :

- LCA (Library of Congress Authorities) : Budapest (Hungary) <http://id.loc.gov/authorities/names/n79091691>

Notice n° : FRBNF11933405

Création : 81/02/12 Mise à jour : 13/03/01

Budapest BNF-rekordja

A BNF -azonosító egy karaktersorozat, amelyben lehetnek betűk is a számok mellett.

BNF: cb119334051

2.9.2.7 A könyvtárak besorolási állománya

Amikor személynevet, testületi vagy földrajzi nevet kell felvenni a bibliográfiai rekordokba, a könyvtárak a neveket írják be, nem pedig névazonosítókat a megfelelő mezőkbe. Ez a megoldás nem teszi lehetővé a nevek egyértelmű kezelését. Lényegében ez az oka annak, hogy kidolgozták és használatba vették a besorolási állományok (authority records) rendszerét. A besorolási állományok bevezetésének és fenntartásának célja egyfelől a nevek egyértelmű kezelése, a szinonimák és homonimák feloldása, tehát az egyértelműsítés, másfelől a kitüntetett névalak kiválasztása abban az esetben, ha egy névhordozóhoz (személy, testület, földrajzi hely) több név is tartozik. A kitüntetett névalakra sokáig azért volt szükség, mert csak ezáltal lehetett egyértelmű (egy az egyhez) kapcsolatot kialakítani a nevek és a névordozók között. Ezt a – jogos – igényt a digitális világban már másként lehet kielégíteni, így a kitüntetett névalak jelentősége csökken. Természetesen addig, ameddig a különböző szakmai (például könyvtárosi) közösségek igényt tartanak a preferált névalak minősítés kiosztására, addig ezt a lehetőséget fent kell tartani, de a nemzeti névterek építése és használata során, amikor többféle szakma képviselőinek kell egymással együttműködve dolgoznia, figyelembe kell venni azt is, hogy a preferált névalak besorolása szakmánként, kontextusonként változhat.

Adding authority Personal Name

0 1 2 3 4 5 6 7 8

000 - LEADER

▲ 00 fixed length control field *

001 - CONTROL NUMBER

003 - CONTROL NUMBER IDENTIFIER

▲ 00 control field *

005 - DATE AND TIME OF LATEST TRANSACTION

▲ 00 control field *

008 - FIXED-LENGTH DATA ELEMENTS

▲ 00 fixed length control field *

010 - LIBRARY OF CONGRESS CONTROL NUMBER

014 - LINK TO BIBLIOGRAPHIC RECORD FOR SERIAL OR MULTIPART ITEM

részlet egy besorolási rekord adatbeviteli űrlapjáról

A könyvtári besorolási állomány névtérnek tekinthető, hiszen ugyanazokat a funkciókat valósítja meg, mint amiket a névtértől elvárunk, még akkor is, ha gyakran más eszközöket, más megoldásokat használ a célok eléréséhez.

2.9.2.8 ISADN

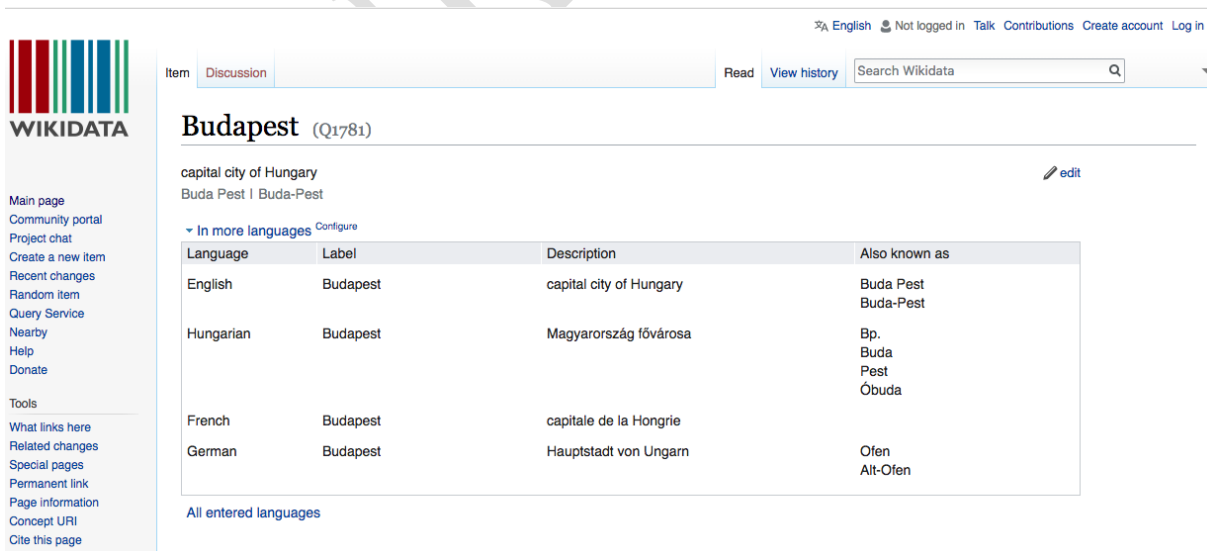
Az 1989-ben elindult projektben az ISADN (International Standard Authority Data Number) céljaként a személyek egyértelmű azonosítását lehetővé tevő –számszerű – azonosító rendszer felállítását és működtetését tűzték ki. A projekt végül nem valósult meg, de mintául szolgált a később sikeressé váló globális személyazonosító rendszerek, mint a VIAF vagy az ISNI számára (Habibzadeh & Yadollahie 2009).

2.9.2.9 Wikidata

A Wikipédia elvileg felkínálhatná az egyértelmű azonosítás lehetőségét kiválasztott entitástípus (mondjuk a személyek) számára, hiszen a webcímek egyedisége garantálni tudná a személyek egyedi azonosítását. A Wikipédia gyakorlati működés módja azonban e lehetőség kihasználását nem teszi lehetővé. A wiki-bejegyzések címének szabad változtatgatása ugyanis teljesen ellentétes a névterekkel szemben támasztott elvárásokkal, amelyek szerint az azonosítónak (a címnek) minden pillanatban, az időben folytonosan egyedinek és perzisztensnek kell lennie.

A Wikipédia tartalmait „adatszerűsíteni” kívánó Wikidata projekt már alkalmasabbnak tűnik az egyedi azonosítói szerep betöltésére, hiszen a kiválasztott entitás (típus, individuum) számára perzisztens azonosítót adnak a projekt keretein belül, de egyelőre kérdéses, hogy az alkalmazott, választott informatikai megoldás alkalmas^{sá} teszi-e a Wikidata egész rendszerét arra, hogy névtérként is működjön.

A Wikidata célja elsősorban az, hogy RDF-tripletek formájában – géppel is értelmezhető – állításokat lehessen tenni adott entitásról, individuumról, de lehet azonosítórendszerként is értelmezni, hiszen minden wikidata-oldalon leírt dolognak egyedi azonosítója van.



The screenshot shows the Wikidata page for 'Budapest (Q1781)'. At the top, there are navigation links for 'English', 'Not logged in', 'Talk', 'Contributions', 'Create account', and 'Log in'. Below the Wikidata logo, there are links for 'Main page', 'Community portal', 'Project chat', 'Create a new item', 'Recent changes', 'Random item', 'Query Service', 'Nearby', 'Help', and 'Donate'. A 'Tools' section includes 'What links here', 'Related changes', 'Special pages', 'Permanent link', 'Page information', 'Concept URI', and 'Cite this page'. The main content area shows the item name 'Budapest (Q1781)' and its description 'capital city of Hungary'. Below this, there is a table of labels in various languages:

Language	Label	Description	Also known as
English	Budapest	capital city of Hungary	Buda Pest Buda-Pest
Hungarian	Budapest	Magyarország fővárosa	Bp. Buda Pest Óbuda
French	Budapest	capitale de la Hongrie	
German	Budapest	Hauptstadt von Ungarn	Ofen Alt-Ofen











Below the table, there is a link for 'All entered languages'.

Budapest Wikidata-bejegyzésének része (benn a Wikidata-azonosítóval: Q1781)

A Wikidata oldalaira RDF-állításokat lehet feltenni, amelyeket aztán – egy másik felületen – SparQL-lekérdezések segítségével értelmezni, hasznosítani lehet további célokra. Az állítások

szintaxisára, szemantikájára vannak belső szakmai útmutatók, de a rendszer nem ellenőrzi ezek betartását, ami bizonyos mértékig kétségeket ébreszt a rendszer konzisztenciáját illetően.

Statements

instance of	 city  edit
	▼ 0 references
	+ add reference
capital	 capital  edit
	▼ 0 references
	+ add reference
town status in Hungary	 town status in Hungary  edit
	► 1 reference
county of Hungary	 county of Hungary  edit
	▼ 0 references
	+ add reference
	+ add value
part of	 Central Hungary  edit
	► 1 reference
	+ add value

Budapest – RDF-állításokat tartalmazó – Wikidata-oldalának része

3 Névstruktúra

A névtér fogalmának alaposabb elemzésére – sok egyéb további szemponton túl – azért van szükség, hogy

- láthassuk azt, hogy a névterek miben térnek el más rendszerektől;
- meg tudjuk mondani, mi a névtér, és mi nem az, és mi alapján tudjuk ezt eldönteni;
- minél pontosabban el tudjuk különíteni egymástól a különböző névtereket, névtértípusokat, tudjuk, hogy mi az azonos, és mi az eltérő bennük;
- tisztázhassuk, hogy mi kapcsolat a köznévtér, a tulajdonnévtér, valamint a tudásszervezési rendszer fogalmai között.

Ezen kérdések megválaszolásához első lépésben arra van szükségünk, hogy röviden felvázoljunk egy relációelméletet, majd az itt definiált relációk segítségével pontosan meghatározzuk a névterekkel, tudásszervezési rendszerek, ezek típusaival kapcsolatos fogalmakat.

3.1 Relációelmélet

A névterek, a tudásszervezési rendszerek (KOS-ok) matematikai értelemben olyan struktúráknak tekinthetők, amelyek a névtérbe sorolt elemek közt adott relációkat definiálnak. A névterek egyfelől a névhalmaz elemeinek típusában, másfelől az elemek közti relációk típusaiban térnek el egymástól. A köznévterek vizsgálatakor, jellemzésekor ezért fontos lehet, hogy milyen relációról beszélhetünk. A tudásszervezési rendszerek (főleg az osztályozási rendszerek és teauruszok) építésének több évszázados történetében nem túl nagy mennyiségű relációtípust kezeltek, ami nyilván nem véletlen, hiszen az emberi tudás terminológiai, fogalmi rendszerbe foglalása nem elsősorban a feldolgozni, leírni kívánt tudásterület sajátosságaitól, hanem az emberi tudástermelés, az emberi gondolkodás általános szabályszerűségeitől függ.

A következőkben egy olyan rövid és formális relációelméletet vázolunk fel, amelyben a névterek építkezésekor szükséges relációk mind benne lesznek, és amely – szükség esetén – szabadon bővíthető. A fogalmak formális meghatározása során használjuk a logikában kiterjedt módon alkalmazott szimbólumokat a logikai alapfogalmakra, műveletekre.

3.1.1 Relációelméleti alapfogalmak

Az itt bemutatott relációelméletet a legalapvetőbb relációs fogalmak (reláció, relációtulajdonságok, elemi relációk, függvény, művelet, nevezetes matematikai relációk, relációművelet) definiálásával kezdjük, majd összegyűjtjük az ontológiai alapfogalmakat. Ezek után és ezek alapján tudjuk meghatározni azokat a KOS-relációkat, amelyeket az egyes tudásszervezési rendszerek, névterek jellemzéséhez, leírásához lehet felhasználni.

3.1.1.1 A reláció fogalma

Mind a reláció fogalmát, mind a hozzá kapcsolódó két fontos fogalom, az értelmezési tartomány és értékészlet fogalmát halmazelméleti alapon határozzák meg. Ezeket a következő módon definiálhatjuk.

reláció

relation(x, y) $\equiv R(x, y)$	$\{\{x\}, \{x, y\}\}$
domain(R)	$\{y: \exists x(R(y, x))\}$
image(R)	$\{y: \exists x(R(x, y))\}$

A relációfogalomra támaszkodva meghatározhatjuk a legegyszerűbb relációtípusokat is. Ezek az univerzális vagy teljes reláció, az azonosság, identitás vagy diagonális reláció, az üres vagy null reláció, valamint a különbözőségi, egyenlőtlenségi, antidiagonális relációk.

elemi relációk

universal_relation(x,y)	$\forall m \forall n (R(x, y))$
identity_relation(x,y)	$\forall m \forall n (R(x, y) \leftrightarrow x = y)$
empty_realtion(x,y)	$\forall m \forall n (\neg R(x, y))$
diversity_relation(x,y)	$\forall m \forall n (R(x, y) \leftrightarrow x \neq y)$

3.1.1.2 Relációtulajdonságok

A relációk szemantikájának kibontása előtt először azt kell felidézni, hogy a különböző relációfogalmak milyen formális, szintaktikai tulajdonságokkal rendelkezik(nek). A névterek számára hasznos relációk jellemzéséhez a matematikában használt relációtulajdonságra tudunk támaszkodni.

relációtulajdonságok

reflexív	$\forall x (R(x, x))$
irreflexív	$\forall x (\neg R(x, x))$
nem-reflexív	$\exists x (\neg R(x, x))$
balról totális	$\forall x \exists y (R(x, y))$
jobbról totális	$\forall y \exists x (R(x, y))$
másodlagosan reflexív	$\forall x \forall y (R(x, y) \rightarrow R(y, y))$

fordított másodlagosan reflexív	$\forall x \forall y (R(x, y) \rightarrow R(x, x))$
koreflexív	$\forall x \forall y (R(x, y) \rightarrow x = y)$
szimmetrikus	$\forall x \forall y (R(x, y) \rightarrow R(y, x))$
aszimmetrikus	$\forall x \forall y (R(x, y) \rightarrow \neg R(y, x))$
antiszimmetrikus	$\forall x \forall y ((R(x, y) \wedge R(y, x)) \rightarrow x = y)$
nem-szimmetrikus	$\exists x \exists y (R(x, y) \wedge \neg R(y, x))$
összefüggő	$\forall x \forall y (x \neq y \rightarrow (R(x, y) \vee R(y, x)))$
erősen összefüggő	$\forall x \forall y (R(x, y) \vee R(y, x))$
gyengén összefüggő	$\forall x \forall y \forall z (R(x, y) \wedge R(x, z) \rightarrow (y = z \vee R(y, z) \vee R(z, y)))$
dichotóm	$\forall x \forall y (x \neq y \rightarrow (R(x, y) \wedge \neg R(y, x)) \vee (\neg R(x, y) \wedge R(y, x)))$
trichotóm	$\forall x \forall y (R(x, y) \wedge \neg R(y, x) \wedge \neg \exists z (R(x, z) \wedge R(z, y)) \vee (\neg R(x, y) \wedge R(y, x) \wedge \neg \exists z (R(x, z) \wedge R(z, y)) \vee x = y)$
balról egyértelmű	$\forall x \forall y \forall z ((R(x, z) \wedge R(y, z)) \rightarrow x = y)$
jobbról egyértelmű	$\forall x \forall y \forall z ((R(x, y) \wedge R(x, z)) \rightarrow y = z)$
transzítív	$\forall x \forall y \forall z ((R(x, y) \wedge R(y, z)) \rightarrow R(x, z))$
intranszítív	$\forall x \forall y \forall z ((R(x, y) \wedge R(y, z)) \rightarrow \neg R(x, z))$
kvázitranszítív	$\forall x \forall y \forall z ((R(x, y) \wedge R(y, z) \wedge x \neq z) \rightarrow R(x, z))$
aliotranszítív	$\forall x \forall y \forall z ((R(x, y) \wedge R(y, z)) \rightarrow (R(x, z) \vee x = z))$
nem-transzítív	$\exists x \exists y \exists z ((R(x, y) \wedge R(y, z)) \wedge \neg R(x, z))$
cirkuláris	$\forall x \forall y \forall z ((R(x, y) \wedge R(y, z)) \rightarrow R(z, x))$
acirkuláris	$\forall x \forall y \forall z ((R(x, y) \wedge R(y, z)) \rightarrow \neg R(z, x))$
euklideszi	$\forall x \forall y \forall z ((R(x, y) \wedge R(x, z)) \rightarrow R(y, z))$
aneuklideszi	$\forall x \forall y \forall z ((R(x, y) \wedge R(x, z)) \rightarrow (\neg R(y, z) \wedge \neg R(z, y)))$
katenáris	$\forall x \forall y \exists z (R(x, y) \rightarrow ((R(y, z) \vee R(z, x)))$
akatenáris	$\forall x \forall y \forall z (R(x, y) \rightarrow (\neg R(y, z) \wedge \neg R(z, x)))$
konfluens	$\forall x \forall y \forall w \exists z ((R(x, y) \wedge R(x, w)) \rightarrow (R(y, z) \wedge R(w, z)))$
akonfluens	$\forall x \forall y \forall w \exists z ((R(x, y) \wedge R(x, w)) \rightarrow \neg (R(y, z) \wedge R(w, z)))$

3.1.1.3 Függvény

A reláció kategóriája segítségével lehet definiálni a függvény vagy leképezés fogalmát, majd erre támaszkodva lehet meghatározni további kategóriákat, mint az inverz függvény vagy inverz leképezés, az egy-egyértelmű, a bitotális függvény vagy 1:1-es reláció, az injektív vagy egyrétű függvény, a szürjektív függvény vagy ráképezés vagy szuperjektív vagy lefedő, a bijektív vagy egyrétűen lefedő függvény, a sorozat függvény, a monoton vagy monoton növekvő függvény, illetve az antiton vagy monoton csökkenő függvény.

függvénytulajdonságok

function $\equiv F$	left-total(F) \wedge right-unique(F)
inverse_function(m, n) $\equiv F^{-1}$	right-total(F) \wedge left-unique(F)
biunique(F)	left-unique(F) \wedge right-unique(F)
bitotal(F)	left-total(F) \wedge right-total(F)
injective(F)	function(F) \wedge left-unique(F)
surjective(F)	function(F) \wedge right-total(F)
bijjective(F)	injective(F) \wedge surjective(F)
series(F)	function(F) \wedge domain(F) = N
monotone(F(x, y))	$\forall x \forall y (x \leq y \rightarrow F(x) \leq F(y))$
antitone(F(x, y))	$\forall x \forall y (x \leq y \rightarrow F(y) \leq F(x))$

3.1.1.4 Művelet

A művelet (operáció vagy belső összekapcsolás) fogalma a függvényből származtatható. A művelet meghatározása után sorban lehet definálni az egyes művelettulajdonságokat, mint az asszociatív vagy társítható vagy csoportosítható művelet, a kommutatív vagy felcserélhető művelet, O disztributív Q-ra vagy szétosztható vagy széttagolható művelet, Q disztributív O-ra vagy szétosztható vagy széttagolható művelet, idempotens vagy önhatványozó vagy azonos hatványú művelet, O abszorbtív Q-ra vagy elnyelő vagy adjunktív művelet, Q abszorbtív O-ra vagy elnyelő vagy adjunktív művelet.

művelet és művelettulajdonságok

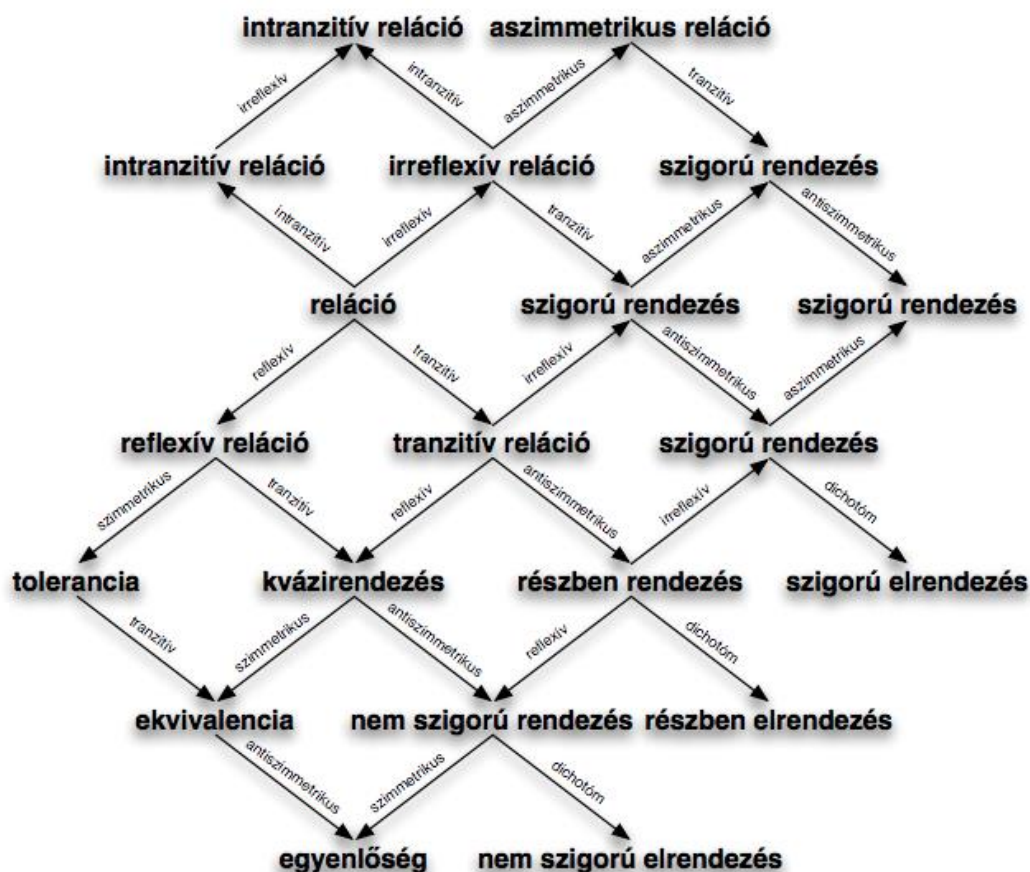
operation(x,y) $\equiv O(x, y)$	function(x, y) \wedge domain = cod
associative(O(x, y))	$\forall x \forall y \forall z (xO(yOz) = (xOy)Oz)$
commutative(O(x, y))	$\forall x \forall y (xOy = yOx)$
distributive(O(x, y) Q)	$\forall x \forall y \forall z ((xO(yQz) = (xOy)Q(xOz))$
distributive(Q(x, y) O)	$\forall x \forall y \forall z ((yQz)Ox = (yOx)Q(zOx))$
idempotent(O(x, y))	$\forall x (xOx = x)$
absorptive(O(x, y) Q)	$\forall x \forall y (xQ(xOy) = x)$
absorptive(Q(x, y) O)	$\forall x \forall y (yO(xQy) = x)$

3.1.1.5 Nevezetes matematikai relációk

A matematikában gyakran használt relációknak saját nevük van, és ezeket a relációtulajdonságok alapján szokás meghatározni. A leggyakrabban használt, nevezetes relációk között van a tolerancia vagy kompatibilitási vagy hasonlósági reláció, a részleges ekvivalencia reláció, az ekvivalencia reláció, a részben vagy parciális vagy részleges vagy gyenge vagy gyenge részben rendezés vagy nem szigorú rendezés, a gyenge rendezés, a lineáris rendezés, a szigorú vagy erős részben vagy irreflexív rendezés, a nem szigorú elrendezés reláció vagy teljes parciális vagy lineáris rendezés, a parciális vagy totális parciális rendezés, a szigorú elrendezés vagy teljes szigorú vagy lineáris szigorú vagy totális szigorú rendezés, az előrendezés vagy kvázirendezés, a gyenge rendezés vagy lineáris előrendezés, a gyenge részben rendezés, a részben elrendezés vagy teljes gyenge részben rendezés, valamint a tournament reláció.

<i>nevezetes matematikai relációk</i>	
tolerance	reflexive \wedge symmetric
partial_equivalence	transitive \wedge symmetric
equivalence	reflexive \wedge symmetric \wedge transitive
partial_order	reflexive \wedge antisymmetric \wedge transitive
weal_order	complete \wedge transitive
linear_order	complete \wedge antisymmetric \wedge transitive
strict_order	irreflexive \wedge transitive
linear_partial_order	reflexive \wedge antisymmetric \wedge transitive \wedge dichotomous
linear_strict_order	irreflexive \wedge transitive \wedge dichotomous
preorder	reflexive \wedge transitive
weak_order	reflexive \wedge transitive \wedge dichotomous
weak_partial_order	antisymmetric \wedge transitive
linear_weak_partial_order	antisymmetric \wedge transitive \wedge dichotomous
tournament	transitive \wedge dichotomous

Az eddig bemutatott relációk közül sokat (bár nem az összeset) kiválasztva a relációtulajdonságok alapján felrajzolhatunk egy ábrát, amiben látszanak a relációk közti kapcsolatok is.



nevezetes relációk közti kapcsolatok a relációtulajdonságok alapján

3.1.1.6 Relációműveletek

A relációkkal is végezhetünk műveleteket, és ezek közül a legfontosabbak saját megnevezéseket kaptak. Itt bemutatjuk a relációkonvertálás vagy konverzképzés vagy inverzképzés vagy megfordítás, a relációinvertálás vagy inverzképzés, a kontrapozíció vagy transzpozíció, a relációkomplementálás, a duálreláció vagy duális reláció, a relációmetszés, a relációegyesítés vagy relációúnió, a relációszorzás vagy relatív szorzat vagy kompozíció vagy multiplikáció vagy szuperpozíció vagy relációkompozíció, valamint a relációösszeadás vagy összeadás vagy relatív összeadás műveletét.

relációműveletek

R-conversion $\equiv R^{-1}(x, y)$	$\forall x \forall y (R^{-1}(x, y) \leftrightarrow R(y, x))$
R-inversion $\equiv R^{-i}(x, y)$	$\forall x \forall y (R^{-i}(x, y) \leftrightarrow R(\neg x, \neg y))$
R-contraposition $\equiv R^{-p}(x, y)$	$\forall x \forall y (R^{-p}(x, y) \leftrightarrow R(\neg y, \neg x))$

R-complementation $\equiv R^{-t}(x, y)$	$\forall x \forall y (R(x, y) \leftrightarrow \neg R(x, y))$
R-dual $\equiv R^{-d}(x, y)$	$\forall x \forall y (R^{-d}(x, y) \leftrightarrow \neg R(y, x) \equiv R^{-c}(x, y))$
R-intersection $\equiv \times(R(x, y), Q(x, y))$	$\forall x \forall y (R \times Q(x, y) \leftrightarrow (R(x, y) \wedge Q(x, y)))$
R-union $\equiv + (R(x, y), Q(x, y))$	$\forall x \forall y (R + Q(x, y) \leftrightarrow (R(x, y) \vee Q(x, y)))$
R-composition $\equiv * (R(x, y), Q(x, y))$	$\forall x \forall y \forall z (R * Q(x, y) \leftrightarrow (R(x, z) \wedge Q(z, y)))$
R-addition $\equiv \dagger (R(x, y))$	$\forall x \forall y \forall z (R \dagger Q(x, y) \leftrightarrow (R(x, z) \vee Q(z, y)))$

3.1.1.7 Reifikáció

A relációk használata során gyakran előfordul, hogy azokat módosított értelemben használjuk, mégpedig úgy, hogy nem vesszük figyelembe a reláció teljes jelentését, csak annak egy részét. Példaként az 'x apja y-nak' kétargumentumú relációra hivatkozhatunk, amelynek reifikált alakjából elhagyjuk az egyik argumentumot és az eredeti relációs minőséget belevetítjük a reláció meghagyott argumentumát adó entitásba: 'x apa'. Ezt műveletet nevezik reifikációnak, aminek van két típusa, a bal-, illetve jobb-reifikáció.

a reifikáció művelete

left-reification $(x, y) \equiv 'R(x, y)$	$\forall x \forall y ('R(x, y) \leftrightarrow R(x))$
right-reification $(x, y) \equiv R' (x, y)$	$\forall x \forall y (R' (x, y) \leftrightarrow R(y))$

3.1.2 Ontológiai alaprelációk

Ebben a fejezetben azokat az ontológiai alapfogalmakat vesszük számba, amelyeket minden adatmodell felépítése során használni kell, és amelyek között többféle kapcsolatot is értelmezhetünk. Azért fontos ez a kitekintés, mert ezeket a kapcsolatokat nem nehéz összekeverni, ami aztán könnyen pongyolasághoz, többértelműséghez vezet.

A filozófiai, ontológiai szakirodalomban régóta használt, széles körben elterjedt a partikularé-univerzálé fogalomkettős, bár a két terminus helyett – különösen a filozófián túli tudásterületeken – gyakran használnak más kifejezéseket. Amíg partikularé, illetve az individuum, az instancia, az előfordulás, az elem, a példány terminusaival mindig a világ valamilyen egyedi, minden mástól elkülöníthető dologára hivatkozunk, addig a univerzálé, illetve a concept, a property, az osztály, a típus terminusaival mindig valamely jellemző mentén közös dolgok csoportját, gyűjteményét jelöljük. Nem szokás azonban használni a gyűjtemény vagy a halmaz terminusát az univerzálé helyett, és ez a mozzanat felveti annak szükségességét, hogy a partikularé-univerzálé kettős mellé felvegyük a gyűjtemény fogalmát is, és persze tisztázzuk a három ontológiai alapfogalom egymáshoz való viszonyát (Bittner et al. 2004).

Bár az elem és az individuum halmazelméletéhez közelebb álló terminusok, belül nem teszünk különbséget a partikularé szinonimái között (bár azért leggyakrabban a partikularé és az individuum terminusokat fogjuk használni). Helytelennek tartjuk viszont a concept, az osztály vagy a property terminusok használatát az univerzálé fogalma helyett, ezért az univerzálé szinonimájaként csak a típus terminusát tartjuk elfogadhatónak. A kérdés tehát az, hogy mi a különbség az individuum, az univerzálé és a gyűjtemény fogalmai között.

Mivel a világban létező dolgok leírására alkalmas fogalmainkat akarjuk valahogyan tipizálni, ezért legelőször a partikularé (individuum) és univerzálé fogalmait kell meghatározni. Ezt egy speciális metarelációval, az instanciálás (példányosítás) relációval tehetjük meg. Ezt a definíciót használják a jelenkori filozófusok (Lowe 1998), de ezt az értelmezést veszi át a Dolce ontológia is (Gangemi et al. 2001). Az instanciálás metarelációval azt fejezhetjük ki, hogy egy fogalommal leírt „dolognak” mennyi példánya lehetséges a világban. Ha a fogalommal leírt dolog csak „egymagában” létezik, tehát nincsenek példányai, legfeljebb csak annyit mondhatunk, hogy „önmaga példánya”, akkor beszélünk a partikularé fogalmáról.

Ezzel szemben az univerzálé fogalmával olyan dolgot illetünk, amelynek lehetséges több példánya is. Mondhatjuk azt is, hogy a partikularé az instanciálás metareláció jobbról egyértelmű szűkítése, míg az univerzálé e szűkítés komplementere. A három eddig tárgyalt fogalmat Gangemiék alapján a következő formulákkal írhatjuk le (Gangemi et al. 2001).

Először az instanciálás vagy instanciája, előfordulása, példánya relációt definiáljuk.

instanciálás reláció

instance = ϵ	asymmetric \wedge aneuclidean
-----------------------	---------------------------------

Az instanciálás reláció segítségével már meghatározhatjuk a partikularé vagy individuum, illetve az univerzálé fogalmait is.

partikularé és univerzálé

particular	$\forall x(\text{particular}(x) \leftrightarrow \neg \exists y(\text{instance}(y, x)))$
universal	$\forall x(\text{universal}(x) \leftrightarrow \neg \text{particular}(x))$

Van tehát három alapvető metafogalmunk, köztük az instanciálás reláció, de amikor továbblépünk, figyelni kell arra, hogy nem szabad összekevernünk az instanciálás metarelációt azzal a relációval, amit a partikularé és az univerzálé fogalmai között értelmezhetünk és aminek az instanciája nevet adhatjuk. Bár az instanciája reláció nagyon hasonlít az eleme relációhoz, mégsem lehet azonos vele, mivel másfajta fogalmak között értelmezzük. Amíg az eleme reláció egy individuum (partikularé) és egy gyűjtemény (halmaz) közötti reláció, addig az instanciája reláció egy partikularé és egy univerzálé között áll fent (Bittner et al. 2004, Gangemi et al. 2001). A meglévő különbségek miatt tehát egymástól különböző relációnak kell tartanunk őket, de a meglévő hasonlóságok miatt az eleme reláció jeléhez (\in) hasonló szimbólummal jelöljük az instanciája metarelációt (ϵ). Az instanciája relációban van valamilyen intenzionális kapcsolat,

ami teljesen hiányzik a halmaz és elemei között fenntartott eleme relációból. Ezért van szükség az univerzálé fogalma mellé felvenni a gyűjtemény kategóriáját, hogy az individuumok és a gyűjtemények közti extenzionális kapcsolatot kifejezhessük az eleme reláció alkalmazásával.

A három fogalomtípus eltérő minőségét szemléltethetjük az időbeli „viselkedésükkel”. Az individuumok elveszthetik részeit vagy épp új részeket „szerezhetnek” (az ember sejtszintjén folyamatosan elhalnak, illetve megszületnek). Az univerzálék is hasonlóképpen viselkednek, csak nem a részeit, hanem a példányaikat veszíthetik el vagy nyerhetnek újakat belőlük. A gyűjtemény (a halmaz) viszont mindig az elemei, összetevői (vagyis az individuumok) által azonosítható, ezért nem lehetséges, hogy különböző időpontokban különböző elemei legyenek.

Van tehát három ontológiai metaegységünk, melyek között különböző relációkat értelmezhetünk. A következő táblázatban először felsoroljuk azokat a kapcsolatokat (1-8. sor), amelyek a három entitástípus között léteznek a partitivitás atommentes (részhalmaza-szerű) és atomos (eleme-szerű) relációtípusa, illetve a genericitás szerint. Ezt kiegészítjük a feltüntetett relációk fontosabb inverzeivel (1-10, 4-9, 7-11, 5-12).

metaszintű keresztrelációk

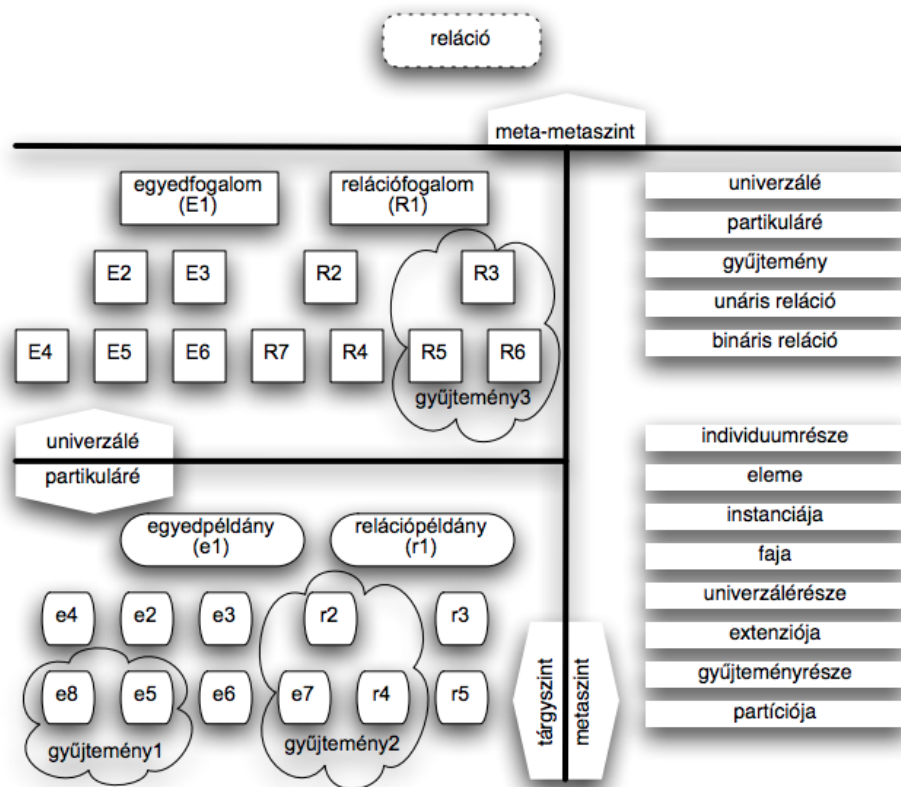
	reláció jele	bal argumentum	jobb argumentum
1	individuumrésze	individuum	individuum
2	eleme	individuum	gyűjtemény
3	instanciája	individuum	univerzálé
4	faja	univerzálé	univerzálé
5	univerzálérésze	univerzálé	univerzálé
6	extenziója	gyűjtemény	univerzálé
7	gyűjteményrésze	gyűjtemény	gyűjtemény
8	partíciója	gyűjtemény	individuum
9	neme	univerzálé	univerzálé
10	individuumegésze	individuum	individuum
11	gyűjteményegésze	gyűjtemény	gyűjtemény
12	univerzáléegésze	univerzálé	univerzálé

A metaszintű keresztrelációkat önálló szimbólumokkal jelöljük, és a relációk argumentumait infix-módban adjuk meg (tehát a két argumentum közé tesszük a relációjelet). A fenti relációk – legalább az időhöz való viszonyukban – különböznek egymástól, és különböző erősségű formulákkal írhatók le (ezeket itt nem mutatjuk be, részletesen kibontva lásd: Bittner et al. 2004).

A táblázat az ontológiák nagyon fontos metareláció-típusait mutatja be, de jeleznünk kell azt is, hogy egy nagyon fontos reláció hiányzik a felsorolásból. Az intenzió kategóriája – az extenzió mellett – a fogalmak nélkülözhetetlen „tartozéka”, amit – ennek ellenére – a kortárs ontológiák és ontológiaelméletek egyike sem tud (vagy akar) megragadni. Ez persze nem véletlen, hiszen az

intenzió, vagyis a fogalmak jelentését, tartalmát kifejező bármilyen állítás, tehát „tetszőleges” logikai formula lehet, amit eddig még nem sikerült egyetlen metafogalomtípushoz rendelni. Mi sem tudjuk formalizálni és a modellbe építeni ezt a fogalmat, de fontosnak tartjuk, hogy egyrészt azért jelezzük a hiányát, másrészt rögzítsük a lehetséges argumentumait: az ‘intenziója’ reláció egy formula és egy univerzálé között teremt kapcsolatot.

Az univerzálé és a partikuláré (illetve a gyűjtemény) kategóriái mellett – még mindig a metaszinten maradva – szükségünk van egy másik fontos fogalmi megkülönböztetésre is. A matematikai szinten definiált legáltalánosabb metafogalmunk, a reláció fogalmához a meta-metaszinten hozzákapcsoltuk az extenzionális relációelméletet, melyből két relációtípust „lehozhatunk” a metaszintre és deklarálnak az unáris reláció és a bináris reláció metafogalmait (meg kell jegyezzük, hogy ez a dichotómia még mindig szintaktikai meghatározáson alapul, tehát ezek miatt sem kell semmiféle plusz ontológiai elkötelezettséget vállalnunk). Az unáris és a bináris reláció két fogalomtípusát messzemenően hasznosíthatjuk a tárgyszintű építkezésben, hiszen a bináris reláció metafogalma adja a tárgyszint relációfogalmának típusát, az unáris reláció metafogalmat pedig a tárgyszinten az egyedfogalomnak feleltethetjük meg (mint a tárgyontológia két legfelső kategóriáját). Ha az eddig bemutatott dimenziókat egymásra vetítjük, akkor meghatározhatjuk az ontológia tárgyszintjének legfontosabb metafogalmait. Az univerzálé-partikuláré kettősség mentén választhatjuk el egymástól a modell generikus, illetve instanciális szintjét, míg az egyed- és relációfogalmak (vagyis az unáris és a bináris reláció metafogalmak) segítségével megadhatjuk a tárgyszint generikus és instanciális hierarchiájának két-két csúcselemét. A generikus szinten tehát lehetnek egyedfogalmaink és relációfogalmaink, az instancia szinten pedig előfordulhatnak egyedpéldányok és relációpéldányok, és természetesen a gyűjtemények mindkét szinten összefoghatnak univerzálékat és partikulárékat egyaránt. A négy ontológiai réteg, a metafogalmak és a köztük levő metarelációk, valamint a tárgyszint legfontosabb összetevőit mutatjuk be a következő ábrán.



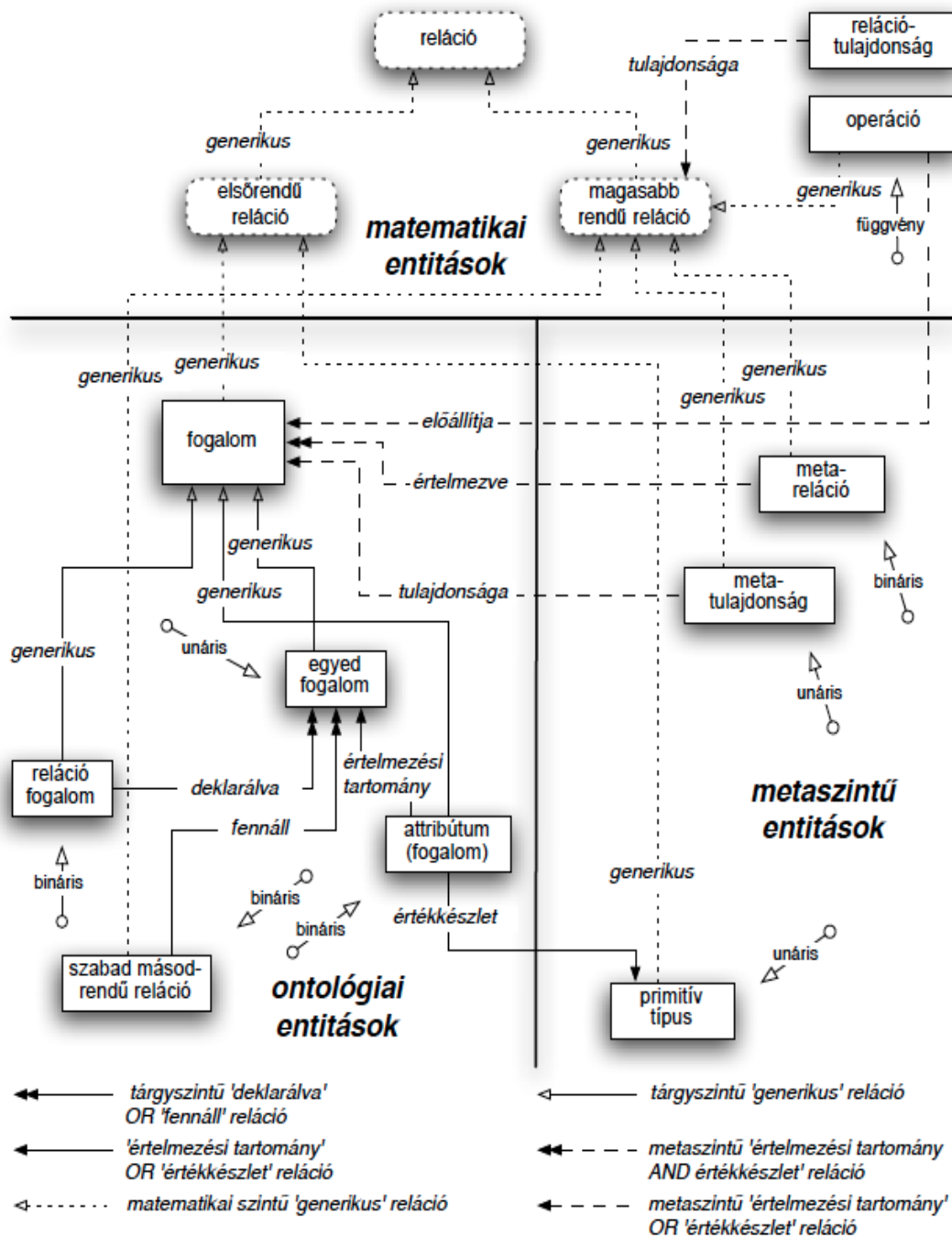
alapfogalmak egymáshoz való viszonya

A keresztrelációk elméletében kellett még definiálni az olyan metafogalmakat, mint az összekapcsolódás (connection), elhelyezkedés (location) vagy a függőség (dependence), melyek közül az utóbbi azért különösen fontos, mert a fogalmak tartalmát jelentő ismertetőjegyek, tulajdonságok formalizálhatóságának ígéretét jelenti.

Az eddig említett elméletek az ontológiatervezési munkával kapcsolatosak. A tényleges ontológiaépítő munka a tárgyszintű építkezést jelenti, amikor az ontológia szerkesztői – saját ontológiai elfogultságaik, elkötelezettségeik alapján – új és új fogalmakat vesznek fel, melyeket ismertetőjegyekkel jellemeznek, illetve meghatározzák az új fogalom helyét a fogalmi struktúrában.

A tárgyszinten a generikus modellnek megfelelően vehetjük fel az univerzálékat, melyekre a metaszinteken előzetesen meghatározott metafogalmak (metatulajdonságok, relációs tulajdonságok, műveletek, generikus alárendeltje reláció stb.), kategóriák (absztrakt, konkrét, egyedfogalom, relációfogalom stb.).

Elsősorban a tárgyszintű ontológiaépítés lehetőségeit szem előtt tartva az alábbi ábrán mutatjuk be a generikus modell legfontosabb fogalmait.



ontológiai szintek és alapfogalmak

3.1.3 KOS-relációk

A tudásszerkezési rendszerek, a névterek építése során sokfajta relációra van szükség – igaz, a különböző tudásszerkezési rendszerek, valamint a kétféle névtér, a tulajdonnévtér és a köznévtér számára más relációk a fontosak. Vannak relációk, amelyek mindegyik névtér esetében kellenek, de vannak olyanok, amelyeket csak a köznévtér vagy tudásszerkezési

rendszerek építése során vehetünk használatba. Ebben a fejezetben ezeket vesszük sorba. Elemzésünket az alábbi relációkra terjesztjük ki.

- faja/neme vagy generikus (föle- és alárendeltje) reláció
- része/egése vagy partitív (föle- és alárendeltje) reláció
- homonima reláció
- szinonima reláció
- meghatározója (előzménye, következménye) reláció
- hierarchikus alá-fölelendeltje (tartalmazás) reláció
- lexikografikus reláció
- rokona (egyéb) reláció

3.1.3.1 Generikus reláció

A tudásszervezési rendszerek építésében (vagy általánosabban: az emberi tudás fogalmakkal való reprezentációjában) a legfontosabb reláció a generikus reláció. Az egyértelműség kedvéért érdemes rögzíteni, hogy itt nem egyetlen relációról, hanem inkább egy relációkettősről kell beszélnünk. A generikus reláció fogalmát általában a generikus alárendeltje reláció értelemben szokták használni, de nem kevés alkalommal szükség lehet a reláció inverzére, a generikus fölelendeltje relációra is. Ha nem zavaró, akkor a generikus reláció kifejezéssel hivatkozhatunk a generikus alá- és fölelendelt relációra együttesen.

A generikus reláció a relációk (predikátumok, univerzálék) közt teremt kapcsolatot, ezért másodrendű relációnak kell tekintenünk. A relációt definiáló formulával azt fejezzük ki, hogy ha veszünk egy általánosabb (Q) és egy speciálisabb fogalmat (P), akkor az utóbbi bármely példányára igaz lesz, hogy az előbbi példány is, miközben ez fordítva nem igaz.

generikus reláció

generic_subordinate(P,Q)	$\forall x \forall P \forall Q ((P(x) \rightarrow Q(x)) \wedge \neg(Q(x) \rightarrow P(x)))$
generic_superordinate(P,Q)	$\forall x \forall P \forall Q ((Q(x) \rightarrow P(x)) \wedge \neg(P(x) \rightarrow Q(x)))$

Más kontextusokban ugyanerre a fogalomra használják az 'osztályozási reláció' kifejezést is, sőt, a 'típusa' relációt is gyakran ebben az értelemben alkalmazzák. A keresztrelációk elméletében korábban a 'neme' és 'faja' kifejezéseket alkalmaztuk erre a két fogalomra. Előfordul még az inklúzió mint magába foglalás, a szubszumpció mint tartalmazás terminusa is, amelyek inkább a generikus fölelendelés értelmét fejezik ki. Olykor felbukkan a hipo- és hipernima kifejezéspár is, aminek jelentése a generikussággal azonos. Az 'osztály-alosztály', illetve a 'típus-altípus' besorolások a generikusság itt bemutatott értelmét fejezik ki, ezért ezek a terminusok tekinthetők egymás szinonimáinak, de vigyázni kell akkor, amikor a – szintén széles körben elterjedt – IS-A relációt akarnánk használni a generikus reláció szinonimájaként, mert ezzel a relációval lehetnek problémák. A generikus reláció nagyon fontos tulajdonsága a tranzititás, és ez a minőség nem minden esetben áll fenn – a maga teljességében – az IS-A reláció esetében. Ennek az az oka, hogy az IS-A reláció használata során előfordul, hogy azt kiterjesztik egyfelől az

'IS-A-TYPE-OF' (vagy az 'IS-A-KIND-OF', illetve 'IS-A-SORT-OF') relációra, másfelől az 'IS-AN-INSTANCE-OF' relációra egyaránt, és ez – olykor – nem várt következményekhez vezethet. Vegyünk egy példát! A következő három mondat mindegyike igaz lehet.

állat IS-A élőlény

puli IS-A állat

Bodri IS-A puli

Az első két esetben a generikus relációval kötünk össze kategóriákat, ám a harmadikban már a példány relációt használjuk, ami furcsa változást hoz az 'is-a' relációval összekapcsolt entitások esetében. Nem pontos azt mondani, hogy a tranzitivitás minősége sérül, inkább egy furcsa változás áll be a tranzitivitási sorban. A generikus ('IS-A-TYPE-OF') reláció esetében elméletileg mindig van lehetőség arra, hogy a generikus relációval összekötött kategóriák közé be lehessen tenni egy újat úgy, hogy a tranzitivitás megmarad. A fenti példát kiegészítve mondhatjuk még a következőt.

puli IS-A kutya

kutya IS-A állat

Ekkor a puli és az állat fogalmi közé betettünk egy újat (a kutya fogalmát), ami által létrehoztunk egy négy lépcsőből álló tranzitív sort.

puli IS-A kutya IS-A állat IS-A élőlény

Amikor a generikusként értelmezhető IS-A relációval összekapcsolt elemek közé beteszünk egy instanciájaként felfogható IS-A relációt, akkor a fent jelzett alábonthatóság lehetősége már nem áll fent. Ha azt mondjuk, hogy:

Bodri IS-A puli

akkor a 'Bodri' alá már nem vehetünk fel semmi mást az IS-A reláció segítségével. Ez nem borítja fel a tranzitivitást, de balról lezárja új elemek felvételének lehetőségét a generikus reláció segítségével. Az IS-A reláció jobb oldalán sosem szerepelhet a 'Bodri' elem. Ezzel szemben a valódi generikus struktúra esetén bármelyik elem szerepelhet a kapcsolat mindkét oldalán. Leszámítva persze a generikus hierarchia csúcspontját álló 'valami' fogalmat.

A névtérépítés során az egyik legfontosabb feladat annak biztosítása, hogy a generikus relációt és a példány relációt mindig elkülönítsük egymástól, mert az összevont kezelésük könnyen zavart okozhat. A példány és a generikus relációk különbsége segít majd nekünk később, a névterek definiálásakor abban, a névterek különböző típusait egyértelműen és pontosan elhatárolhassuk egymástól.

A generikus reláció kapcsán fontos figyelembe venni azt a megkülönböztetést, miszerint a hierarchikus relációnak két fajtája is van: a **polihierarchikus** és a **monohierarchikus** reláció. A monohierarchikus struktúrán belül minden elemnek egyetlen közvetlen felettese van, míg a polihierarchiára ez nem igaz, vagyis itt megengedett a több szülő létezése. A tranzitivitást nem változtatja meg ez a fogalmi felbontás, de a struktúra elemei közti kapcsolatok minősége már más a két rendszeren belül. A monohierarchia fagráfokkal reprezentálható (aminek legfőbb jellemzője „körmentesség”, vagyis az, hogy két pont között mindig csak egyetlen útvonal létezik), míg a polihierarchiára csak annyit mondhatunk, hogy azok tranzitív, irányított gráfok, amiben két elem között többféle útvonal is létezhet.

3.1.3.2 Partitív reláció

A rész/egész relációval vagy másként a partitív alárendeltje/partitív fölérendeltje relációval a mereológia elmélete foglalkozik. A része (partitív alárendeltje) reláció az egésze (partitív fölérendeltje) inverze, és ez relációkettős is tranzitív, hierarchikus – akárcsak generikus reláció. A tranzitivitás azonban csak a pontosan definiált – és egymástól is jól elkülönített – altípusokra érvényes, nem általában a partitív relációra. Sok esetben érvényesül a tranzitivitás. Vegyünk két példát!

Ha Érd Pest megye része, és Pest megye Magyarország része, akkor Érd is Magyarország része.

Ha az ujjam a kezem része, és a kezem a karom része, akkor az ujjam a karom része is.

Könnyen lehet azonban olyan példát találni, amikor a partitív reláció intranzitivitást tapasztalhatunk.

A zongorista keze része a zongoristának, a zongorista része a zenekarnak, de a zongorista keze nem része a zenekarnak.

Az ellentmondásosnak tűnő helyzet magyarázata egyszerű: többféle partitív reláció létezik, és ezek között nem érvényes a tranzitivitás. Olyan mereológiai elméletre van ezért szükség, amely kezelni, magyarázni tudja ezt a jelenséget. A versengő mereológiai elmélet közül itt a Peter Simons és Archello Varzi nevéhez köthető elméletet fogadjuk el (Simons 2000, Varzi 1996).

A partitív relációra igaz az alábbi három tulajdonság.

mereológiai axiómák

(a1)	$\forall x(P(x, x))$
(a2)	$\forall x\forall y(P(x, y) \wedge P(y, x) \rightarrow x = y)$
(a3)	$\forall x\forall y\forall z(P(x, y) \wedge P(y, z) \rightarrow P(x, z))$

Az első három axióma rögzíti, hogy a P partitív reláció reflexív (a1), antiszimmetrikus (a2) és tranzitív (a3). A következő lépésben a P relációval néhány új mereológiai relációt definiálunk, majd az újakkal még újabbakat hozunk létre:

(1) PP(x, y)	$\forall x\forall y(P(x, y) \wedge \neg P(y, x))$	valódi része
(2) O(x, y)	$\forall x\forall y\exists z(P(z, x) \wedge P(z, y))$	átfedése
(3) U(x, y)	$\forall x\forall y\exists z(P(x, z) \wedge P(y, z))$	kívüllevősége
(4) OX(x, y)	$\forall x\forall y(O(x, y) \wedge \neg P(x, y))$	részhatáros átfedése
(5) UX(x, y)	$\forall x\forall y(U(x, y) \wedge \neg P(x, y))$	részhatáros kívüllevősége
(6) PO(x, y)	$\forall x\forall y(OX(x, y) \wedge OX(y, x))$	valódi lefedése
(7) PU(x, y)	$\forall x\forall y(UX(x, y) \wedge UX(y, x))$	valódi kívüllevősége

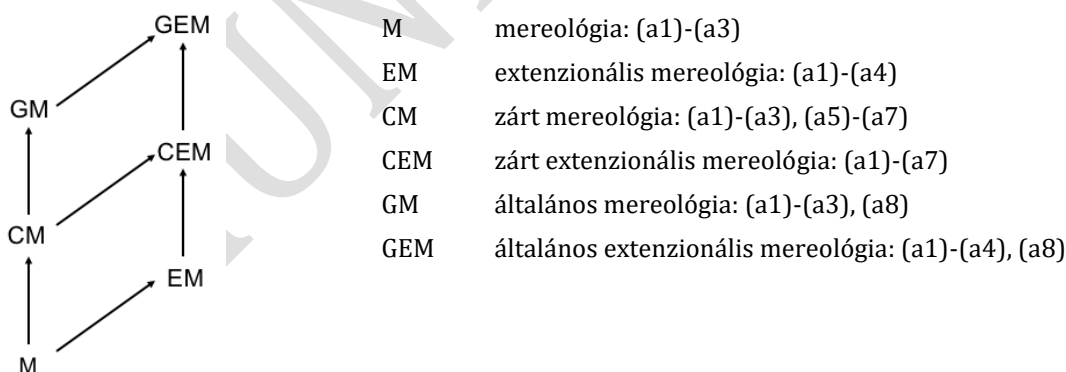
Sorban definiáljuk a következő relációkat: valódi része (1), átfedése (2), kívüllevősége (3), részhatáros átfedése (4), részhatáros kívüllevősége (5), valódi lefedése (6), valódi kívüllevősége (7). Ezek után a P része, O átfedése és U kívüllevősége relációk segítségével először meghatározhatjuk a kiterjeszhetőség (erős) elvét (a4), majd az (a5)-(a7) mereológiai axiómák elfogadásával definiálhatjuk a zárt (extenzionális) mereológia elméletét.

(a4)	$\forall x \forall y (\neg P(x, y) \exists z (P(z, x) \wedge \neg O(z, y)))$
(a5)	$\forall x \forall y (U(x, y) \rightarrow \exists z \forall w (O(w, z) \leftrightarrow (O(w, x) \vee O(w, y))))$
(a6)	$\forall x \forall y (O(x, y) \rightarrow \exists z \forall w (P(w, z) \leftrightarrow P(w, x) \wedge P(w, y)))$
(a7)	$\forall x \forall y \exists z ((P(z, x) \wedge \neg O(z, y)) \rightarrow \exists z \forall w (P(w, z) \leftrightarrow (P(w, x) \wedge \neg O(w, y))))$

További axiómák felvételével egyrészt definiálhatjuk az általános (extenzionális) mereológia elméletét (a8) másrészt meghatározhatjuk az atomos (a9), illetve atom-nélküli mereológiákat (a10).

(a8)	$\exists x \phi \rightarrow \exists z \forall y (O(y, z) \leftrightarrow (\exists x (\phi \wedge O(y, x))))$
(a9)	$\forall x \exists y (P(y, x) \wedge \neg \exists z PP(z, y))$
(a10)	$\forall x \exists y (PP(y, x))$

Nem szükséges a fenti axiómák mindegyike ahhoz, hogy egy mereológiai elméletet definiálni lehessen. Varzi többfajta mereológiai elméletet különített el más és más axiómák együttes elfogadásával.



Ezek a mereológiai elméletek különböző erősségű és tartalmú ontológiai elkötelezettségek mentén érvényesek. A fenti hat elmélet egyikében sem szerepel azonban az az utolsó két mereológiai axióma, amelyek alapján elkülöníthetjük az atomos és nem-atomos partitív relációkat egymástól. Az (a9) és (a10) jelű tételek alkalmazásával a fenti elméletek tovább bonthatók atomos vagy atom-nélküli aeleméletekre.

A partitív reláció is tranzitív. Jeleztük azonban, hogy a tranzitivitás nem általában a partitív reláció egészére, hanem annak valamelyik típusán belüli kapcsolatokra igaz. A fenti formulák alapján most már mondhatjuk azt, hogy a partitív reláció tranzitivitása mindig csak valamelyik – fent definiált – mereológiai elméleten belül érvényesül. A partitív reláció tranzitivitása akkor válik érvénytelenné, amikor egy atomos és atom-nélküli partitív relációra akarjuk érvényesíteni. Az emberi test és a testrészek között része viszonyt atom-nélküli partitív relációval írhatjuk le, ezzel szemben a testületek és személyek közti szervezeti tagságot atomos partitív reláció alkalmazásával tudjuk megragadni. A kétféle rész-egész relációt azonban nem szabad összekeverni, különben ahhoz a tranzitivitási deficethez jutunk, amit korábban már bemutattunk a zenekari tag zongorista keze kapcsán. Természetesen lehet többféle partitív relációt használni egy rendszeren belül, de ilyen esetekben a tranzitivitás létezését csak adott partitív reláción belül várhatjuk el, a – részben – eltérő minőségű partitív relációk közt nem számíthatunk rá. Mindebből az következik a névtérépítésre vonatkozóan, hogy – szükség esetén – többféle partitív relációt kell definiálni és használni, hogy az ilyen keveredést elkerüljük.

3.1.3.3 Szinonima

A szinonima reláció a nyelv formai többértelműségét mutató nyelvi reláció. Nyelvi konstrukciók (tehát nyelvi formák) közötti kapcsolatot fejez ki azáltal, hogy azok jelentései (tartalmi) közti azonosságot fejezi ki adott nyelven belül. A magyar nyelvben a 'kutya' és az 'eb' egymás szinonimái. Bármelyiket használjuk, ugyanarra az állatra utalunk vele. A következő két mondatnak – durván – ugyanezt a jelentést tulajdonítjuk.

A postást megharapta a kutya.

A postást megharapta az eb.

A szinonima relációban álló terminusok egymással felcserélhetők egy nyelvi konstrukción belül anélkül, hogy a megnyilatkozás jelentése ezáltal módosulna. Ebből az is következik, hogy a szinonima reláció szimmetrikus, meg az is, hogy valamifajta egyenlőséget fejezhetünk ki vele. A relációnak az a lényege, hogy a szinonim egységek ugyanarra a jelentésre mutatnak. A nyelvhasználatunk lényege az, hogy szintaktikailag jól-formált nyelvi konstrukciókat (szavakat, kifejezéseket, mondatokat) – konvencionális módon, tehát teljesen önkényesen – összekapcsolunk jelentésekkel. Amikor ezt tesszük, akkor a jelentések tartalmát nyelvi formák reprezentálják. A szinonimitás azt jelenti, hogy két – egymástól különböző – nyelvi formátum ugyanazt a tartalmat hordozza.

A szinonima reláció esetében fontos megkötés az egynyelvűség, hiszen két különböző nyelv egységei is mutathatnak ugyanarra a jelentéstartalomra, és ezeket nem szokás egymás szinonimáinak tekinteni. Az angol 'dog' ugyanazzal a jelentéssel bír, mint a magyar 'kutya' szó, de ezek inkább egymás nyelvi fordításai, mint egymás szinonimái.

Fontos megjegyezni, hogy a szinonima reláció által összekapcsolt nyelvi egységeknek a valóságban sosincs teljesen azonos jelentése, bizonyos nyelvhasználati kontextusokban nem lehet feltétlen kicserélni egymással a – máshol, máskor – szinonim egységeket. A „kutya hideg van odakint” mondat helyett nem mondhatjuk azt, hogy 'eb hideg van odakint'.

3.1.3.4 Homonima

A homonima reláció a nyelv tartalmi többértelműségét mutató nyelvi reláció. Jelentések (tehát nyelvi tartalmak) közötti kapcsolatot fejez ki azáltal, hogy azok nyelvi megjelenései (formái) közti azonosságot fejezi ki adott nyelven belül. A homonimának két típusát el lehet különíteni aszerint, hogy írásos vagy szóbeli nyelvi egységekről van-e szó. Előbbi esetben homográfáról, utóbbi esetben homofóniáról beszélnek. Mivel itt rögzített, tehát írás-alapú kommunikációról van szó, a homonima relációt mindig írásos nyelvi egységek között értelmezzük. A homonima esetén arról van szó, hogy ugyanazzal a nyelvi alakkal más jelentés fejezünk ki, tehát különbözik a tartalom. A homonima leggyakrabban idézett, klasszikus, magyar példája az 'ár' szóval kapcsolatos. Az alábbi négy nyelvi egység alakilag megegyezik, de jelentését tekintve különböző.

ár mint árhullám
ár mint értékmérő a piacon
ár mint szűrőszerszám
ár mint területi mértékegység

A homonima reláció meghatározásakor ugyanazokat a megoldásokat kell alkalmaznunk, mint a szinonima reláció esetében, csak itt az azonosságot az eltérő jelentésű megnyilatkozások nyelvi alakjaira kell vonatkoztatnunk.

3.1.3.5 Meghatározója (előzménye, következménye) reláció

A tudásszervezési rendszerek közül a teauruszokban használják azt a relációpárt, amelyet nem igazán lehet egyértelmű és egzakt jelentéssel ellátni, inkább többféle jelenséget egybefogó ernyőfogalomként lehet értelmezni. Ebben a – sokszor az okság fogalmához kapcsolódó – szövevényes jelentéshálóban is lehet inverzrelációt definiálni, annak érdekében, hogy a különböző alkalmazási helyzetekben a kényelmesebb megoldást (a kényelmesebb irányban kifejezett relációt) lehessen választani.

A meghatározója relációpár nagy általánosságban az ok-okozat, cél-eszköz-következmény, tárgy-folyamat-funkció fogalomcsoporttal jellemezhető kapcsolatokat fejezi ki a dolgok között. Valamiféle meghatározottságot jelölhetünk vele. Amikor ennek a meghatározottságnak az eredmény oldalára fókuszálunk, akkor beszélhetünk a **rezultánsa** relációról. Ezzel az jelöljük, hogy a bal-oldali relátumnak (dolognak, tárgynak, folyamatnak) mi a meghatározottja (rendeltetése, okozata, eredménye, terméke, célja, tárgya, következménye), amit a jobb-oldali relátummal fejezünk ki. Az inverzreláció, amit szoktak **előfeltétele** vagy kiindulása relációnak nevezni, azt fejezi ki, hogy a bal-oldali relátum (dolog, tárgy, folyamat) mi a meghatározója (létének, létrehozásának, működésének, meghatározásának oka, előidézője, eszköze, alapja), amire a jobb-oldali relátummal utalunk.

Ha a relációkettőst szemantikailag kellő mértékben szűk jelentéstartományon belül tartjuk, akkor elképzelhető, hogy hierarchikus struktúrákat lehet építeni ebben a dimenzióban is. Például bizonyos típusú események esetében lehet beszélni az egymást követő események hierarchikus láncolatáról.

3.1.3.6 Hierarchikus alá-fölérendeltje (tartalmazás) reláció

A generikus és a partitív reláció mentén a tudásszervezési rendszerek elemei hierarchikus struktúrába, alá-, fölérendeltségi kapcsolatrendszerbe szervezhetők. Mivel ez a két reláció szabatosan definiált, mindkettő egyértelmű, konzisztens struktúrát eredményez. Vannak azonban olyan tudásszervezési rendszerek is, amelyben csak annyit követelnek meg, hogy a rendszer egységei között létezen valamilyen tartalmazási, hierarchikus alá- és fölérendelési kapcsolat úgy, hogy a relációk szemantikai egyértelműségét nem követelik meg, csak az elemek között értelmezhető hierarchikus jelleg létét. Angolul pontosan fejezi ki ezt a minőséget a 'broader term', illetve 'narrower term' kifejezőpár, ami csak arra utal, hogy valamilyen szempont szerint az így összekapcsolt kategóriák között valamilyen mennyiségi, tartalmazási kapcsolat létezik, semmi több.

Amikor egy hierarchikus rendszert úgy építünk fel, hogy közben a tartalmazásnak többféle értelmét engedjük meg egyszerre, akkor egy pongyola, mert jelentését tekintve többértelmű, de ettől még jól rendezett struktúrához jutunk. Ezt a pongyolást és a belőle fakadó inkonzisztenciát ezért szokás bevállalni, mert ilyen esetekben a rendszer építése, használata egyszerűbb, hiszen kevesebb – olykor szemantikai értelemben nehéz – intellektuális döntést kell hozni.

3.1.3.7 Lexikografikus rendezési reláció

A lexikografikus (ábécé szerinti) rendezési reláció nyelvi egységek lineáris sorba rendezését jelenti, matematikai értelemben **gyenge teljes rendezésnek** felel meg a **nyelvi egységek** halmazán. A könyvek, tanulmányok végén található hivatkozáslisták bibliográfiai tételekből állnak (általában a szerzők neveivel kezdődve), és ezeket a tételeket lexikografikus rendezés szerint sorolják be egymás mögé. A tudásszervezési rendszerek egységeinek, terminusainak megjelenítését, felsorolását általában a legfontosabb struktúraképző reláció mentén szokták megoldani, de sokszor adódhat olyan helyzet, amikor az egységeket vagy azok egy részét (például egy hierarchián belül az azonos szinten levő, azonos szülő alá tartozó egységeket) a lexikografikus rendezéssel teszik ki egymás után.

3.1.3.8 Rokona (egyéb) reláció

A tudásszervezési rendszerek világában a rokona reláció nem a biológiai értelemben vett leszármazási rokonságot vagy a társadalmi értelemben társadalmi viszonyt jelenti, hanem pusztán csak arra szokás utalni vele, hogy az általa összekötött két relátum között van-e valamilyen – kellő pontossággal nem meghatározható – hasonlóság. Ezt a relációt akkor veszik fel két elem között, ha az adott tudásszervezési rendszeren belül a jelezni kívánt kapcsolat nem feleltethető meg egyik rendszeren belüli, jól-definiált relációnak sem (például teauruszokon belül a generikus vagy partitív relációnak), de egy gyengébb hasonlósági, rokonsági relációval érdemes rögzíteni ezt a gyenge, csak egy pongyola, nem jól-definiált relációval kifejezhető kapcsolatot.

3.2 A használatba vehető relációk listája

Az eddig elemzett relációk közül az alábbiakat fogjuk a tudásszervezési rendszerek leírásakor használni. Az egyszerűség kedvéért itt „egyként” hivatkozunk azokra a relációkra is, amelyeket az inverz párjukkal együtt szoktak alkalmazni. Az egyértelműség kedvéért jelezzük azt is, ha az elméleti bemutatás során alkalmazott névhez képest másfajta megnevezés mentén kerülhet használatba egy reláció.

KOS-relációk listája

R_1	lexikografikus rendezési reláció	
R_2	ekvivalenciareláció	
R_3	hierarchikus alárendeltje (tartalmazás) reláció	alárendeltje
R_4	különbözőségi reláció	
R_5	generikus alá- és fölrendeltje relációpár	fölrendeltje-alárendeltje, faja-neme relációpár
R_6	partitív alá- és fölrendeltje relációpár	része-egésze, pars-totum relációpár
R_7	következménye-előzménye relációpár	meghatározója (rezultáns)-meghatározója (előfeltétel) relációpár
R_8	rokona (egyéb) reláció	
R_9	lásd és helyette szinonimareláció-pár	
R_{10}	címkegyakorisági függvény	
R_{11}	címke-címke együttjárasi reláció	
R_{12}	felhasználó-címke együttjárasi reláció	
R_{13}	felhasználó-dokumentum együttjárasi reláció	
R_{14}	felhasználói aktivitás függvény	
R_{15}	dokumentumcímkezési gyakorisági függvény	
R_{16}	tetszőleges relációfogalom	
R_{17}	tetszőleges osztályfogalom	
R_{18}	tetszőleges attribútumfogalom	

3.3 Tudásszervezési rendszerek

A nevek rendezett gyűjteményét, vagyis nevek adott halmazát és a rajtuk értelmezett relációk együttesét matematikai értelemben **struktúrának** nevezzük. Más ismeretterületeken ugyanerre a képződményre a **tudásszervezési rendszerek** (Knowledge Organization System, KOS) fogalmát alkalmazzák. Többféle tudásszervezési rendszert szokás elkülöníteni az egyszerűbb struktúráktól az egyre összetettebbek felé haladva, úgymint a **terminuslista**, az **osztályozási rendszer**, a **tezaurusz** és az **ontológia**, illetve az építkezési logikát tekintve el lehet választani az előző kontrollált rendszereket a kontroll-nélkül szervezett **folkszonómiától**.

A tudásszervezési rendszerek a világ leírásához szükséges terminusokat gyűjtik egybe, és a terminusok közötti relációkon keresztül a megkönnyítik a hozzáférést magukhoz a terminusokhoz. A különböző típusú tudásszervezési rendszerek igazából csak a relációikban különböznek egymástól, a felvehető terminusokra nincs semmilyen ontológiai megkötés. Elvileg mindegyik rendszer alkalmas arra, hogy a terminusaival tetszőleges tudásterületet lefedhessünk.

A tudásszervezési rendszerek bizonyos értelemben érzéketlenek arra, hogy tulajdonneveket vagy közneveket kell-e kezelniük, gyakran elő is fordul, hogy egy konkrét tudásszervezési rendszerben keverten szerepelnek tulajdon- és köznevek.

A tudásszervezési rendszerekre vagy legalábbis egy részükre lehet (szokták) alkalmazni a névtér fogalmát, ami a fenti értelmezések alapján érthető, kezelhető. Szerencsésebb lenne a tulajdonnév és köznévi rendszereket jobban elkülöníteni egymástól, de inkább az a fontos, hogy adott rendszerről tudjuk, milyen névállományból épül fel. Ha fontos, a tulajdonnévtérrel megfelelő jelző felvételével lehet jelezni.

3.3.1 Formális keretrendszer

A tudásszervezési rendszerek abban a különböznek egymástól, hogy a rendszer elemei között milyen relációkat engedélyeznek felvenni. Éppen ezért matematikai értelemben a tudásszervezési rendszerek matematikai struktúráknak tekinthetők.

Matematikailag a KOS a tárgyszavak tartóhalmazán értelmezett struktúra, amelyből természetesen többféle is lehetséges annak megfelelően, hogy milyen relációkat engedünk meg a tárgyszavak között. A struktúra az alábbi formában írható fel.

$KOS = \langle D, R_1, R_2, \dots, R_i, R_n \rangle$, ahol

D a tudásszervezési rendszer tárgyszavaiból álló tartóhalmaz,

R_i az elemeken (tárgyszavakon) értelmezhető reláció.

A matematikai leírást azonban ki kell egészítenünk, ugyanis később látni fogjuk, hogy a tudásszervezési rendszerek különböző típusai nemcsak a tárgyszavaikban és a relációikban, hanem más jellemzőikben is eltérhetnek egymástól. Mielőtt azonban bővítenénk a

tudásszervezési rendszer fogalmának összetevőit, a tudásszervezési rendszerek típusairól kell pár szót szólnunk. A különböző archívumépítési gyakorlatokban, a hálózati kultúra időszakát megelőző évszázadban, érdemben és kiterjedt módon háromféle tudásszervezési rendszert vettek használatba, majd a hálózati kommunikáció további típusok megjelenését tette lehetővé. Az alábbi öt KOS-típusról van szó.

- hagyományos archívumokban (is)
 - terminuslista
 - taxonómia
 - tezaurusz
- csak digitális archívumokban
 - ontológia
 - folkszonómia

Ezek mindegyike a rá jellemző matematika struktúrával, valamint a rendszerek gyakorlati működtetésére vonatkozó szabályrendszerrel írható le. A tudásszervezési rendszerek alaphalmaza (D) azokat a szavakat, kifejezéseket (terminusokat) tartalmazza, amelyeket a dokumentumokhoz lehet rendelni. A tudásszervezési rendszerek különbségeit az a tény határozza meg, hogy milyen – szintaktikai, szemantikai vagy pragmatikai – relációkat (R_i) engedünk meg felvenni a rendszer elemei között. Nem elég azonban csak a matematikai struktúrára figyelni, ha igazán meg akarjuk érteni a tudásszervezési rendszerek jelenségét. Azt is fel kell vennünk e rendszerek jellemzői közé, hogy van-e, és ha igen, akkor milyen felügyelet, milyen kontroll van a metaadat-hozzárendelési munka folyamatában. Előbb persze meg kell mondanunk, miért is van szükség ennek fi gyelembevételére. Nos, ha a tartalmi metaadatok dokumentumokhoz rendelésének az a fő funkciója, hogy egyértelműen jellemezzük általuk a dokumentumok tartalmát, akkor az egyértelműséget (vagyis a többértelműségek elkerülését) biztosítanunk kell valahogy. A többértelműségek elkerülése pedig megfelelő szakértelmet, fegyelmezett munkarendet, szakmai kontrollt, kontrollált szótárakat és tudásszervezési rendszereket kíván. Ezt persze nem olyan könnyű formalizálni, hiszen olyan kérdésekre kell választ találnunk e szempont alapján, mint:

- $Q_1 = \{\text{támasztanak-e bármilyen szakmai feltételt, szaktudással kapcsolatos elvárást a munkát végzők személyével szemben?}\}$
- $Q_2 = \{\text{van-e bármilyen munkaszervezési szabályrendszer, ellenőrzési mechanizmus a munka menetére vonatkozóan, azaz kik, milyen jogosultságokkal vehetnek részt a munka egyes részfolyamataiban?}\}$
- $Q_3 = \{\text{kik rendelhetik a KOS-rendszer elemeit a dokumentumokhoz?}\}$
- $Q_4 = \{\text{kik szerkeszthetik, módosíthatják, bővíthetik a tudásszervezési rendszer elemeit, relációit?}\}$

A tudásszervezési rendszerek építésének és alkalmazásának kontrolljára vonatkozó fenti Q_i kérdésekre különböző válaszokat adhatunk, és ezt a feltételegyüttest vagyis az S_j társadalmi normák összefüggő rendszerét érdemes felvenni a tudásszervezési rendszer jellemzői közé. A dolgokat kissé leegyszerűsítve a következő tevékenységekre vonatkozó normákat kell rögzítenünk:

- S_1 – kinek szabad új tárgyszót létrehozni a tudásszervezési rendszerben
- S_2 – kinek szabad új relációt létrehozni a tudásszervezési rendszerben
- S_3 – kinek szabad két tárgyszót relációba állítani a tudásszervezési rendszerben
- S_4 – kinek szabad tárgyszót dokumentumhoz rendelni a katalógusban
- S_5 – csak a tudásszervezési rendszer elemeit szabad-e a dokumentumokhoz rendelni

Az öt norma közül az első három a tudásszervezési rendszerek építésével, az utolsó kettő a katalógusok bővítésével kapcsolatos. A számítógépek világában a fenti normák mind kezelhetők azáltal, hogy a digitálisan szabályozzuk, kinek van írási joga a tudásszervezési rendszer és/vagy a katalógus elemeire, illetve milyen adatokat lehet egymással összekapcsolni. Az írási, szerkesztési jogosultságokat is figyelembe véve már felírhatjuk a tudásszervezési rendszerek teljesebb formuláját, amelynek segítségével aztán majd megmutathatjuk az egyes típusok közti különbségeket is:

$KOS = \langle D, R_1, R_2, \dots, R_n, S_1, S_2, S_3, S_4, S_5 \rangle$, ahol

D a tudásszervezési rendszer tárgyszavaiból álló tartóhalmaz,

R_i az elemeken (tárgyszavakon) értelmezhető reláció ($i=1, \dots, n$)

S_j a tudásszervezési rendszer

A fent bemutatott összetevőkkel már adott az az általános keret, amelyre támaszkodva elég pontosan megragadhatjuk a történelmileg létező, szélesebb körben elterjedt tudásszervezési rendszerek legfontosabb jellemzőit. Minden tudásszervezési rendszerben van egy közös reláció, a lexikografikus rendezés, amely a tárgyszóhalmaz elemeinek ábécé szerint való sorba állítását jelenti. A kérdés az, hogy melyek azok a relációk (és normák), amelyek alapján különbséget tehetünk az egyes típusok között. A KOS-ok leírásához pontosan definiált relációkat kell használnunk, amelyeket korábban már bemutatunk, és most felsoroluk az alábbi táblában.

KOS-relációk listája, 2.

R_1 lexikografikus rendezési reláció	R_{10} címkegyakorisági függvény
R_2 ekvivalenciareláció	R_{11} címke-címke együttjárési reláció
R_3 hierarchikus alárendeltje (tartalmazás) reláció	R_{12} felhasználó-címke együttjárési reláció
R_4 különbözőségi reláció	R_{13} felhasználó-dokumentum együttjárési reláció
R_5 generikus alá- és fölrendeltje relációpár	R_{14} felhasználói aktivitás függvény
R_6 partitív alá- és fölrendeltje relációpár	R_{15} dokumentumcímkézési gyakorisági függvény
R_7 következménye-előzménye relációpár	R_{16} tetszőleges relációfogalom
R_8 rokona (egyéb) reláció	R_{17} tetszőleges osztályfogalom
R_9 lásd és helyette szinonimareláció-pár	R_{18} tetszőleges attribútumfogalom

A bemutatott keretrendszer és a fenti relációk alapján az egyes tudásszervezési rendszereket a következő relációk segítségével definiálhatjuk.

3.3.2 Terminuslista

A *terminuslistáknak* van a legegyszerűbb szerkezetük, mert ott a tartóhalmaz elemein csak egy *lexikografikus rendezést* engedünk meg. Ilyenek a könyvek végén található indexek (név- és tárgymutatók), amelyek a könyv legfontosabb kulcsszavait sorolják fel ábécé szerint, minden kulcsszóhoz hozzárendelve az oldalszámokat, ahol a kulcsszavak a szövegben előfordulnak. De ide tartoznak a különféle egységesített névlisták, sőt a keresőmotorokban használt 'invertált index' technológiája is.

Copyrighted material

⇒⇒⇒

L

Licensed Practical Nurse, 46
Living trusts, 72–73, 80
Living will, 87–89, 230
Long-term care insurance policies, 59–62

M

Massage, 206–207
Meals and feeding, 192–93—
Alzheimer's patients, 194
appetite, 194–95, 238, 240
eating aids, 127–28
solutions to problems, 195–96
see also Nutrition and diet
Medicaid, 12, 13, 15, 30, 56–58—
eligibility, 57, 79–80
services not covered, 57
waiver programs, 56
Medicare, 13, 16, 22, 53–55—
claims disputes, 55
eligibility, 53
hospice coverage, 311
what is covered, 20–21, 22, 30, 34, 41–42, 54, 119
what is not covered, 53, 54
Medication, 26–29, 30–31—
cautions, 144–45
directions for taking, 28
drug interaction, 28
generic drugs, 27, 28
label abbreviations, 144
over-the-counter drugs, 28
overdoses, 29
record keeping, 143–44

Medigap (Medicare supplemental) insurance, 55
Mental health considerations, 31—
Alzheimer's, 288–305
boredom, 219
dementia, 288
depression, 217–18, 233–34
fear of aging, 218
Moving patients—
helping to stand, 279
in bed, 270–77
transfer to and from wheelchair, 280–85
using mechanical lift, 278–79

N

Nail care, 164, 170
Nursing homes, 4, 11–14
Nutrition and diet—
fluid intake, 242
food safety, 239
increasing calories and protein intake, 240–41
nutritional guidelines for elderly, 240
prevention of bone loss, 242–43
sources of nutrients, 243–44

O

Occupational therapy, 204
Odor prevention, 187–88
Older Americans Act, 59

P

Pain management, 234—
reduction techniques, 224–26
types of pain, 224

362 THE COMFORT OF HOME

tárgyszómutató mint terminuslista egy könyvből

Mivel a terminuslisták elemei között csak kevés relációt vesznek fel, igazából csak arra lehet használni őket, hogy egyfajta leltárként funkcionáljanak, mutassák, hogy az adott rendszeren

belül milyen terminusok, nyelvi egységek fordulnak elő. Ehhez persze lehet további információkat is hozzárendelni (mint a fenti ábrán látható tárgyszómutató esetében), de ez a hozzárendelés már olyan hozzáadott értéként tekinthető, amely a terminuslista használatából fakad, nem pedig magából az alaprendszerből. Ettől még érdemes megbecsülni ezt az egyszerűséget, mert sok esetben komoly előny származhat abból, ha ilyen terminuslistákat „nyilvántartási” célokra használunk. **A szűkebben értelmezett tulajdonnévterek is terminuslistának minősíthetők.**

3.3.2.1 Formális jellemzés

A terminuslista a legegyszerűbb struktúrával rendelkező tudásszervezési rendszer. Ez a lapos struktúra nem kínál túl sok funkciót, viszont a rendszer nagyon egyszerűen kezelhető. A terminuslista formulája a következőpp írható fel.

$KOS_{tlist} = \langle D, R_1, R_2, S_1, S_4, S_5 \rangle$, ahol

R_1 lexikografikus rendezési reláció

R_2 ekvivalenciareláció

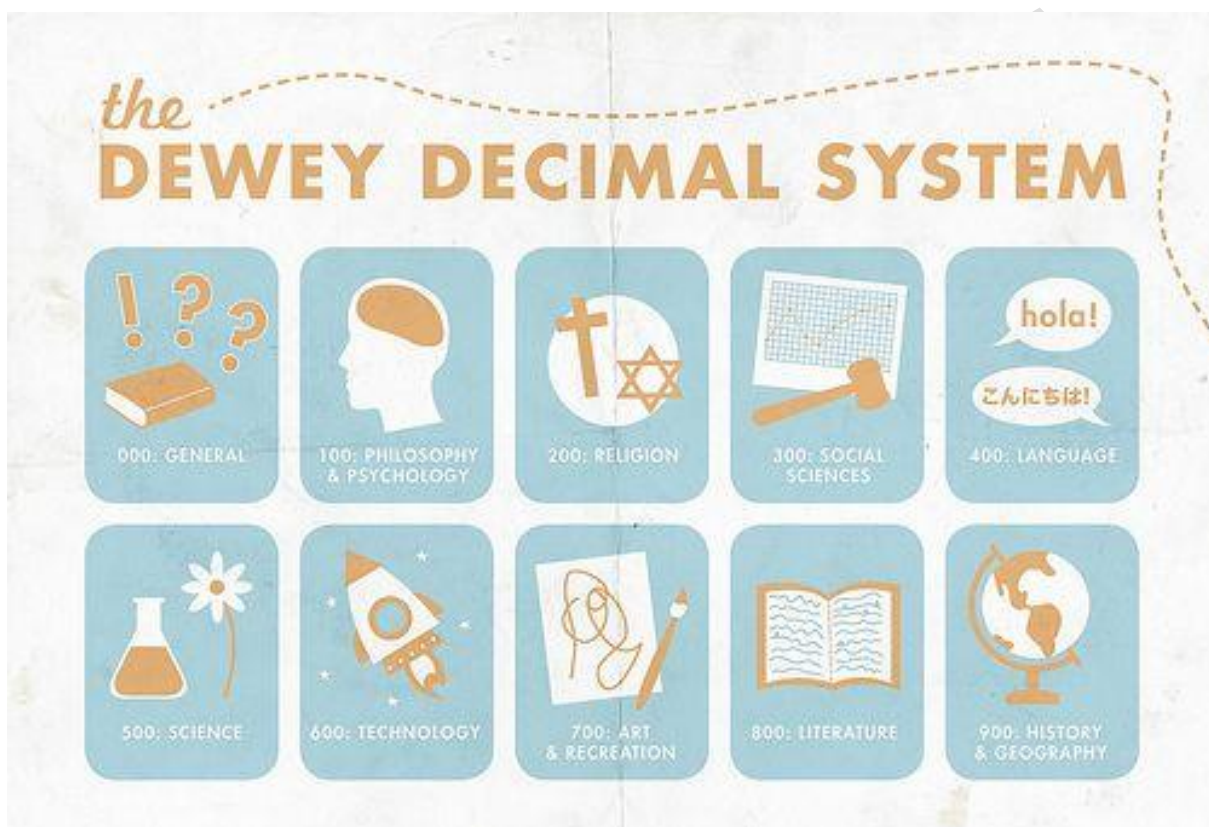
A terminuslista esetében az S_2 és az S_3 szabály nem érvényesíthető, hiszen ebben a rendszerben új relációt nem lehet definiálni. Az ilyen rendszernek a lexikografikus rendezés mellett van még egy másik relációja: a lista elemeit az R_2 ekvivalenciareláció kapcsolja össze. A földrajzi nevek egységesített adatbázisát például az kapcsolja össze egyetlen rendszerré, hogy minden tételéről azt állítjuk, hogy ekvivalensek egymással abban a tulajdonságukban, hogy valamennyien földrajzi entitások tulajdonnevei. Az ekvivalenciareláció fenntartásával azt kell „garantálnunk”, hogy a terminuslista elemei – az alkalmazott szempont szerint – azonos minőségűek lesznek (tehát nem keverednek különböző típusú elemek, mondjuk személynevek a földrajzi nevekkel). Más relációt nem lehet a terminuslista elemei közé felvenni. Attól függően, hogy milyen típusú terminusokról van szó, változhat az a gyakorlat, hogy fenntartják-e a a szavak, kifejezések bekerülését szabályozó S_1 normát. Az igazán komolyan vett két előírás a katalógusépítésre vonatkozik, ezek a tárgyszó és a dokumentum összekapcsolását szabályozó S_4 és S_5 normák.

3.3.3 Taxonómia

Elmélyült elemzések segítségével különbséget lehetne tenni a taxonómia, az osztályozási rendszer, a klasszifikáció fogalmai között, de itt nem szükség figyelembe vennünk ezeket a finom eltéréseket, így a továbbiakban azonos jelentéssel használjuk ezeket a kategóriákat.

A klasszifikációs rendszereknek a klasszikus példája a könyvtárak világában használt ETO, az Egységes Tizedes Osztályozási rendszer (UDC, Universal Decimal Classification), vagy annak őse a DDC, a Dewey Decimal Classification system. A könyvtárakban a kezdetektől fogva elemi igény volt arra, hogy a könyveket könnyen és gyorsan meg lehessen találni tartalmi leírások, címkék alapján. Ennek az igénynek a kielégítésére dolgozták ki és tartották fent a decimális osztályozási

rendszereket, amelyek a könyvek tartalmának jellemzéséhez, leírásához adtak segítséget azzal, hogy egy kontrollált és hierarchikusan összeszervezett osztályozó rendszert kínáltak fel. Az ETO mindent átfogó információs (indexelő és kereső) nyelv, decimális jelzetelésű osztályozási rendszer, amely egyrészt a 10 főosztályból álló táblázatból (amelyek 10-10 osztályra, azok mindegyike 10-10 alosztályra, azok pedig 10-10 szakcsoportra oszthatók, fogalmaikhoz alosztások fűzhetők), másrészt az osztályozó fogalmak összekapcsolására szolgáló szabályokból áll.



a DDC legfelső szintű kategóriái

Elvileg tetszőleges mélységig le lehet menni a hierarchiában, ami nagyon pontos leírás lehetőségét kínálja, bár ez a másik oldalról komoly szerkesztési és fenntarthatósági kérdéseket is felvet.

Dewey has a rich notational structure

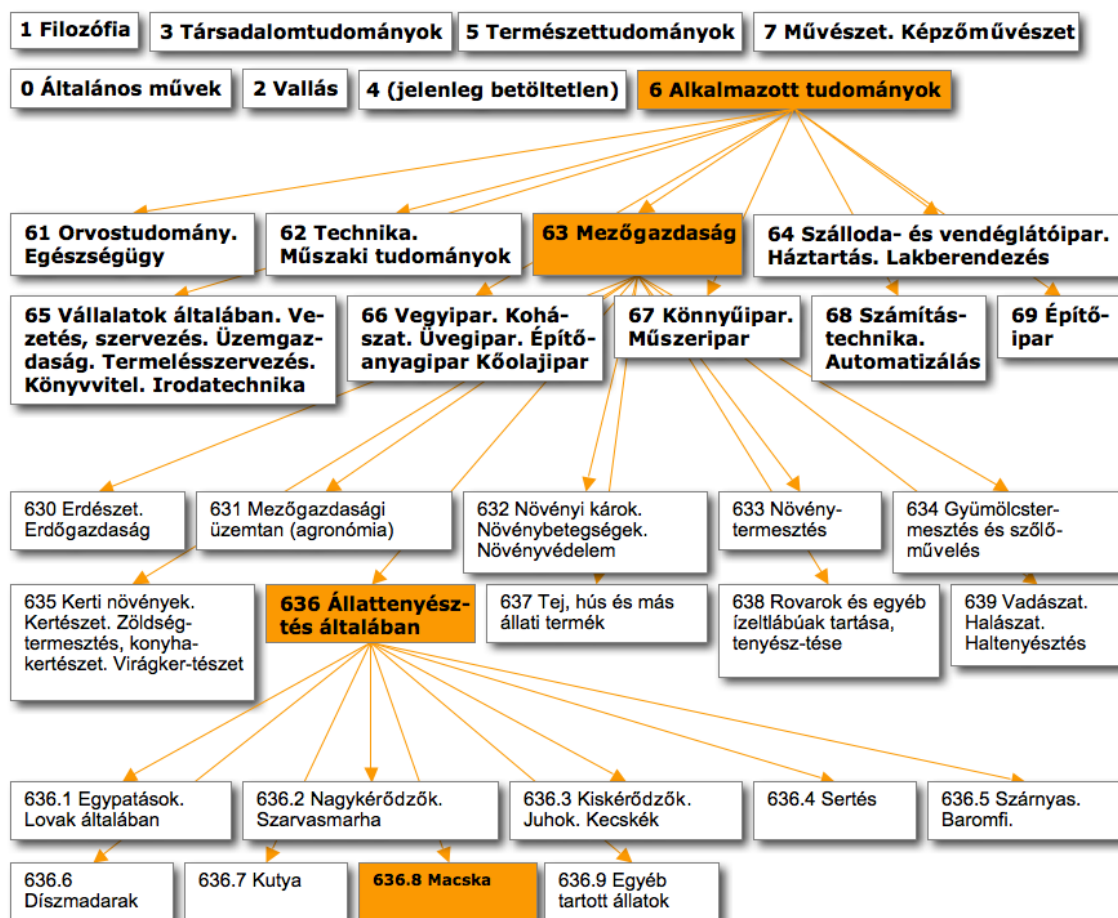
The Lancashire cotton industry : a study in economic development

Assigned DDC Code: 338.4767721094276



a DDC hierarchiájának alábontása több, mint húsz mélységű tételig

A hierarchikus szerveződés miatt nagyon sok osztályozási jelzetet lehet a rendszerbe felvenni, és a tizes osztályozási logika miatt gyorsan lehet benne keresni. Lehet olyan nézetben is megmutatni a rendszer egységeit, amelyek szemléletesek lehetnek a használók számára, bár ez az elvi lehetőség a gyakorlatban sokszor nem áll rendelkezésre.



a 'macska' lexikai egység helye az ETO-ban

A taxonómiák népszerűsége az egyszerű kezelhetőségükben rejlik. Ez az egyszerűség persze viszonylagos. A terminuslistákhoz képest ugyanis itt már szemantikai elvárásokat kell figyelembe vennünk, hiszen a hierarchikus alárendelési reláció alkalmazása (akármi legyen is az értelme egy konkrét taxonómia esetében) szükségszerűen mindig szemantikai kényszerek betartásával jár együtt. Ezért van az, hogy ezen rendszerek használata esetében már megkövetelnek valamilyen szaktudást, és szabályozzák a munkafolyamat menetét is. Utóbbi mozzanat több részre osztható. A taxonómia elemeinek halmazát felfoghatjuk olyan *kontrollált szótárként* is, amelynek elemeit nem tetszőleges módon, hanem csak adott szabályokhoz igazodva, tehát csak kontrollált módon lehet bővíteni. Ez egyfelől korlátot jelent a tárgyszó-hozzárendelési munka során, mert előírásokhoz igazodó, tehát *fegyelmezett* munkavégzést követel meg az erre a feladatra előzetesen felkészített, *képzett archivátoroktól*, másrészt az ilyen rendszernek szüksége van egy olyan folyamatra, amely során a folyamatos változtatási igényeket ki lehet elégíteni, vagyis bővíteni, módosítani kell a rendszer valamely részét. Ez azt is jelenti, hogy a rendszer fenntarthatóságának érdekében szükség van *taxonómiaépítő szaktudásra*. Ezen feltételek, pontosabban az ezek teljesülésére vonatkozó kérdés azonban felvet két újabb, nagyon fontos tudásszociológiai, tudományfilozófiai kérdést:

- $Q_5 = \{ \text{mennyire egységesen értelmezi az osztályozó közösség a taxonómia elemeit?} \}$
- $Q_6 = \{ \text{lehet egyetlen egységes rendszerbe rendezni valamely dokumentumgyűjteményt jellemző tudásterület fogalomkészletét?} \}$

A kérdések megválaszolása szétfeszítené jelen tanulmány kereteit, ezért a részletes kifejtéstől el kell tekintenünk. A hierarchikus jelleg elvileg a gyors kereshetőséget támogatja, de azt meg kell itt jegyezni, hogy az ilyen struktúrák sok esetben merevekké válhatnak, és olykor a létrehozóik elfogultságai is visszatükröződnek bennük. Kiváló példát hozott erre Clay Shirky, aki a DDC elemzése közben vette észre, hogy a történelmi régiók tárgyszó alá sorolt címkék között „egy szinten” szerepelt Ázsia, Afrika és a Balkán félsziget (Shirky 2005).

D: History (general)

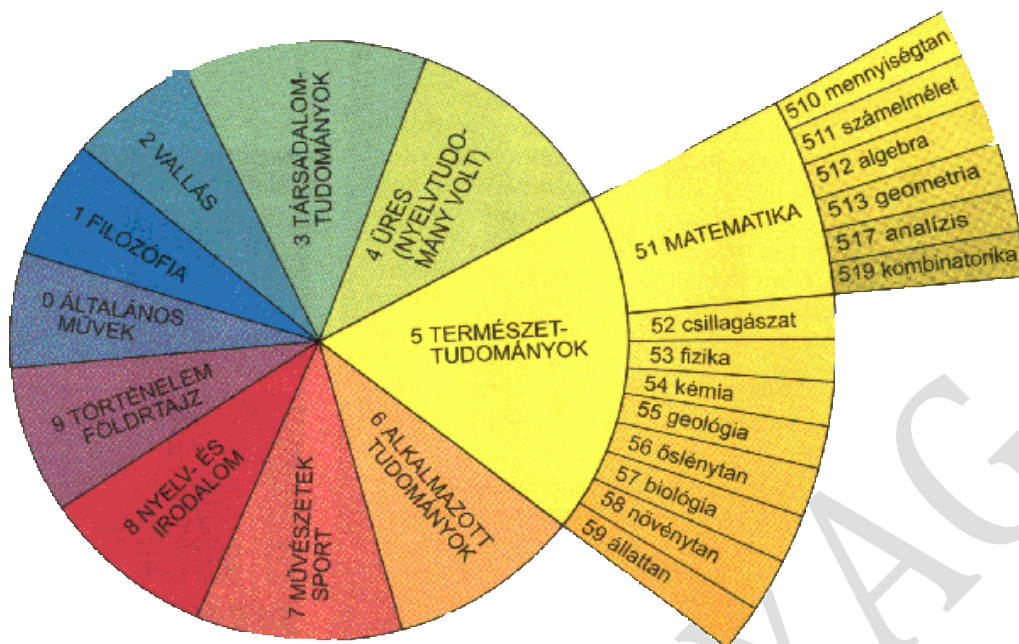
DA: Great Britain	DK: Former Soviet Union
DB: Austria	DL: Scandinavia
DC: France	DP: Iberian Peninsula
DD: Germany	DQ: Switzerland
DE: Mediterranean	DR: Balkan Peninsula
DF: Greece	DS: Asia
DG: Italy	DT: Africa
DH: Low Countries	DU: Oceania
DJ: Netherlands	DX: Gypsies

Ez azért volt feltűnő, mert a méretét (és ezáltal kicsit a jelentőségét) tekintve a Balkán félsziget jóval kisebb a másik két kontinenshez képest. Ezt az aránytalanságot jelezte azzal Shirky, hogy bekarikázta a három területet egy földgömbön.



Clay Shirky ábrája a DDC kapcsán

Shirky végül rájött arra, mi lehet az oka ennek a furcsaságnak. Azért volt azonos szinten a három régióra mutató egység, mert nagyjából azonos számban adtak ki könyvet Amerikában a három földrajzi régióról, tehát a könyvtári polcokon elfoglalt helyigényük alapján voltak egyenrangú kategóriák a könyvtárosok számára. Akármennyire is jogos és elfogadható szempont ez a könyvtáros világ számára, ez a szempont nyilván elfogult. Ezek után pedig fel lehet tenni a kérdést, hogy vajon lehet-e elfogultság nélkül tudásszervezési rendszert építeni? A valószínűsíthető válasz pedig az, hogy nem nagyon. Minden tudásszervezési rendszernek lehet létjogosultsága egy adott tudásterületen, az arra jellemző elfogultságokat figyelembe véve, de bajos olyan rendszert építeni és feltételezni, amely univerzális igényvel léphetne fel, vagyis azzal a céllal, hogy minden tudásterületen, minden alkalmazási célra egyaránt felhasználható legyen.



az ETO rendszerének főbb kategóriái a természettudományokon belül

Az ETO folyamatosan karbantartott, klasszifikációs rendszer, amely alapvetően köznévtérnek minősíthető, még ha vannak benne tulajdonnevek is.

3.3.3.1 Formális jellemzés

A *taxonómiák* (más néven *osztályozási* vagy *klasszifikációs rendszerek*) már két fontos szemantikai relációt tartalmaznak (a lexikografikus rendezés – „kötelező” – szintaktikai relációján túl). Ez a tudásszervezési rendszer úgy van felépítve, hogy az elemei hierarchikusan egymás alá vannak rendelve – valamilyen *tartalmazási reláció* alapján. Ezt az alárendelési relációt lehet tiszta és pongyola értelmezésben is használni (a tiszta értelmezés esetben az alárendelési reláció a *generikus alárendeltje* relációval egyezik meg, a pongyola megközelítés keverten alkalmazza a generikus és a *partitív*, vagy más hierachikus relációt, például az *előzménye* viszonyt). A *hierarchia* leírásához azonban nem elégséges egyetlen relációt értelmeznünk a rendszeren, noha a közvélekedés gyakran megelégszik ezzel a megoldással. Arra is szükség van, hogy egy második relációval biztosítani lehessen azt, hogy az azonos felettes elem alá rendelt elemek különbözzenek egymástól, vagyis definiálni kell egy *különbözőségi relációt*.⁶ A taxonómia rendszerét a következőképpen írhatjuk le:

⁶ Lehetne még erősebb feltételt is előírni és a JEPD-elv teljesülését elvárni. Ez annyival több a közvetlenül függő elemek különbözőségének elvárásától, hogy azt is megköveteli, hogy az azonos szinten levő fogalmak „együttes terjedelme” megegyezzen a fölöttes elem terjedelmével. A JEPD-elv (Jointly Exhaustive and Pairwise Disjoint) magyar fordítása 'együttesen kimerítő és kölcsönösen kizáró' lehetne. Bővebben lásd: [Bittner et al. 2004]

$KOS_{tax} = \langle D, R_1, R_3, R_4, S_1, S_3, S_4, S_5 \rangle$, ahol

R_1 lexikografikus rendezési reláció

R_3 hierarchikus alárendeltje (tartalmazási) reláció

R_4 különbözőségi reláció

Ebben a szisztémában az S_j szabályok közül már négyet érvényesítenek, csak az S_2 norma hiányzik, hiszen a két szemantikai kapcsolaton túl nincs mód más reláció alkalmazására a rendszerben, ebből következően pedig nincs is szükség az új relációk felvételét szabályozó normára.

3.3.4 Tezaurusz

Ahogy a közgyűjtemények világán belül a taxonómiák a könyvtári világban jöttek létre először, ott terjedtek el széles körben, úgy a tezauruszok használata is ehhez az intézménytípushoz köthető. A tezauruszok iránti igényt a taxonómiák gyenge és könnyen inkonzisztenssé váló struktúrájával való elégedetlenség teremtette meg. Az osztályozási rendszerben csak egy lényegi relációtípust, a hierarchikus alá-fölérendeltséget lehet alkalmazni, és ez rákényszerítette, rákényszeríti a mindenkori felhasználókat arra, hogy minden hierarchikus relációt egyként kezeljenek e rendszeren belül. A rendszer elemei között pedig csak a hierarchikus láncok mentén lehetett kapcsolatot teremteni és keresgélni, ami könnyen merevvé tette az ilyen rendszereket. A tezauruszok megteremtésekor éppen e hátrányok leküzdése érdekében előre definiáltak több, pontosan meghatározott relációt, amelyek egyfelől a kapcsolatok sokkal pontosabb kifejezésére voltak alkalmasak, másfelől a rendszer elemei között több relációtípust is alkalmazva, többféle kapcsolati, bejárési útvonal alakulhatott ki, és ez jóval nagyobb rugalmasságot biztosított a tezauruszok számára.

A pontosabb szabályok szerinti működésmód igénye megkövetelte, hogy a tezauruszokra pontos és számon kérhető elvárásokat fogalmazzanak meg. Idővel a tezauruszok építésére vonatkozó szabályokat nemzetközi (és magyar) szabványokban rögzítették. A legfrissebb nemzetközi szabvány ISO 25964, amit két részletben publikáltak, 2011-ben és 2013-ban. A magyar szabványt 1987-ben adták ki (MSZ 3418-87 Magyar nyelvű információkereső tezauruszok szerkezete, részei és formái).

A tezauruszok tartalmas tételeit (vezér)deszkriptornak nevezik, amelyekhez a tezaurusz relációinak segítségével további tételeket lehet kapcsolni. A rendszerbe fel lehet venni nem-deszkriptorokat is, amelyek a szinonimitás kezelésére alkalmasak. A nem-deszkriptorok arra használhatók, hogy segítségükkel jelezzék, hogy a rendszeren belül van már olyan deszkriptor, amely be van kötve a tezaurusz struktúrájába, a nem-deszkriptornak így csak annyi a funkciója, hogy elvezessen a hozzá rendelt deszkriptorhoz, ahol további tartalmas továbblépési lehetőséget lehet találni.

A magyar tezaurusz szabványban az alábbi relációkat definiálták.

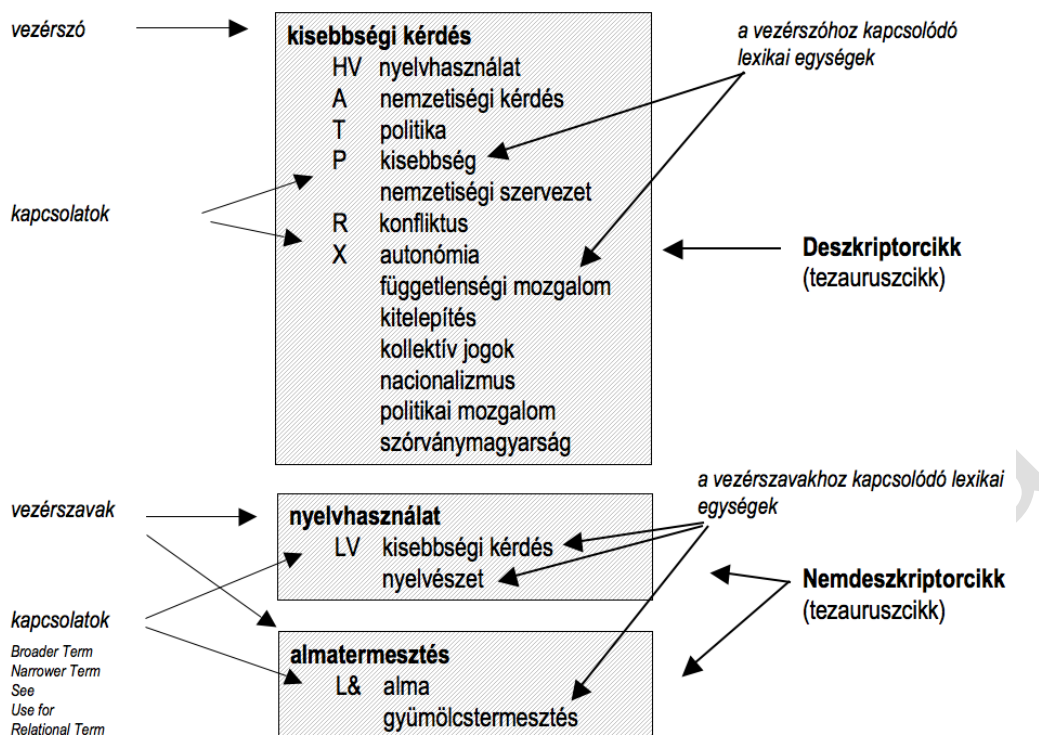
a magyar teauruszszabvány által hdefiniált relációk

H	szinonimája (lásd, helyette)	R ₉
F	fölérendeltje	R ₅
A	alárendeltje	R ₅
T	egésze (totum)	R ₆
P	része (pars)	R ₆
R	meghatározója (rezultáns)	R ₇
E	meghatározója (előfeltétel)	R ₇
X	rokonsági kapcsolat	R ₈

Ezeket a relációkat a következő módon határozzák meg a szabványban.

- H – szinonima reláció: a jelet követő nemdeszkriptor a deszkriptorcikk élén álló vezérdeszkriptort helyettesíti
- F/A – föl- és alárendeltségi (generikus, nem-faj) reláció: a két fogalom egymásból lineárisan származtatható (valódi részhalmaz)
- F – fölérendeltje reláció: a mögötte levő fogalom fölérendeltje a vezérdeszkriptornak
- A – alárendeltje reláció: a mögötte levő fogalom alárendeltje a vezérdeszkriptornak
- T/P – egész-rész (totum-pars; szerkezeti) reláció: a fizikai, szerkezeti kapcsolat meghatározása
- T – egésze reláció: a mögötte levő fogalom egésze a vezérdeszkriptornak
- P – része reláció : a mögötte levő fogalom része a vezérdeszkriptornak
- R/E – rezultánsa és előfeltétele reláció:
- R – rezultánsa reláció: „a vezérdeszkriptor által jelölt tárgy, folyamat, stb. rendeltetése, okozata, eredménye, terméke, célja, tárgya, következménye (együttvéve: meghatározottja) a jelet követő deszkriptor által jelölt folyamat, tárgy, stb. (együttvéve: meghatározó).”
- E – előfeltétele, kiindulása reláció: a vezérdeszkriptor által jelölt folyamat, tárgy, stb. létének, létrehozásának, működésének, meghatározásának oka, előidézője, eszköze, alapja, kialakulása, azaz meghatározója.
- X – rokonsági reláció: más módon ki nem fejezhető kapcsolat, ellentét, hasonlat vagy személynév.

A szabványban rögzített elvárások alapján a teauruszok teauruszcikkekből, deszkriptorokból állnak, és ezek a tételek sokféle módon kapcsolódhatnak egymáshoz. Egy adott tétel több komponense a vezérdeszkriptor, amihez a relációk mentén vannak hozzárendelve a további tételek (amelyek lehetnek deszkriptorok vagy nem-deszkriptorok).



tezaursz bejegyzés, magyarázó szövegekkel

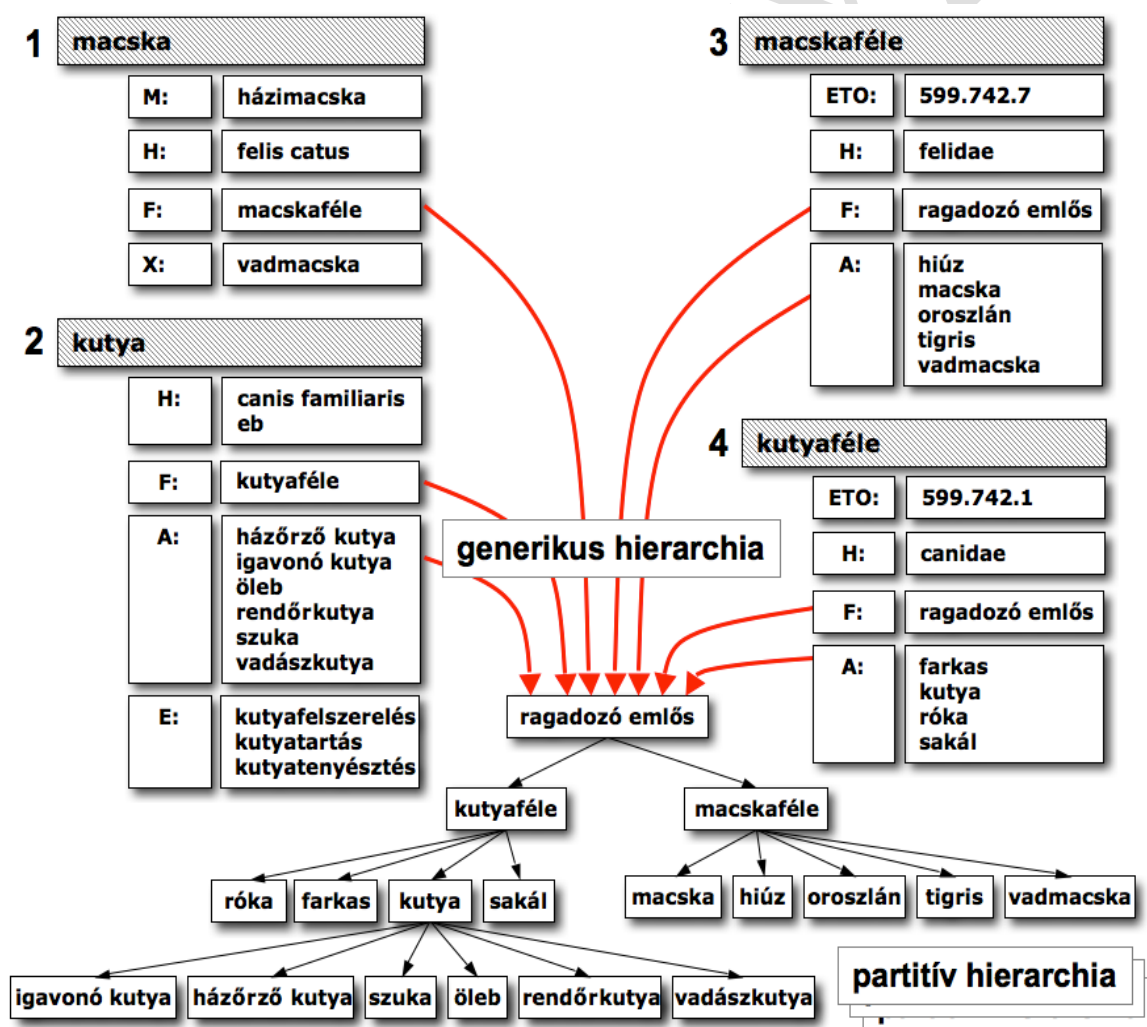
Ez az előny a kezdetektől arra a szerepre predesztinálta a tezauszt, hogy a könyvtári, archívumi világ legjobb, és ezért a legelterjedtebb tudásszervezési rendszerévé váljék, ez azonban nem valósult meg. A tezausznak az osztályozási rendszerekkel kellett viaskodnia, de a taxonómia-tezaursz csatát már szinte az első pillanatban az előbbi nyerte meg. A csata kimenetelét a könnyebb kezelhetőség döntötte el.⁷ Mindez persze felveti azt, hogy egy alaposabb tárgyalás során figyelniük kellene arra a szempontokra is, hogy milyen kötelezettségei vannak a katalogizálást, metaadat-hozzárendelést végző embernek.

Ha a taxonómiák építéséhez kontrollált szótárra, kompetenciafeltételek fenntartására, fegyelemre, a munkafolyamatok ellenőrzésére van szükség, akkor ez még inkább így van a tezaurszok esetében, hiszen ott jóval bonyolultabb struktúrát kell fenntartani, több szempontra kell figyelni, és nagyobb szaktudást kell elvárni a rendszert építőktől. Két – együttesen nehezen vagy egyáltalán nem teljesíthető – elvárás áll itt szemben egymással. Minél inkább szakterületi kérdéssről van szó, annál megbízhatóbbnak lehet tartani a szaktezaurszok (szaktaxonómiák) használatát az adott tudásterület leírásában, de annál inkább szükség van szakképzett, fegyelmezett, és ezért drága munkaerő alkalmazására is. Ha pedig valami sokba kerül, akkor mindig felmerül a kérdés, hogy ki fogja hosszútávon megfizetni. A kontrollált rendszerek hanyatlásának magyarázatában ez a döntő mozzanat: egyre inkább az látszik, hogy a web kontextusában egyre kevésbé hajlandóak pénzt áldozni erre. Annál is inkább nehéz megfizettetni a kontroll árát, mert az utóbbi években megjelent új jelenség sokak számára egy alternatív megoldás lehetőségét sejteti.

⁷ ETO-jelzet alapján nem csak a művet, hanem a példányt is meglehetősen találni: a szakkatalógusban kinézett 943.9"1848" : 323.1 jelzetű könyvért csak oda kellett menni a 9-essel kezdődő polcokhoz, az ETO jelzet magára a könyvre is rákerülhetett. Az ETO-nak az analóg világban ilyen gyakorlati haszna is volt.


Bár a teauruszok elbuktak mind a hagyományos világban a taxonómiákkal folytatott versenyben, mind a hálózati világ újonnan megjelent KOS-rendszereivel szemben, a teauruszok szemantikailag pontos és gazdag belső struktúrája megmaradt vonzó és követendő példát jelentő alternatívának, és ezzel részben előképét jelentette a számítógéppel kezelt formális ontológiáknak, és kezdetektől fogva táplálta a hozzájuk fűzött reményeket.

A tezusz pontos relációi mentén ugyanis elméletileg mindig adva volt a lehetőség arra, hogy logikai következtetéseket lehessen végezni e relációk segítségével. Az alábbi ábrán azt mutatjuk be, hogy az egyes teauruszcikkekben rögzített generikus relációs kapcsolatok arra voltak alkalmasak, hogy ezek mentén következtetéseket lehessen levonni a különböző tárgyszavakkal leírt dokumentumokra vonatkozóan (például ha egy könyvhöz a kutya terminus volt hozzárendelve, akkor a generikus reláció mentén lehetett ajánlani olyan könyveket is, amelyek a kutyafélékről, a ragadozó emlősökről szólnak – ha esetleg a kutyákról nem találtak megfelelő könyveket – azon az alapon, hogy minden kutya kutyaféle, valamint minden kutyaféle ragadozó emlős).



teauruszcikkek közti kapcsolatok a generikus reláció mentén

Az ilyen következtetéseket a hagyományos világban az ember mindig képes volt elvégezni, de a humán erőforrással kapcsolatos kapacitási gondok pont ezen a területen jelentkeztek a legerősebben. A géppel végzett következtetéseket pedig már a formális ontológiák segítségével próbálták, próbálják megoldani. A tezausz túl bonyolult volt ahhoz, hogy emberekkel hatékonyan kezelni lehessen, viszont még nem volt elég bonyolult ahhoz, hogy a gépek érdemben használni tudják.

 **cable-stayed bridges** (<bridges by construction>, bridges (built works), ... Built Environment (hierarchy name))

Note: Bridges having a deck that is directly supported from towers or pylons by straight cables without vertical suspenders. Construction of cable-stayed bridges usually follows the cantilever method. After the tower is built, one cable and a section of the deck are constructed in each direction. Each section of the deck is pre-stressed before continuing. The process is repeated until the deck sections meet in the middle, where they are connected. The ends are anchored at the abutments.

Terms:

cable-stayed bridges (preferred, C,U,English-P,D,U,PN)

cable-stayed bridge (C,U,English,AD,U,SN)

bridges, cable-stayed (C,U,LC,English,UF,U,N)

cable braced bridges (C,U,English,UF,U,N)

stayed bridges (C,U,English,UF,U,N)

cable-stayed girder bridges (C,U,English,UF,U,U)

stayed girder bridges (C,U,English,UF,U,N)

tuibruggen (C,U,Dutch-P,D,U,U)

tuibrug (C,U,Dutch,AD,U,U)

kabelbrug (C,U,Dutch,UF,U,U)

kabelbruggen (C,U,Dutch,UF,U,U)

Schrägseilbrücken (C,U,German,D,PN)












Schrägseilbrücke (C,U,German-P,AD,SN)

puentes de cables oblicuos (C,U,Spanish-P,D,U,PN)












punte de cables oblicuos (C,U,Spanish,AD,U,SN)

Facet/Hierarchy Code: V.RK

Hierarchical Position:

 Objects Facet
 Built Environment (hierarchy name) (G)
 Single Built Works (hierarchy name) (G)
 single built works (built environment) (G)
 <single built works by specific type> (G)
 <single built works by function> (G)
 transportation structures (G)
 <transportation structures by form> (G)
 bridges (built works) (G)
 <bridges by construction> (G)
 cable-stayed bridges (G)

Additional Parents:

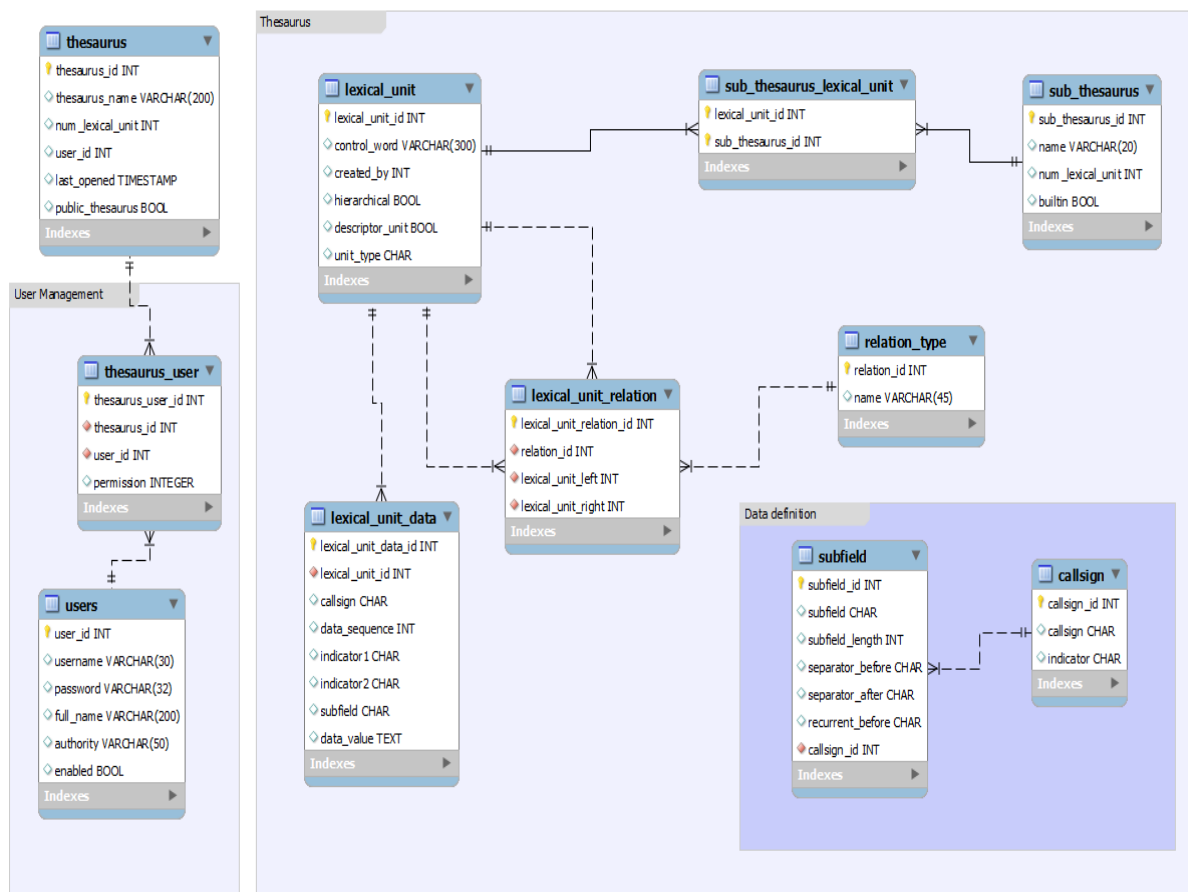
 Objects Facet
 Components (hierarchy name) (G)
 components (objects parts) (G)
 <components by specific context> (G)
 architectural elements (G)
 <structural elements and components for structural elements> (G)
 structural elements (G)
 structural systems (G)
 cable structures (G)
 cable-stayed structures (G)
 cable-stayed bridges (G)

Additional Notes:

German Brücke mit einer Tragfläche, die mittels Kabel ohne vertikale Unterstützung direkt von Türmen oder Pylonen getragen wird. Der Bau einer Schrägseilbrücke erfolgt meist im Freivorbauverfahren. Nach Errichtung des Turms werden in jede Richtung jeweils ein Kabel und ein Segment der Tragfläche gebaut. Jeder Teil der Tragfläche wird vorgespannt, bevor mit dem weiteren Bau fortgefahren wird. Der Prozess wird wiederholt, bis sich die Tragflächenteile in der Mitte treffen, wo sie verbunden werden. Die Enden werden in den Stützpfählern verankert.

Getty Art & Architecture Thesaurus Online: 'cable-stayed bridges' tétel

A tezasauruszokat nyilván sokféleképp lehet implementálni, de nagy különbség nem lehet köztük, hiszen a tezasauruszokat definiáló relációk maguk olyan kényszereket jelentenek a megvalósítás



a Relex adatmodellje

A sémák áttekintése alapján látható, hogy az adatmodell központjában a tezaszuscikkek magját jelentő relációs tábla áll, amelyben a vezéreszkriptor és a különböző relációk mentén hozzá kapcsolható lexikai egységek kapcsolatát rögzítik.

Mindig is adta magát, hogy a tezaszuszban rejltó gazdagabb struktúra valahogyan hasznosítani kellene az egységek és kapcsolataik megjelenítése során. Ezt az elvi lehetőséget, és régi kívánalmat akkor lehetett megvalósítani, amikor a tezaszuszokat is számítógéppel lehet kezelni, és a grafikus felületeken már el lehetett szakadni a hagyományos íráskultúra vizuális kereteitől. Részben ilyen „elszakadási” kísérletként is értékelhetjük az olyan kezdeményezéseket, amelyek a tezaszuszok egységeit, struktúráját a térben szabadabb módon, a lineáris írástechnikák korlátaitól elszakadva akarták megjeleníteni. Az egyik legrégebbi és legsikeresebb ilyen kísérlet a vizuális tezaszusz, amelynek felületén a kétdimenziós felületet szabadon és teljesebb módon kihasználva lehet feltüntetni a kiválasztott tételeket és a hozzájuk tartozó további egységeket. Ez a megoldás természetesen csak a felhasználói élményben okoz változást, a rendszer belső felépítése, működése ugyanolyan marad.

R_8 rokona (egyéb) reláció

R_9 lásd/helyette szinonimareláció-pár

A teaurusz esetében az S_i szabályok mindegyikét be kell tartani (igaz, hogy az S_2 -es szabályt nem mindig, sőt gyakrabban nem érvényesítik, vagyis nem engedik, hogy új relációt lehessen felvenni a rendszerbe, de a formális modellbe mégis be kell emelnünk ezt a szabályt, mert előfordulhatnak olyan teauruszok is, amelyekben a szabványokban rögzített relációkhoz képest további relációt is definiálnak). Vannak azonban a formulában jelzett relációk, relációpárok, amelyek más minőséget adnak a teauruszoknak. Mivel a teaurusz több, pontosan definiált relációt tartalmaz a taxonómiához képest, ezért összetettebb struktúrát képezhetünk le vele, és a gazdagabb szemantika, a nagyobb kifejezőerő miatt sokkal pontosabban, rugalmasabban és megbízhatóbban lehet vele a dokumentumok tartalmát leírni.⁸

3.3.5 Folkszonómia

Alvin Toffler 1980-ban írt könyvében már előrevetítette, hogy a termelés és fogyasztás hosszú időszakon keresztül egymástól elváló társadalmi szerepei újra közelebb kerülnek egymáshoz. A jelenség fontosságát azzal is hangsúlyozni akarta Toffler, hogy új terminust alkotott: a *prosumer* kifejezést [Toffler 2001]. Toffler nem látta, nem láthatta előre a hálózati kommunikáció web2-es trendjeit, de ettől még tény, hogy ezek a fejlemények messzemenően igazolták a toffleri jóslatot. Persze, ezt a jelenséget – a web2.0 kifejezésen túl – nagyon sokféle terminussal próbálták meg leírni, Toffler *prosumer* kategóriája csak egy a sok közül. A felhasználók által létrehozott tartalom (user generated content), *producer*, *social media*, *social web*, *peer production*, *crowdsourcing*, *collective intelligence*, *wikinómia*, *a tömegek bölcsessége* – sokáig lehetne sorolni a versengő kifejezéseket. Ha a navigáció támogatására szolgáló tudásszervezési rendszereket vizsgáljuk, akkor a megfelelő web2-es kategória a *folkszonómia* (folksonomy) fogalma.

Maga a 'folkszonómia' terminus Thomas Vander Wal egyik blogbejegyzésében jelent meg először 2004-ben (azóta kétszer, két helyen is megpróbálkozott a fogalom definiálásával (Vander Wal 2007, 2005b), de érdemes megemlíteni azt a tényt, hogy az *etnoklasszifikáció* kifejezéssel Susan Leigh Star már 1996-ban nagyon hasonló értelmű fogalmat hozott létre (Star 1996). Ezeket a rendszereket gyakran *címkézési rendszerekként* (tagging system, tagsonomy), olykor *közösségi címkézési rendszerként* is emlegetik, magát a metaadat-kezelési tevékenységet pedig címkézésnek hívják, de használnak még további neveket is, mint például *grassroots classification*, *cooperative classification* vagy *distributed classification system*, *folk classification*, *folk taxonomy*, *social classification*, *open tagging*, *free tagging*, *faceted hierarchy*.

⁸ A teaurusz R_5 , R_6 , R_7 relációi egyenként" mind hierarchikus struktúrát biztosítanak, ezért ezek úgy tekinthetőek, mint amelyek egy-egy alosztályát adják a taxonómiák R_3 'hierarchikus alárendeltje' relációjának. Ez azt (is) jelenti, hogy ha e három teauruszreláció mentén összekapcsolt elemeket egyetlen – általánosabb – relációval kötnénk össze, akkor az R_3 relációval képezhető struktúrához nagyon hasonló képződményt kapnánk. A teauruszban tehát egy „rejtett” taxonómia is található.

Az új terminust Vander Wal a 'folk' és 'taxonomy' szavak összevonásából gyártotta. Azt akarta kifejezni vele, hogy a tárgyszavazási tevékenységet nem csak az erre felhatalmazott, kontrollált tárgyszórendszert kezelő szakember, hanem bárki (a „nép” bármely tagja) végezheti. Tartalmilag nem volt szerencsés erre a taxonómia kifejezést ráhúzni, mert a folkszonómiák nem hogy egy osztályozási rendszert, de semmilyen más struktúrát nem képesek felépíteni, fenntartani. Ennek pedig pont az az oka, amit az új kifejezés előtagja fejez ki, hogy ti. ha bárki felvehet új címkét, bárki bármilyen címkét a dokumentumokhoz rendelhet, akkor semmi sem garantálja az , hogy a címkék között konzisztens kapcsolat alakulhasson ki (és maradjon fenn).

A folkszonómia nagyon lapos struktúrájának tekinthető, csak abban különbözik az egyszerű terminuslistáktól, hogy a címkék gyakorisági értékeire lehet relevanciakezelést építeni. A folkszonómiák alapján épített címkefelhők a címkék méreteivel, forgalmukkal kapcsolatos számosságával jelzik azt, hogy a rendszeren belül melyik a közösség által fontosabbnak ítélt (mert gyakrabban használt) címkék. Ez a gyakoriságjelzés mint relevanciakezelő kritérium sok esetben valóban használható, de olyan is előfordul, nem is kevés, hogy ez a megoldás gyakorlatilag használhatatlan. Erre később mutatunk példákat.

A folkszonómiák hatalmas előnye a rendszer könnyű kezelhetősége. Nem követel meg semmilyen feltételt a címkét kiosztó felhasználótól a rendszer a címkék használatakor, szabadon lehet új címkét felvenni, régit használni. Ebből persze többféle probléma is előállhat. Ha nincs semmi szabály, akkor ugyanazt a tárgyszót fel lehet venni egyes számban, de többes számban (lány, lányok, vagy ugyanazt a jelentésű tartalmat fel lehet venni más szófajú (főnévi, melléknévi vagy igei) alakban (futok, futás, futó), vagy alkalmazni lehet szinonim jelentésű tárgyszavakat (kutya, eb), és ezek mind külön címkéknek fognak megfelelni. A folkszonómiák védelmében szokták felhozni, hogy a hibás, „nem helyes” alakú címkék kisebb számban fordulnak majd elő, és a gyakoriságuk alapján majd nem lesznek láthatók, így nem is fognak zavarni. Ez azonban nem feltétlen igaz. Az Instagram hashtag-jei kapcsán a későbbiekben bemutatjuk azt, hogy ez effajta szűrés nem feltétlen működik minden esetben eredményesen.

A folkszonómiák esetében további komoly problémáját jelenti a *kontextuskötöttség*, ami alatt azt a jellegzetességet értjük, hogy a felhasználók gyakran olyan címkéket rendelnek a dokumentumokhoz, amelyeknek csak a saját kontextusokban van igazi referenciája, és annak ismerete nélkül semmilyen értelemben nem ad tartalmi eligazítást a címke. A klasszikus példája ennek az, amikor a felhasználók a 'me' címkét ragasztják a – mondjuk – az általuk feltöltött képekhez, de ugyanilyen semmitmondó a 'selfie' (és még sok minden más, hasonló) címke is.

A jelenség első híres példája a del.icio.us webcímen működő közösségi címkézési rendszer volt. Azóta a kép- és videómegosztók, a mikroblogok és a közösségi oldalak is ezt a technikát használják: a Twitter vagy a Facebook hashtagjei, a Flickr, Instagram vagy a YouTube címkézési gyakorlata is folkszonómiaként minősíthetők.



a Twitter legnépszerűbb hashtagjei adott időpontban

A Twitter rövid tartalmai alapján nehezen lehetne összekapcsolni az elküldött posztotokat, ráadásul a Twitteren küldött képek esetében semmilyen kereshető, linkelésre alkalmas metaadat sem áll rendelkezésre, ezért a Twitter platformja számára kiváló eszköz a hashtagek használata, mert ez ad esélyt a linkelésre.



a Twitter személyes hashtag-ajánlata

Persze, a hashtagelés nem mindig, nem feltétlenül nyújt segítséget az eligazodásban, kapcsolódásban. A bal oldalon látható az a címkeajánlat, amit a Twitter kínált fel egy olyan felhasználónak, aki hosszú évek inaktivitása után belépett újra a szolgáltatás felületére. A 'Budapest' címkén kívül semmilyen értékes, értékelhető címkét nem láthatunk, a többségüknél talán még a véletlenszerű kiosztás is több relevanciát nyújthat.

Az ajánlott címkék listájában jól látszik a folkszonómiák kontextuskötöttsége, felszínes tartalmassága. A '#mustread' címke csak annyit képes üzeni nekünk, hogy fontos, de persze ez is csak akkor lenne számunkra, ha a címkét kiosztó személy és köztünk valamilyen értékközösség lenne, aminek jelzésére ez a hashtag nem alkalmas. A kontextuskötöttség mint hátrány a személyes névmások ('me', 'my mom'), a személynevek ('Gabriel') gyakori alkalmazásában érhető tetten, amikor tudjuk, hogy a címke referenciája csak a címkéző kontextusában határozható meg pontosan, annak ismerete nélkül a címke semmire sem mutat.

A hashtagek tartalmatlanságát szemléltethetjük még az Instagram képmegosztó száz legnépszerűbb címkéjével. A listában vannak olyan címkék, amelyek bírnak némi tartalmi eligazító erővel (dog, food, makeup, sky, flower, cat, gym), de ezek vannak kevesebben. A kontroll hiánya megmutatkozik abban is, hogy nem lehetséges a szintaktikai egységesítés sem. A címkék felbukkanhatnak egyes és többes számban vagy szinonim alakban, és ilyen esetekben különböző címkéknek számítanak, példa lehet erre a girl (17.) és a girls (59.) vagy a picoftheday (13.) és a photooftheday (3.) címke is. Sok olyan címke van a listában, amely teljes mértékben kontextusfüggő, érdemi tartalommal nem rendelkezik, például a my (85.), a me (7.), a family (26.), a selfie (12.), followme (10.) vagy a baby (63.). A népszerű címkék nagy része pedig valamilyen állapotot vagy értékelést jelöl, ami megint csak kevés eligazító erővel bír, például a fun (18), a funny(66.), beautiful (4.), happy (8.) vagy a love (1.). Vannak olyan címkék is, amelyek valamilyen divatjelenség miatt kerülhettek be a topszázba, vagy egy közösség szlengjeként funkcionálhatnak, és a kontextus ismerete nélkül nem is lehet tudni a jelentésüket: f4f (46.), lfl (48.), tbt (4.), vsco (50.), ootd (62.).

100. #birthday	75. #black	50. #vsco	25. #art
99. #red	74. #gym	49. #dog	24. #food
98. #workout	73. #motivation	48. #lfl	23. #instalike
97. #sweet	72. #yummy	47. #pretty	22. #igers
96. #wedding	71. #healthy	46. #f4f	21. #repost
95. #blackandwhite	70. #instacool	45. #travel	20. #smile
94. #instalove	69. #hot	44. #sky	19. #tagforlikes
93. #fit	68. #iphoneonly	43. #music	18. #fun
92. #instafood	67. #instapic	42. #swag	17. #girl
91. #instasize	66. #funny	41. #followforfollow	16. #instadaily
90. #flowers	65. #instagram	40. #beach	15. #friends
89. #iphonesia	64. #night	39. #beauty	14. #summer
88. #tweetgram	63. #baby	38. #sun	13. #picoftheday
87. #design	62. #ootd	37. #vscocam	12. #selfie
86. #instafollow	61. #makeup	36. #bestoftheday	11. #follow
85. #my	60. #cat	35. #fitness	10. #followme
84. #webstagram	59. #girls	34. #life	9. #fashion
83. #followback	58. #cool	33. #amazing	8. #happy
82. #christmas	57. #lol	32. #follow4follow	7. #me
81. #work	56. #party	31. #nofilter	6. #cute
80. #blue	55. #foodporn	30. #style	5. #beautiful
79. #instacool	54. #tflers	29. #instagood	4. #tbt

78. #work	53. #hair	28. #nature	3. #photooftheday
77. #blue	52. #photo	27. #likeforlike	2. #instagood
76. #pink	51. #sunset	26. #family	1.#love

az Instagram 100 legnépszerűbb hashtagje (2018)

Bár azt egyáltalán nem mondhatjuk, hogy a folkszonómiák tudatos tervezés eredményeként jöttek volna létre, kialakulásukat és elterjedésüket mégis a taxonómiákkal való összehasonlítás révén érthetjük meg. Ahogy azt az osztályozási rendszerek leírásakor jeleztük, hamar kiderült, hogy hálózati környezetben ezek a tudásszervezési rendszerek sikertelennek bizonyultak, mert rugalmatlanok voltak, kiépítésük és használatuk nehéznek és költségesnek bizonyult. Jóval rugalmasabb és egyszerűbb, a felhasználók számára elfogadhatóbb, használhatóbb megoldásra volt szükség. Ezt hozták el a folkszonómiák, amelyeket épp ezért minősíthetünk felhasználóbarát tudásszervezési rendszereknek is.

A folkszonómiák használatakor ahelyett, hogy a dokumentumokat egy hierarchikus struktúrájú osztályozási rendszer elemeivel próbálnánk meg leírni, egyszerűen címkéket ragasztunk rájuk úgy, hogy a címkék formátumára, jelentésére semmilyen megkötés nem teszünk. A címke kifejezheti a dokumentum jellemző tulajdonságát ('film', 'magyar', 'dokumentumfilm'), lehet emlékeztetőként használni ('megnézni', 'fontos', 'tanulmányhoz'), vagy utalhat akár a dokumentumhoz fűződő érzelmekre is ('tetszett', 'izgalmas', 'hatásvadász').

Mivel nincs semmilyen központi előírás, szabályozás, ezért a dokumentumok címkézését maguk a felhasználók végezhetik. Emiatt persze a címkézés erősen szubjektív lesz. Ez az elfogultság természetesen a címkézést végző felhasználók gondolkodásmódját tükrözi. A szubjektivitás hatása csökkenthető, ha az archívum felhasználóit érdekeltté tesszük abban, hogy a már mások által felcímkézett dokumentumokat maguk is ellássák új címkékkel (Speroni 2006). Minél több felhasználó címkézi a tartalmakat, annál több és relevánsabb címke fog rákerülni egy dokumentumra, megkönnyítve annak megtalálását. Ha ez sikerül, akkor a sokak által címkézett dokumentumokon nagy eséllyel lesz néhány olyan, amelyet kiugróan sokan ragasztottak rájuk. Ezek lesznek a dokumentumra valamilyen szempontból jellemző tulajdonságok. És persze akad majd nagyon sok olyan címke, amelyet csak a címkézők töredéke ragasztott fel. Ezek lesznek azok a tulajdonságok, amelyek csak kevesek számára fontosak, illetve ebbe a csoportba kerülnek a személyes érzelmek, emlékeztetők is. Adam Mathes már az első folkszonómiák megjelenésekor megállapította, hogy a címkék eloszlása hatványfüggvény jelleget követ (Mathes 2004). A címkék kisebbik hányadát nagyon sokszor rendelik hozzá a dokumentumokhoz, míg a címkék túlnyomó részét jóval ritkábban (legtöbb esetben nagyon kevés alkalommal) használják.

Mindkét fajta címke nagyon jól használható kereséskor. Az *erős címkék* a keresési eredmények rangsorolásakor nyújtanak nagy segítséget. Például egy sportkocsikat bemutató film esetén jellemző lehet az 'autó', és a 'film' címke. Ha ezen kulcsszavak alapján keresnénk dokumentumokat, akkor szeretnénk, ha ez a film a találatok között előkelő helyet foglalna el, és ehhez hatalmas segítséget jelent az erős címkék jelenléte. A *gyenge címkék*, amelyek az ún. „hosszú farok” részről származnak (Vander Wal 2005a), biztosíthatják azt, hogy az adott dokumentum a legváltozatosabb keresési feltételek esetén, különféle gondolkodásmódokhoz

igazodóan is, megtalálható legyen. A fenti sportkocsis film példájánál maradva ilyen címke lehet például a 'nyolchengeres', 'izgalmas', 'sebességváltó'.

A dokumentumokhoz való hozzáférés is a címkék közti kulcsszavas keresésen alapul: a rendszer azokat a tartalmakat fogja megjeleníteni, amelyek a keresett címkét viselik. Kereshetünk az összes dokumentum, illetve csak az általunk megjelöltek között, így amilyen könnyen megtalálunk egy 'ház'-at ábrázoló képet, amit sokan 'szép'-nek tartottak, olyan könnyen visszakereshető a mi 'ház'-unkról készült kép, vagy egy olyan, amit mi gondoltunk 'szép'-nek. Ezért a magunk számára érdemes megjelölnünk egy dokumentumot akár egy olyan, mások számára semmitmondó címkével is, mint például az 'elintézni való'. A címkézés azért működik, mert – a taxonómiába való besorolással szemben – a végletekig leegyszerűsített, intuitív művelet, és ezért a felhasználók hajlandóak használni. Sőt, nem csak hogy hajlandóak rá, de motiváltak is, hogy maguknak is felcímkézzék a számukra fontos dokumentumokat – hiszen azok később a saját fogalmi rendszerüket reprezentáló címkéken keresztül sokkal könnyebben megtalálhatóak lesznek.

A HP Lab két kutatója a Del.icio.us címkézési gyakorlatának adatait elemezve több szempontból is jellemezni próbálta a felhasználók aktivitását (Golder & Huberman 2006). Az elemzés során tipizálták a címkéket is, és az alábbi csoportokat (pontosabban címkézési funkciókat) különítették el egymástól.

- azonosítani, kiről vagy miről van szó (köznevek vagy tulajdonnevek, amelyek általános és egyes, illetve absztrakt és konkrét dolgokat, objektumokat, eseményeket stb. jelölnek: 'dog', 'computer')
- azonosítani, milyen fajtáról, típusról van szó (köznevek, melyek valamilyen szinten általánosítva jelzik azt, hogy milyen dolgokról van szó: 'article', 'blog', 'book' – általában az előző címketípus megadásával párhuzamosan)
- azonosítani, kinek a címkéjéről van szó (ki „birtokolja” a címkét: 'Peter')
- pontosítani a kategóriákat (már megadott címkék terjedelmét, pontosságát jelzőkkel, számokkal, értékekkel pontosítani: '12,500')
- értékelni a dolgokat (jelzőkkel kifejezni a felhasználó véleményét: 'scary', 'funny', 'stupid')
- önhivatkozás (kifejezni a személyes kapcsolatot: 'mystuff', 'mycomment')
- teendőket szervezni (utalás elvégzendő feladatokra: 'todo', 'toread')

Ezek a címketípusok abban az értelemben nem összevethetők, hogy különböző célokra lehet/érdemes őket használni, más tehát a funkciójuk. Később kitérünk a folkszonómiák gyakorlatával szemben felvethető kifogásokra is, előbb azonban nézzük meg a címkézési rendszerek előnyeit. A címkézés ugyanis számos előnnyel jár. A hagyományos taxonómia alapú archívumok esetén két dokumentumot akkor mondhatunk hasonlóknak, ha nagyjából egy kategóriába esnek a hierarchikus rendszerben. A hasonlóság fogalmát folkszonómia alapú archívum esetén sokkal természetesebben, ráadásul több szinten is értelmezhetjük. Címkék használata esetén két dokumentum akkor lesz hasonló, ha a rájuk ragasztott címkék között sok az átfedés. A hasonlóság természete ebben az esetben tehát nem a taxonómia megalkotójának fogalmi rendszere mentén nyilvánul meg, hanem azt a közösség által felállított címkék fogják meghatározni. Egy hagyományos fotóarchívumban ritkán kerülne egymás mellé egy piros sportkocsit és egy piros almát ábrázoló kép. Ez folkszonómia esetén teljesen természetes fogalmi társítás a 'piros' címkén keresztül. A hasonlóság ráadásul tovább általánosítható, hiszen nemcsak a dokumentumok között definiálhatjuk azt, hanem feltérképezhetők a gyakran együtt

előforduló címkék is. Ezzel a rendszer felismerheti a szinonimákat, vagy az egy fogalomkörből származó szavakat, például az archívum valamiképp 'tudni' fogja, hogy a 'piros' és a 'vörös' címke hasonló dolgokat ír le. Ennek segítségével a rendszer megtalálja a valamiképpen kapcsolódó tartalmakat. És ezzel még mindig nincs vége, hiszen az archívumból az is kideríthető, hogy kik azok a felhasználók, akik hasonló címkékkal láttak el dokumentumokat. Ők, szintén a hasonló fogalmi kör használata miatt valamiképp ugyanúgy gondolkodnak, azaz hasonlítanak egymáshoz. A 'hasonló' felhasználók könnyen megtalálhatják egymást, így nem csak a közösség építi a rendszert, hanem a rendszer is a közösséget. A címkézés segíthet megoldani az eltérő nyelvi környezetben élő felhasználók eltérő nyelvhasználatának problémáját is. A legnépszerűbb címkéket egyszerre megmutató címkefelhő a folkszonómiák egyik kedvenc szolgáltatása. Mivel a címkefelhő eligazító, navigációs ereje akkor „értékesül”, amikor a felhasználók ránéznek, könnyen szelektálni tudják a nem az általuk ismert nyelvből származó címkéket, ahogy ezt az alábbi ábrán láthatjuk.



többsz nyelvű elemek egy címkefelhőben

Az ábrán az angol címkék mellett szerepel néhány más (talán kínai) írásjel is, amelyeken az angolul olvasó felhasználó szeme könnyedén „átugrik”, és nem veszi figyelembe, illetve ugyanez igaz a kínai felhasználókra is, csak fordítva.

A címkék, a címkézés rugalmassága nagy előny. Ennek szemléltetésére hasonlítsunk össze a fotozz.hu szolgáltatását a Flickr.com által kínált lehetőségekkel. Mindkét rendszer képek feltöltését, megosztását, online fényképalbum létrehozását, valamint mások albumainak, képeinek böngészését, kereshetőségét biztosítja. A fotozz.hu szolgáltatás egy új kép feltöltésekor megköveteli, hogy a képet soroljuk be a megadott kategóriák egyikébe. A kategóriarendszer rögzített, nem lehet változtatni, bővíteni, és erősen fotószakmai elfogultságú kifejezésekből áll.

a fotozz.hu oldal legfontosabb kategóriái

absztrakt	digit. illusztráció	légifotók	tárgyfotók
abszurd	divat	makró	természet
akt	életképek	panoráma	város, építészet
állatfotók	elkapott pillanatok	portré	vízalatti fotók
barlangfotók	glamour	riport	feldolgozott fotók
családi/emlékkép	humor	sport	így készült
csendélet	infravörös fotók	szociofoto	egyéb
csillagászat/égbolt	koncert - színpad	tájkép	pályázat

Lehet, hogy ez a kategóriarendszer fotószakmai szempontból teljes és kimerítő, de ezzel biztosan nem lehet a képeket kereső felhasználók szempontjait kielégíteni. Olyan keresési igény

nehezen képzelhető el, amikor valaki 'pályázat' vagy 'így készült', netán 'egyéb' témakörben szeretne képeket találni magának. Ez a rendszer túl merev. A kép mellé a technikai metaadatokon kívül nem lehet mást felvenni, így kereséskor egyedül a képek kategóriáját lehet kiválasztani, majd arra bizonyos szűkítési feltételeket megadni, de akkor sem tartalmi, hanem egyéb szempontok szerint lehet csak szűrni (a feltöltés ideje vagy a képek értékelései alapján). Az értékelés mindig segíthet a relevánsabb találati listák létrehozásában, viszont a képek tematikus kereséséhez semmit sem képes hozzátenni. Ha kutyákról keresnénk képeket, akkor választhatjuk az 'állatfotók' kategóriát, de ezután már nem tudunk további szűkítést elérni, ezért ha 'briard' kutyafajtáról szeretnénk képeket, végig kellene lapoznunk az összes állatfotót, hogy találhassunk köztük nekünk megfelelő képet.

The screenshot shows the 'fotózz!' website interface. At the top, there are navigation links for 'digitális fényképezőgépek tesztek', 'objektív adatbázis', and 'photoshop tippek-trükkök'. The main header features the 'fotózz!' logo and user statistics: 46,826 felhasználó, 157,063 aktív fénykép, 1,749,948 értékelés, and a date of 2003. október 1. óta. Below the header, there's a search bar with 'syt' and 'KATEGÓRIA: abszurd'. The main content area displays a photo of a bulldog wearing a bow tie, with a 'KÖZEPES EREDETI' label. To the right of the photo, there's a section titled 'válassz géptípust!' with details: 'Feltöltve: 2009. március 18. 21:03', 'Feldolgozható: nem', and 'Pontozható: igen'. Below this, a message reads: 'limóra várva aki az élességet keresi, annak üzenem, hogy a kutyus segglyukán van. köszl.'. On the right side, there's a 'PMA09 HELYSZÍNI TUDÓSÍTÁS LAS VEGASBÓL' section with a list of items like '2009-es PMA: zárzó', 'HP újdonságok', etc. At the bottom, there's a 'FOTÓ ÉRTÉKELÉSE' section with a form for rating and a 'Kredit átutalása:' field.

keresés a 'briard'-ra a fotozz.hu oldalon

A flickr.com másként működik. Az oldal alapvetően ugyanazokat a funkciókat kínálja, mint a fotozz.hu, de itt a kép feltöltője nem egy merev kategóriarendszerbe sorolja be a képét, hanem szabadon, saját kulcsszavaival írhatja le azt. Ezáltal természetesen nagyságrendekkel több címkét rendelnek a képekhez (a fotozz.hu esetében láthattuk, hogy mindössze 32 kategória alá lehetett a képeket besorolni). A szabad és „bőséges” címkézésnek köszönhetően a keresési lehetőségek is jobbak lesznek. Szinte „mindenre” kereshetünk, mert a Flickr már elérte azt a kritikus méretet, amikor már a legextrémebb címkéket is használta valaki. Természetesen a 'briard' címke alapján is szép számmal kapunk találatokat.

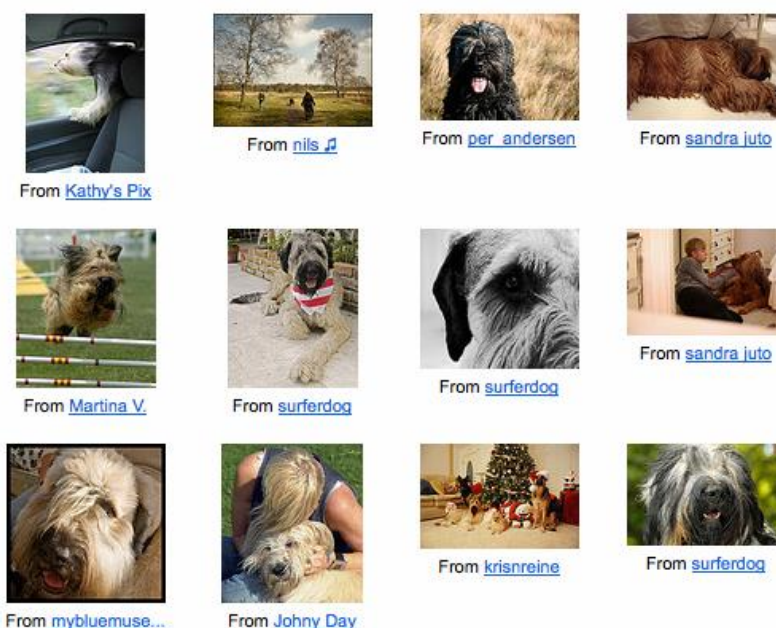
flickr®
 Home The Tour Sign Up Explore ▾

Search Photos Groups People

briard dog

Full text Tags only

✔ We found **1,299 results** matching **briard** and **dog**.
 View: [Most relevant](#) • [Most recent](#) • [Most interesting](#) Show: [Details](#) • [Thumbnails](#)



keresés a 'briard'-ra a flickr.com oldalon

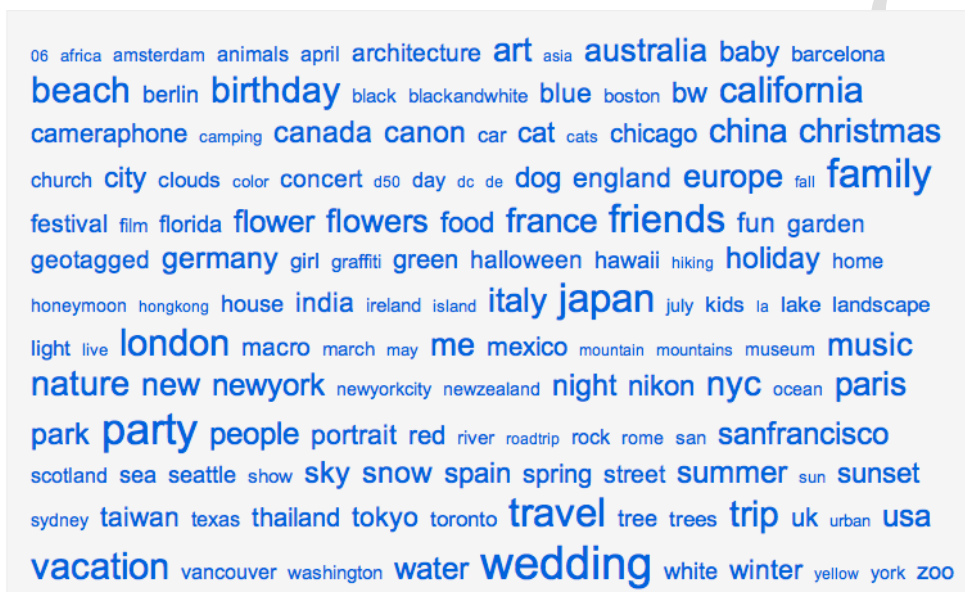
A címkézés tehát jóval rugalmasabb, mint a hagyományos osztályozási rendszer, a keresést sokkal jobban tudja támogatni. Sőt ezen felül egyfajta személyes „emlékezettechnikát” is biztosít a felhasználók számára. Említettük már, hogy a használt címkék egy része a kép tartalmát jelöli ('family', 'party', 'wedding', 'birds', 'dog', 'cat'), de a címkék utalhatnak a készítés módjára ('cameraphone'), idejére ('2006'), stílusára ('art'), helyszínére ('France') vagy akár egészen szubjektív dolgokra is ('cute', 'cool'). Vannak azonban olyan címkék is, amelyek annyira „személyesek”, hogy azok senki másnak nem hordoznak érdemi információt a címkézőn kívül ('friends', 'me', 'todo', 'toread'). Ezekkel a címkékkal nyilván senki sem akar keresni, hiszen ezek csak a feltöltő személyének ismeretében lennének értékelhetők. Ezt a kontextust csak maga a feltöltő személy ismeri, neki viszont értékesek lehetnek az ilyen címkék, amennyiben a célja nem a keresési lehetőség biztosítása bárki számára, hanem a kép személyes célú megjelölése a saját fotóalbumon belül.

A leggyakrabban használt címkék megjelenítését *címkefelhőnek* nevezik, amelyekben a címkék betűméretével, vagy egy számértékkel jelzik a címkék népszerűségét, azok gyakorisági értékeit. Már csak azért is érdemes kicsit alaposabban megvizsgálnunk ezt a jelenséget, mert közben bemutatathatjuk a folkszónómiákkal kapcsolatban felmerülő problémákat is. Könnyen találhatunk

ugyanis olyan példákat, amelyek azt mutatják, hogy a képzetlen és fegyelmezetlen felhasználók nehezen tudnak megbirkózni (ha meg tudnak egyáltalán) a tárgyszavazás olyan – régóta ismert – szemantikai alapproblémáival, mint:

- szinonimakezelés
- poliszémiakezelés
- homonímiakezelés
- granularitáskezelés

Kiinduló, szemléltető példaként vegyük a Flickr egy korábbi időszakból származó címkefelhőjét:



Flickr címkefelhő

Ha figyelmesen nézzük a fenti ábrát, a folkszónómiák több komoly problémáját is észre vehetjük rajta: a kontroll hiánya többféleképpen tetten érhető.

Több olyan címkepár van, amely ugyanazon fogalom egyes- és többesszámú alakjára vonatkozik ('cat' és 'cats', 'flower' és 'flowers', 'tree' és 'trees'), amelyeket a kontrollált szótár használatával természetesen ki lehetne szűrni. Az egyes- és többesszám kevert használatának problémáját általánosíthatjuk, és azt mondhatjuk, hogy a folkszónómiák nem képesek kezelni a morfológiai többalakúság egészének kérdéskörét. A nyelvi megnyilatkozásaink megformálásakor ugyanazt a szó-szintű nyelvi egységet (szót, kifejezést) sokféle morfológiai alakban használjuk a mondatok képzése során, ám ez a morfológiai többalakúság nem jelent szemantikai többértelműséget. Egy szó egyes- vagy többesszámú alakja – első megközelítésben – ugyanazt a jelentést hordozza csakúgy, mint egy szótó és annak bármelyik más, toldalékolt alakja ('ló' vagy 'lovat', 'lónak', 'lóval', 'lovon'). Ha viszont ezek az eltérő morfológiai alakok címkékké válnak (és nyugodtan válhatnak), akkor ezek mind különböző címkék lesznek anélkül, hogy bármilyen szemantikai különbséget ki tudnának fejezni. Ez a jelenség nyilván rontja a keresés hatékonyságát.

Persze, meg lehet kérdezni, hogy a felhasználók miért lennének olyan „buták”, hogy a 'ló' címke helyett a 'lovon' címkét adnák meg egy kép leírásakor. Nem arról van szó, hogy a felhasználók tényleg buták lennének (bár néha lehetnek azok is), hanem inkább arról, hogy a folkszónómiák sajátos működésmódja eredményezi ezt a nem kívánt jelenséget. Ezek a rendszerek a címkéket automatikusan detektálják úgy, hogy két szóköz közti karaktersorozatot tekintenek egy

címkének. Ebből következően a több szavas kifejezéseket „feldarabolják”, amit az emberi feldolgozás nyilván nem tenne meg. Példa lehet erre a 'New York' vagy a 'black & white' tárgy szó. Ezek ebben a formájukban nem szerepelnek (nem is szerepelhetnek) a rendszerben, hiszen több szóból álló „kifejezések”. A kifejezések tagjai, a 'new' és a 'york', illetve a 'black' és a 'white' címkék viszont előfordulnak külön-külön és egybeírva is: 'newyork' és 'newyorkcity', valamint 'bw' és 'blackandwhite'. Mivel a folkszonómiák az ún. „single-word” megoldást alkalmazzák a címkék behatárolására, elkülönítésére, a felhasználók – ezt tudva – az előző „egybeírós” technikát használják a több szóból álló kifejezések bevitelére.⁹ Persze, nem mindenki alkalmazza ezt a megoldást, és emiatt megint csak pontatlanabbá válik a rendszer, amire a fenti címkefelhőben az utal, hogy a 'sanfrancisco' címke mellett ott szerepel a 'san' is, ami részben annak köszönhető, hogy sokan a 'San Francisco' kifejezést két szóval írták le, ezt pedig a rendszer felbontotta (ebbe persze a többi 'San' előtagú városnév (mint 'San Diego') is „belejátszhatott”). Ugyanígy jelenhetnek meg azok a címkék is, amelyek az összetett szavak esetleges – helytelen – különírásával válnak „több szavas kifejezésekké”.

Az előző – New Yorkkal kapcsolatos – példa rávilágít a folkszonómiák egy másik problémájára is. Az ilyen rendszerekben nem lehet jól megragadni és kezelni a szemantikai hasonlóságot, illetve a szemantikai azonosság problémáját. A 'New York' példából látszik, hogy a folkszonómiák nem tudják kezelni a *szinonimitás* kérdését, vagyis az azonos/hasonló jelentésű, de eltérő alakú szavak problémáját, hiszen a 'newyorkcity' és 'nyc' címkék ugyanarra a fogalomra mutatnak, mégis külön címkéként szerepelnek a rendszerben. A folkszonómiák ugyanilyen módon érzéketlenek a hagyományos szinonimitás megragadására, amelynek „klasszikus” példája a 'kutya'-'eb' szópár.

A folkszonómiáknak a szinonimák mellett kezelni kellene a *poliszémiákat* is, de ezzel a nyelvi jelenséggel sem tudnak mit kezdeni. A poliszéma azonos szóalakhoz több, valahogyan összefüggő, de azért részben eltérő jelentést rendel hozzá. A címke pontos jelentéséhez nyilván szükség lenne a különböző jelentésrétegek elválasztására, amely a folkszonómiák esetében – a felhasználói szabadság miatt – lehetetlen. A poliszémára a következő példák adhatók:

- (1) windows (mint falban lévő lyukak, amelyeken ki lehet látni)
- (2) windows (mint a falban lévő lyukakba mint keretbe helyezett üvegezett nyílászáró szerkezetek, amelyeken szintén ki lehet látni, de a levegőáramlást meg lehet akadályozni)
- (3) Windows (mint operációs rendszer, amely „ablakot nyit a világra”)

A fenti példa utolsó eleme természetesen már csak nagyon laza kapcsolatban van az eredeti jelentésekkel, de a kapcsolat léte nem vitatható, ezért poliszémiának tarthatjuk a harmadik értelmezést is.

Vannak esetek azonban, amikor két teljesen azonos szóalak jelentései már teljesen elválnak egymástól. Ekkor beszélhetünk *homonimáról*. A homonímia annyiban hasonlít a poliszéma problémaköréhez, hogy azonos szóalakok között fennálló viszonyt fejez ki, de abban már különböznek, hogy a szavak jelentésében már nincs semmi közös.

⁹ Előfordulhatnak más technikák is: például az alsóvonás jellel vagy a vesszővel elválasztás szóköz helyett.

A fenti három nyelvi probléma a szóalakok és szójelentések között tetten érhető többértelműségről szól. Ezeket a többértelműségeket úgy lehet feloldani, ha kontrollált szótárakat és kezelési útmutatókat hozunk létre, amelyek pontosan szabályozzák a szinonimák, poliszemiák, homonímiák kezelésének feladatait. Ez az elvárás felkészültséget és fegyelmet követel meg, amely nem áll rendelkezésre a közösségi archiválás körülményei között.

Van azonban a tárgyszavazási tevékenységen belül még egy másik nehézség is. A jelenségek leírására alkalmas fogalmaink, szavaink, kifejezéseink eltérő pontosságúak, és mindig kérdés, hogy milyen pontossági szinten kell/lehet leírni az elemzett dolgokat. Ezt nevezik a *granularitás* problémájának. Amikor a saját kutyánkról készített fotót kell jellemeznünk, akkor ezt sokféle módon megtehetjük.

- (a) állat
- (b) kutya
- (c) briárd
- (d) briárdszuka
- (e) fekete szőrű briárdszuka

A nagy kérdés itt az, hogy mikor melyik szintet válasszuk. Ez azért különösen komoly probléma a folkszonómiák számára, mert ha a felhasználók egymáshoz képest eltérő pontossági, granularitási szintet választanak, akkor nem ugyanolyan pontosságú címkékkel fogják ugyanazokat a dolgokat leírni, ez pedig megnehezíti a visszakereshetőséget.

Persze a címkézési rendszereket idővel sok mindenre meg lehet tanítani, és a fent kifejtett szemantikai problémák legalább egy részére lehet ilyen-olyan megoldásokat találni. Ha összehasonlítjuk a Flickr címkefelhőjének másfél év eltéréssel készült két változatát (2007-ből és 2009-ből), láthatjuk, hogy a fent bemutatott problémák egy része már nem jelentkezik a későbbi változatban.

animal animals architecture art australia autumn baby band barcelona beach berlin
 bike bird birthday black blackandwhite blue boston bw california camping canada
 canon car cars cat chicago china christmas church city clouds concert dance
 day dc de dog england europe fall family fashion festival film florida flower flowers
 food football france friends fun garden geotagged germany girl girls graffiti green
 halloween hawaii holiday home house india island italia italy japan kids la lake landscape
 light live london losangeles love macro may me mexico mountain museum music
 nature new newyork night nikon nyc ocean old oregon parade paris park
 party people photo photography photos portrait red river rock rome san sanfrancisco
 school sea seattle show sky snow spain spring street summer sunset sydney taiwan
 texas thailand tokyo toronto tour travel tree trees trip uk urban usa vacation
 washington water wedding white winter woman yellow zoo

Flickr címkefelhő

A személyes emlékezettechnikákat – amelyet korábban a folkszonómiák egyik előnyös vonásaként mutattunk be –, egy másik szempontból kritizálni is lehet. A fent bemutatott címkefelhő 'me' címkéje csak egyvalaki számára bír jelentéssel és haszonnal (aki a címkét „felragasztotta”), ám a közösség többi tagja számára nyilván kezelhetetlen és haszontalan kategória. A címkék egy jelentős része személyes használatra való, és a közösség egésze számára értéktelen (Guy & Tonkin 2006). Az is gyakori jelenség a folkszonómiák gyakorlatában, hogy egyes felhasználók a többiek számára meglepő, gyakran érthetetlen címkéket aggatnak bizonyos dokumentumokra (például egy macskát, kutyát mutató videót a 'bb' címkével ír le valaki), vagy nem kevés esetben a felhasználók hibás alakban adják meg a címkéiket ('cat' helyett 'cad' címkét rendelnek a „macskás” dokumentumhoz).

Az önhivatkozások, a kontextusfüggő hivatkozások (me, mypicture, todo), az értékshivatkozások (good, cool), a címkék többértelműsége, személy- és kontextusfüggő, ezért a közösség szempontjából vett jelentésnélkülisége, a címkék pongyolasága, zavarossága mind-mind növelik a címkézési rendszerek zajosságát és pontatlanságát. Akármennyi hibát (mégpedig rendszerhibát) találunk is a folkszonómiák világában, mégis működőképesnek tűnik az egész. A nagylétszámú közösség tagjainak aprómunkája „szervesül”, a sokaság eltünteti az egyének egyedi „hibáit”. Erre utal az a megfigyelés, amely szerint a címkék relatív gyakorisága igen hamar beáll egy állandó értékre, vagyis a dokumentumokhoz rendelt címkék megoszlása stabilizálódik (Golder & Huberman 2006). Ez annyit tesz, hogy a közösség egésze végül is konszenzusosnak mondható címkekészletet képes a dokumentumokhoz rendelni.

A szubjektív címkék azért nem igazán okoznak problémát, mert azok mentén úgysem akarnak keresni a felhasználók, így nem is zavarja őket az ilyen címkék jelenléte. Hasonlóképpen, a hibás alakok vagy az érthetetlen, egyéni címkék is „lesüllyednek címketenger mélyére”, és nem igazán látszanak (tehát nem is zavarják) a sokak által használt címkékhez képest.

Az is javíthat a helyzeten, ha adott dokumentum címkézését többen is végzik, végezhetik (nemcsak a dokumentum feltöltője). A Flickr-en adott a lehetőség, hogy a felhasználó barátai újabb címkékkel lássák el a felhasználó képeit, ezzel csökkentve a torzítást, amelyet az okoz,

hogy egy címke felragasztása egyfajta önkényes kategorizálás. Ha több ember címkézi ugyanazt a képet, nyilván nagyobb eséllyel kerülnek rá azok a címkék, amelyeket a nagy átlag is bejelölne (és így a legtöbben meg is találják a képet).

A folkszonómiák esetében a szemantikus vakság problémája sem jelent akkora gondot, mivel ha a keresési oldalon nem kapunk egyértelmű minősítést, megkülönböztetést a keresőfeltételek megfogalmazásakor (márpedig nem kapunk, hiszen a felhasználók nem adják meg az egyértelműsítéshez szükséges többletinformációt, amikor például beírják a 'cica' keresőfeltételt), akkor a keresések feldolgozása során sem tudjuk igazán feloldani a nyelvi többértelműségeket.

* * *

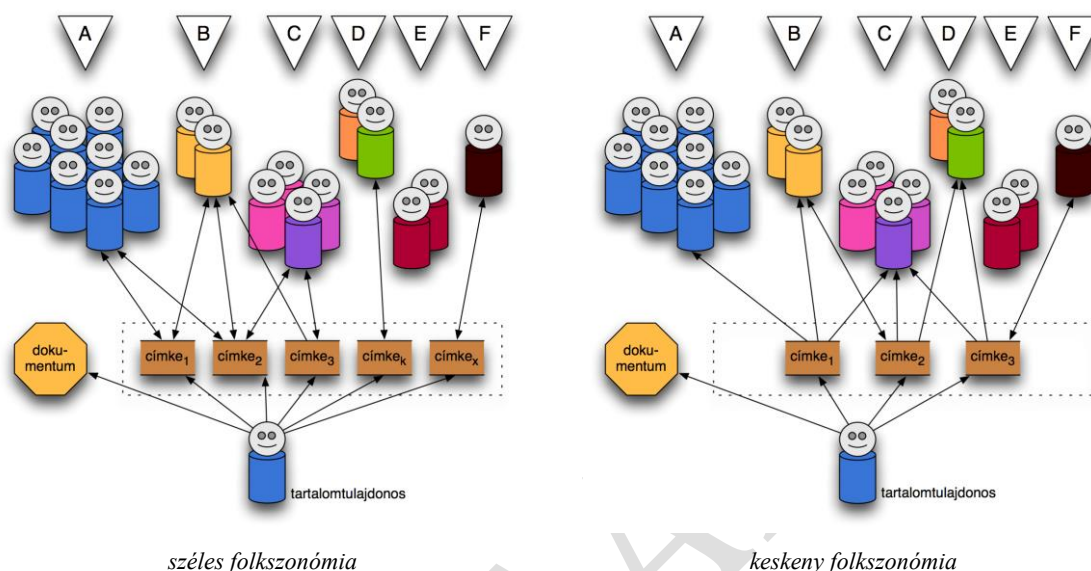
Ha megvizsgáljuk a folkszonómiák „termelési” és fogyasztási” mintázatait és sajátosságait, akkor észre vehetjük azokat a lehetőségeket, amelyek mentén tipizálhatjuk ezeket a rendszereket. Vander Wal elkülönítette egymástól a *folkszonómiák széles és keskeny* típusát (broad and narrow folksonomies) azon az alapon, hogy kik végzik (végezhetik) a dokumentumok, források címkézését, illetve a keresések során a felhasználók milyen „címkeszótárakat” használnak (Vander Wal 2005a). Mindkét folkszonómatípus esetében abból indulhatunk ki, hogy azok a felhasználók, akik a folkszonómiák szkópjába tartozó dokumentumokat feltöltik, maguk is címkéket rendelnek saját dokumentumaikhoz. Minden felhasználónak van valamilyen címkekészlete (címkeszótára), amelynek elemeivel leírhatja a dokumentumokat, illetve keresheti azokat. A címkéket itt tehát kétféleképpen lehet használni. A címkék *aktív* használata a címkézést magát (azaz a címke valamely forráshoz való hozzárendelését) jelenti, míg a *passzív* címkehasználat fogalmával a keresésekhez használt címkék használatát jelölhetjük. Az egyedi felhasználókra fókuszálva kézenfekvő azt feltételeznünk, hogy az aktív és passzív címkeszótárak durván megegyeznek. Azt is feltehetjük, hogy a felhasználók csoportokba rendezhetők címkeszótáraik hasonlósága (azonosság) alapján.

Az igazi kérdés az, hogy a „feltöltést végző” felhasználók mellett a többiek vajon címkézhetik-e mások erőforrásait vagy sem. Amennyiben a válasz 'igen', akkor adott dokumentum címkézését sokan végzik (végezhetik). Ebből fakad az, hogy ilyenkor kialakulhatnak nagy számasságok, amikor a címkék gyakoriságai egyre inkább hatványfüggvény szerinti eloszlást mutatnak, egyre inkább érvényessé válik a hosszú-farok hatás, valamint megjelennek az erős címkék (sokszor használt címkék) a rendszeren belül. Ekkor beszélhetünk *széles folkszonómiákról*. Ilyen esetekben az azonos/hasonló címkeszótárral (és ebből következően: azonos/hasonló érdeklődéssel) rendelkező felhasználók (felhasználócsoporthoz) saját (közös) szótáraik alapján címkézhetnek, illetve kereshetnek. Ez növeli a címkézés erejét és szubjektív hatékonyságérzetét.

Másként működnek a *keskeny folkszonómiák*. Az ilyen rendszerekben a címkézést csak kevesek végzik, végezhetik (a feltöltők és – esetleg – még olyan felhasználók, akiknek a feltöltők engedélyt adnak). Emiatt jóval kevesebb címke kerül be a rendszerbe, és természetesen kevesebb címkézési művelet valósul meg. Ez nehezíti a hatványfüggvény szerinti eloszlás kialakulását és mindazon előnyök hasznosítását, amely ebből fakadhatna. További fontos következménye ennek a helyzetnek az, hogy a keresések során nehezebben alakulnak ki közös szótárak a felhasználók között (mivel az aktív és passzív címkehasználat teljesen elválik

egymástól), ezért a keresés is nehezebb lesz az ilyen rendszerekben. Persze a feltöltők számára nem jelentkezik ez a címkészótár-probléma, ezért személyes használat esetében ezek a rendszerek továbbra is kiválóan működnek (a felhasználók egyéni címkézési tevékenységét *perszonómiának* (personomy) is nevezik – Vander Wal 2005b).

A kétféle rendszert az alábbi két ábra segítségével jellemezhetjük (Vander Wal 2005a).



a folksonómiák típusai

A széles folksonómiára a Delicious, a szűkre a Flickr szolgáltatót példát.

Abból a tényből, hogy a címkéző felhasználók (prosumerek) a folksonómiák építéskor, használatukor nem kontrollált szótárból választanak kulcsszavakat, az is következik, hogy nem lehet tudni, vajon ugyanabban (vagy legalább hasonló) értelemben használnak-e egy adott címkét vagy sem. Látnunk kell azonban azt, hogy ez nem kizárólag a folksonómiák esetében van így. A *szótárprobléma* (vocabulary problem) már a hagyományos kommunikáció vizsgálatok is felmerült a kutatókban, amikor azt vizsgálták, hogy a felhasználók milyen kulcsszavakat használnak ugyanazon objektumok leírására (Furnas et al. 1987). A felmérések eredménye az volt, hogy a hagyományos tárgyszóhasználat világában is gyakran elfordult, hogy a különböző felhasználók nagyon eltérő szavakat, kifejezéseket használtak, vagy ugyanazokat a tárgyszavakat nagyon eltérő jelentés mentén értelmezték. Tanulságos, ahogy Furnas és szerzőtársai a tanulmányukat zárják.

„A hatalmas, komplex rendszerek esetében nagyon sok – belépési pontként funkcionáló – alternatív kifejezésre szükség lehet ahhoz, hogy a felhasználók megkaphassák azt, amit keresnek.” (Furnas et al. 1987)

Erre hivatkozva persze fel lehetne tenni a kérdést, hogy – ebből a szempontból – mennyire volt hatékony a könyvtári, archívumi világ évszázados gyakorlata. Vajon a mindig is létező és ható

szótárprobléma nem teremtett-e és nem tartott-e fenn párhuzamos értelmezési univerzumokat. Szerencsére azonban ezt a kérdést itt nem kell megválaszolniuk.

A folkszonómiák leírásakor a keskeny és széles folkszonómiák elkülönítése mellett figyelembe vehetünk más tipizálási szempontokat is (Marlow et al. 2006). Figyelhetünk arra, hogy milyen vonásai vannak a rendszer működésének magának, de értékelhetjük azt is, hogy a rendszerek milyen eszközökkel próbálják meg ösztönözni a felhasználókat arra, hogy aktív címkézési tevékenységet folytassanak. A folkszonómiák rendszerszintű leírásakor az alábbi szempontokat érdemes figyelembe venni:

- Milyen címkézési jogosultságai vannak a felhasználóknak
 - Ki címkézhet?
 - öncímkézés (self-tagging) – Technorati
 - szabad címkézés (free-for-all tagging) – Yahoo! Podcasts
 - megosztott címkézés (például: barátok, családtagok, ismerősök) – Flickr
 - Ki vonhat vissza címkét?
 - senki – Yahoo! Podcasts
 - címkebejegyző (tag creator) – Last.fm
 - forrástulajdonos (resource owner) – Flickr
- Milyen támogatást adnak a címkézési tevékenységhez?
 - vak címkézés (blind tagging) – Del.icio.us
 - látható címkézés (viewable tagging) – Yahoo! Podcasts
 - ajánló címkézés (suggestive tagging) – Yahoo! MyWeb2.0
- Milyen aggregálás?
 - Zsákmodell: megengedi a címkék többszörös kiosztását (bag-model) – Del.icio.us
 - Halmazmodell: egy címkét egyszer enged egy forráshoz rendelni (set-model) – YouTube, Flickr
- Milyen típusú szociális tárgyról (objektumról, forrásról, dokumentumról) van szó?
 - weboldalak – Del.icio.us, Yahoo! MyWeb2.0
 - bibliográfiai tételek – CiteULike
 - blogposztok – Technorati, LiveJournal
 - képek – Flickr, ESP Game
 - felhasználók – LiveJournal
 - videók – YouTube
 - zeneszámok – Last.fm
 - podcast anyagok – Yahoo! Podcasts, Odeo
 - helyszínek, események – Upcoming, Yelp
- Ki biztosítja a szociális tárgyat (objektumot, forrást, dokumentumot)?
 - résztvevők – YouTube, Flickr, Technorati, Upcoming
 - rendszer – ESP Game, Last.fm, Yahoo! Podcasts
 - globális: bármilyen webforrás beemelhető – Del.icio.us, Yahoo! MyWeb2.0
- Milyen kapcsolat van a források között (resource connectivity)?
 - összekapcsolhatók (linked) – Web
 - csoportba rendezhetők (grouped) – Flickr
 - semmilyen
- Milyen kapcsolat van a felhasználók között (social connectivity)?
 - összekapcsolhatók (linked) – Flickr
 - lehet minősített (barát, ismerős)
 - lehet irányított vagy nem (szimmetricitás)
 - csoportba rendezhetők (grouped) – Del.icio.us (részben)
 - semmilyen

Ha a felhasználói ösztönzés eszközeit, technikáit keressük, akkor az alábbi motivációkat mindenképpen figyelembe vehetjük:

- jövőbeni visszakeresés (future retrieval)
- megosztás (contribution and sharing)
- figyelemfelkeltés (attract attention)
- játék, verseny (play and competition)
- önreprezentáció (self presentation)
- véleménykifejezés (opinion expression)

Nyilván lehetne még másfajta szempontokhoz is igazodni, de nincs most szükségünk arra, hogy a folkszonómiák alaposabb tipizálását elvégezzük. Érdemesebb az eddig leírtak alapján összehasonlítani és értékelni a tudásszervezési rendszereket. Említettük, hogy a terminuslisták nagyon egyszerűen használhatók, de a belső struktúra hiánya miatt nem igazán tudnak támogatást nyújtani az eligazodáshoz. Ehhez képest nagy előrelépést jelent a taxonómiák megjelenése, hiszen azokban egy hierarchikus reláció mentén egy már nagyon jól navigálható struktúrát képezhetünk, amely – elméletileg – sokkal inkább képes lehet a navigáció támogatására. Ennek azonban az az ára, hogy megfelelő tudású és fegyelmezett szakemberek kelljenek a működtetéséhez. Még inkább így van ez a teauruszok esetében, amelyeknek jóval bonyolultabb struktúrájuk van. A teaurusz sokkal rugalmasabb, pontosabb leírásra, barangolásra ad lehetőséget, de az építése, fenntartása „még többbe kerül”, a specializációt keményen meg kell fizetni. A folkszonómiák más munkamegosztási séma szerint működnek, amelynek egyaránt vannak előnyei és hátrányai. Az előnyei közé sorolhatjuk a következőket.

- alacsony „humán erőforrás” költség: nincs szükség megbízható, szakmailag képzett, fegyelmezett dolgozó, s ezért „drága” katalogizáló szakemberekre;
- erős önszabályozási képesség, demokratikus, decentralizált működésmód;
- rugalmas, adaptív, friss;
- azonnali visszacsatolási lehetőséget biztosít;
- könnyű használni;
- a „szabadpolcos felfedezés” támogatása (angolul: 'serendipitous discovery').¹⁰

A folkszonómiák hátrányai között azokat a vonásokat említhetjük meg, amelyek a felhasználók „képzetlenségéből” és „fegyelmetlenségéből” fakadnak. A leggyakrabban emlegetett hátrányok:

- alacsony relevanciamérték (pontosság és felidézés) értékek – alapvetően a szinonimák és poliszémák miatt;
- alapszint-hangolási, granularitási probléma: túl speciális, túl általános címkék együttes kezelésének nehézségei;
- a hierarchia hiánya, vezérlési gondok;

¹⁰ A „szabadpolcos felfedezés” kifejezés alatt azt a jelenséget értjük, amikor a könyvtárak olvasótermeiben kiteszik egymás mellé a frissen érkezett könyveket (vagy egy másik polcra az olvasóktól épp visszaérkezett könyveket), és a könyveknek ilyen egymás mellé kerülése gyakran érdekes „felfedezésekre”, „találatokra” ad alkalmat. A folkszonómiák ugyanezt az élményt képesek nyújtani azzal, hogy a különböző felhasználók által a dokumentumokhoz rendelt címkéket megmutatva érdekes kapcsolatokra „hívják fel” a figyelmet.

- túlérzékenység, védtelenség a rosszindulatú felhasználói tevékenységgel (csalással, rejtőzködéssel, spamtevékenységgel, tendenciózus tévedésekkel) szemben;
- gyenge kereséstámogatási képesség speciális keresési igények esetén.

A pro és kontra érvek más szempontokat emelnek ki illetve tartanak hangsúlyosnak, ezért nem könnyű az összehasonlításuk. A végérvényes döntést talán ma még nem érdemes meghozni, ha választani akarunk a kétfajta rendszer között, mert vélhetőleg még mindkét oldalon sokat fognak tanulni egymástól, és a jövőben inkább a valamilyen mértékig kevert rendszerek jelennek majd meg. Ezt erősítik azok az elképzelések is, amelyek a kétfajta elv ötvözését keresik, s már új terminust is generáltak erre. A 'collabulary' új terminust az alábbi "képlettel" magyarázzák.

folksonomy + controlled vocabulary = collaborative vocabulary = collabulary

Látnunk kell azonban azt is, hogy a folkszonómiák terjedésével párhuzamosan válik egyre jelentősebbé az ún. *szemantikai rés* problémája. A folkszonómiák terjedésének, dominássá válásának legfőbb okaként azt adtuk meg, hogy a metaadat-hozzárendelés nehéz, fáradságos emberi munkáját sokak számára lehetővé téve, nagyszámú „szabad” és – ami a legfontosabb – „ingyenes munkát” várhatunk az önkéntes felhasználók seregétől. Ahhoz, hogy ez működjön, nem lehet semmilyen megkötést előírni a címkézési munkát végzők számára, tehát fel kell adni az előzetes szakképzésre, szakértelemre vonatkozó elvárást, a kontrollált szótárak fenntartásának igényét és a munka ellenőrzésének lehetőségét. Mindebből a fontos mozzanat most az, hogy kontroll hiányában nem lehet semmiféle szemantikai struktúrát elvárni és fenntartani a rendszer elemei között, ugyanakkor a struktúra hiánya miatt egyre nagyobb mértékben nő a szemantikai rés nagysága. E probléma megszüntetésére, csökkentésére megoldást jelenthet egyrészt a szemantikai relációk helyettesítése másféle (forgalmi-használati) relációk bekapcsolásával az egyik oldalon, másrészt a hiányzó szemantikai relációk pótlása speciális emberi szaktudás, valamint gépi intelligencia bevonásával. A kérdés az, hogy milyen módon lehet ez utóbbi hiányzó tudást a rendszerbe táplálni. A válasz pedig a számítógépes szemantikus technológiák, ipari ontológiák, következtető rendszerek felépítése, kifejlesztése és alkalmazása lehet, de azt tudnunk kell, hogy ennek az útnak csak az elején járunk.

A metaadatok dokumentumokhoz rendelésének, és ezáltal a visszakereshetőség biztosításának évszázados hagyománya olyan előfeltevéseken nyugszik, amelyek megkérdőjelezésével az egész építmény könnyen utópisztikus jellegűvé válhat (Doctorow 2001). Cory Doctorow írt egy elgondolkodtató esszét a 'metahadováról', a metaadat-utópia téveszméiről. Nézzük meg röviden, mit is gondolt problémásnak Doctorow. Szerinte a nagy metaadat-elképzelések azért maradtak mindig utópiák, mert az ilyen víziók kidolgozásakor nem vették figyelembe az alábbiakat.

- az emberek hazudnak
(a figyelemért folytatott versenyben „kiszínezik” a metaadatokat)
- az emberek lusták
(ha egy kattintásnál többet kell tenniük)
- az emberek figyelmetlenek ("hülyék")
(akkor sem figyelnek oda eléggé a metaadatok helyességére, ha elemi érdekük lenne, így sokszor elírják azokat)

- „lehetetlen küldetés – ismerd meg önmagad”
(az emberek becsapják önmagukat is, másokat is, sokszor nem megbízhatóak az állításaik)
- a sémák nem semlegesek
(a metaadatsémákban kialakított hierarchiák, struktúrák csak bizonyos kontextusokban érvényesek, máshol elvesztik a relevanciájukat)
- az értékelési-elemzési szempontok sokszor meghatározzák az értékelés kimenetét
(a fontosnak tartott metaadatok kiválasztása jelentős részben meghatározza, hogy mit lehet és mit nem lehet leírni velük, így bizonyos jelenségek „eltűnhetnek” a metavilágból, vagy be sem kerülhetnek oda)
- mindig mindent többféleképpen írhatunk le
(ugyanazt a jelenséget különféle szempontok szerint sokféle módon értelmezhetjük, értékelhetjük)

Erősen megfontolandó Doctorow listája, különösen a szótárprobléma kapcsán korábban írtak fényében. Persze nem feltétlenül kell/lehet minden megállapításával egyetértünk. Azt azért megállapíthatjuk, hogy ezek a kifogások alapvetően pragmatikai jellegűek, ez pedig arra ösztönöz, hogy a továbbiakban elsőbbséget adjunk ennek a szempontnak.

3.3.5.1 Formális jellemzés

A folkszonómiák – a web2-es paradigmának megfelelően – a metaadat-hozzárendelési tevékenységet teljes mértékben saját felhasználóikra bízják. Ezt természetesen csak akkor tehetik meg, ha nem várnak el semmilyen fegyelmet (és semmilyen speciális szaktudást) azoktól, akik a metaadatokat a dokumentumokhoz rendelik (azaz a felhasználói közösségük tagjaitól). Nem is tehetnek mást. Ha bármit megkövetelnének a felhasználóktól, akkor azok vélhetőleg tovább állnának olyan helyekre, ahol szabadon „mozoghatnak”. Ebből következően viszont nem számíthatunk arra, hogy a – „képzetlen és fegyelmezetlen” – felhasználók bármiféle szemantikai relációt felvonnának a rendszerbe, vagy ha mégis, akkor sem remélhetjük, hogy konzekvensen kezelnék azokat. Az ilyen rendszerekben a szemantikai relációk helyett tehát másfajta relációk alkalmazására van szükség, ez pedig azt jelenti, hogy a folkszonómiákat az eddigiekhez képest nagyon másként kell leírunk.

$KOS_{folk} = \langle D, R_1, R_2, R_{10}, R_{11}, S_4 \rangle$, ahol

R_1 lexikografikus rendezési reláció

R_2 ekvivalenciareláció

R_{10} címkegyakorisági reláció

R_{11} címke-együttjárési reláció

A folkszonómiák megjelenésével a legfontosabb változás az, hogy azok az S_i munkaszervezési szabályok közül csak egyet vesznek figyelembe (az S_4 -es normát). Azt sem mindig, sőt talán még az is megkockáztatható, hogy nagyobb azoknak a folkszonómiáknak a száma, amelyek esetében ettől a szabálytól is eltekintenek (amikor figyelembe veszik az S_4 -es szabályt, az akkor is „csak” annyit jelent, hogy a felhasználók kizárólag a saját maguk által feltöltött dokumentumokat

címkézhetik, másokét nem). Azt mondhatjuk, hogy a folkszonómiák egyetlen igazi szabálya az, hogy „nincs szabály” – mindent szabad.

Ebből a „kötetlenségből” sok minden következik. Ha nem követelünk meg semmilyen rendszerépítési szabályt, akkor egyrészt nem biztosíthatjuk a terminusok egyértelműségét a rendszeren belül (vagy másként mondva: nem lesz kontrollált szótárunk), másrészt szükségszerűen elveszítünk minden relációt a rendszerből, hiszen a szabad terminusfelvétel lehetősége mellett a címkézést végző személyektől nem követelhetjük meg azt, hogy az új címkéket hozzárendeljék a rendszerben már létező elemekhez. Ekkor viszont nem is tudunk komolyabb struktúrát értelmezni a címkehalmazon, és emiatt nem számíthatunk a szemantikai struktúra meglétéből fakadó – navigációs és következtetési lehetőségeket biztosító – előnyökre sem. Fontos újdonság azonban a folkszonómiákban az új – se nem szintaktikai, se nem szemantikai, sőt, nem is nyelvi, hanem *forgalmi* – relációk megjelenése (R_{11} és R_{12}). Az R_{11} reláció egy olyan, a címkékhez rendelt gyakorisági viszony, amely azt mutatja, hogy a felhasználók milyen gyakran használják az adott címkét a dokumentumok leírására. Ennek a relációnak azért van kiemelt jelentősége, mert ezáltal újfajta relevanciakezelési lehetőséget lehet biztosítani a folkszonómiák számára. A címkegyakoriság ugyanis megmutatja az adott címkének a felhasználói közösségen belüli „népszerűségét”, fontosságát, és ennek az értéknek a figyelembevétele már elég jó alapot nyújt a relevanciakezeléséhez. Másféle segítséget képes nyújtani az R_{12} reláció, amely azt mutatja, hogy a többi felhasználó korábban milyen más címkéket rendelt a dokumentumokhoz az éppen használatban levő címkével együtt. A címkéknek ezt a fajta együtt járását megmutatva segíteni, orientálni lehet a felhasználókat a megfelelő címkék megtalálásában.

3.3.6 Formális ontológia

Az ontológia tudományát a filozófia egyik ágaként művelték sokezer éven át, mígnem előbb a mesterséges intelligencia kutatások, majd később a szemantikus web kezdeményezés hatására megjelentek az ipari, formális ontológiák fejlesztésével kapcsolatos projektek. E tanulmány írásának időpontjában még nem történt átütő siker ezen a téren. A 2000-es évek elején indult szemantikus web kezdeményezés sokakban nagy reményeket ébresztett, de kezdettől fogva hallhatók voltak azok a kritikus hangok, melyek a formális ontológiák építésének és alkalmazásának nehézségeit hangsúlyozták.

A formális ontológiát is tudásszervezési rendszerként kell értelmeznünk, hiszen – ha eltekintünk az ontológia egyik fontos kérdésétől, a nyelvfüggetlen fogalmak és a nyelvfüggő lexikai egységek elkülönült kezelésének problémájától, és elmossuk a fogalmak és terminusok közti határvonalakat, akkor – az ontológia is tárgyszavakból (fogalmakból) és a köztük értelmezett relációkból áll. A kérdés az, hogy milyen relációkat veszünk bele egy ontológiába, és erre egyelőre még nincs általánosan elfogadott válasz a világban az ezzel a kérdéssel foglalkozó szakemberek között. Pontosabban jelenleg az a válasz, hogy az ontológia belső relációi részben az ontológiát építő és használó emberek *ontológiai elkötelezettségein*, részben az *ontológiahasználati célokon* múlnak. Ebből következően nem, vagy csak nehezen beszélhetünk egyértelmű és széles konszenzuson alapuló ontológiarelációkról (Feilmayr & Wöß 2016).

3.3.6.1 Formális jellemzés

Az ontológiák formális leírása azért nem lehetséges igazán, mert még nem alakultak ki széles körben elfogadott szabványok, megoldások.

$KOS_{font} = \langle D, R_1, R_5, R_{16}, R_{17}, R_{18}, S_1, S_2, S_3, S_4, S_5 \rangle$, ahol

R_1 lexikografikus rendezési reláció

R_5 generikus alá- és fölérendeltje relációpár

R_{16} tetszőleges relációfogalom

R_{17} tetszőleges osztályfogalom

R_{18} tetszőleges attribútumfogalom

A meghatározatlanságot az a tény fejezi ki, hogy az ontológiák általános leírásába csak egy teljesen általános, így praktikus jelentés nélküli reláció-, osztály- és attribútumfogalmakat vehetünk csak fel. Az ontológiák építésére, használatára sokféle kezdeményezés, projekt indult és indul világszerte, de a nagy áttörés egyelőre várat magára, az ontológiák fejlesztésére és használatára vonatkozóan egyelőre nem beszélhetünk sem igazán komoly eredményekről, sem széleskörű konszenzusról.

Érdemes viszont hangsúlyozni itt azt, hogy az ontológiák (formális ontológiák) esetében már nem emberi, hanem gépi használatban kell gondolkoznunk. Az ontológiaépítés még sokáig emberi közreműködéssel valósulhat meg (még ha vannak olyan elképzelések, amelyek szerint az ontológiaépítés feladatát is lehetne „gépesíteni”), de az ontológiahasználat már a számítógépek feladata. Emiatt az ontológiák kicsit ki is lógnak a tudásszervezési rendszerek sorából.

3.4 Dokumentumleíró rendszerek

Mind a névterek, mind a tudásszervezési rendszerek valamilyen struktúrának feleltethetők meg, melyek szemantikája részben a rendszerben alkalmazott relációktól, részben a terminuskészlet összeállításakor érvényesített ontológiai elköteleződésektől függ. Ezek a rendszerek önmagukban, "üresen" minden tudásterületen alkalmazhatóak, nincs olyan szakterület, ahonnan ontológiailag ki lennének zárva.

Vannak azonban olyan tudásrendszerek, amelyek valamilyen ontológiai elkötelezettség mentén adott szakterületre vonatkozó tudáskészletet akarják csak kezelni. A vizsgálódásunkat a könyvtári világra szűkítve, a példákat erről a szakterületről keresve hivatkozhatunk a MARC21, az RDA vagy a BibFrame rendszerekre. Ezeknek is van szemantikája, ezek is felfoghatók egy összetett struktúrának, de ezek funkcionalitása más az eddig tárgyalt rendszerekhez képest. Ezek is használnak, használhatnak tulajdonnévtereket, köznévtereket, tudásszervezési rendszereket, de a fő céljuk speciálisabb: valamilyen dokumentumtípusba tartozó dokumentumpéldányok tartalmi és formai leírása. Ezt a célt azzal érhetik el, hogy felvázolnak egy olyan – a szakma ontológiai elkötelezettségeit reprezentáló, leszűkített – világsémát, amelyben minden elmondható a kiválasztott dokumentumtípusról. Ebben a sémában benne

vannak mindazok az entitások, amelyek szükségesek a terület szakszerű leírásához, és rögzítve vannak a modell entitásai közti relációk, illetve azok a pontok, ahol – ha szükséges – igénybe lehet venni "külső" névtereket, tudásszervezési rendszereket. Ezek a dokumentumleíró rendszerek kapcsolódhatnak (sőt, kapcsolódniuk kell) névterekhez, de más a céljuk, más a tartalmuk.

A névterek igazi értelme az általuk leírni kívánt ontológiai szegmensben felmerülő nevek egyértelmű azonosítása. Ha ezt a feladatukat sikeresen látják el, akkor azok a – könyvtári, múzeumi, levéltári, "tetszőleges nyilvántartási" – rendszerek, amelyek a névterek azonosítóit összerendelik saját névazonosítóikkal, átjárhatókká válnak,

4 Tulajdonnév és megnevezés

Korábban már elkülönítettük a névterek két fő típusát, a köznévtereket és a tulajdonnévtereket egymástól. Ebben a fejezetben a tulajdonnevek jellegzetességei közül egyet kiemelve rámutatunk arra, hogy tulajdonnevek esetében nem a nevet, tehát valamilyen nyelvi szimbólum által közvetített információt kell a névtér központi kategóriájának tekinteni, hanem a név valakire vagy valamire való vonatkoztatását mint eseményt. A tulajdonnevek esetében a névtér nem nevek mint megnevezések tereként értelmezhető. Ezt a fontos tézist Saul Kripke által bevezetett merev jelölő fogalma köré épített elméletre támaszkodva mutatjuk be (Kripke 2010). Ennek során tisztázni fogjuk, hogy a névterekre így tekintve miért fontos – elméleti és gyakorlati szempontból egyaránt – a megnevezői közösség kontextust biztosító szerepe, és kibontjuk azt is, hogy ennek az eseményszemantikai megközelítésnek az alkalmazásából milyen következmények fakadnak.

4.1 A névtér összetevői

A következő fontos kérdés az, hogy miket tartunk a névterek elemeinek. Az nyilvánvaló, hogy akárhogyan is értelmezzük a névtereket, a neveknek mindig a névtér elemei közt van a helyük. A névtereknek a lehető legszűkebben vett értelmezése az lehet, amikor a névtereket nevek egyszerű gyűjteményeként, felsorolásaként definiáljuk megadva azt, hogy milyen típusba sorolható dolgokat írhatunk le velük. Ezen értelmezés szerint a névtér a – valamilyen célra – kiosztható, használható nevek egyedi listája. Ebben az esetben egy földrajzi névtérben az 'Öreghegy' név csak egyszer szerepelhetne, és csak annyit állíthatnánk vele kapcsolatban, hogy ez a név rámutathat egy vagy több földrajzi helyre. A legszűkebben értelmezett családnévtérről is csak annyit várhatnánk el, hogy tartalmazza az összes lehetséges családnevet úgy, hogy minden név csak egyszer forduljon elő. Egy ilyen névtérnek egyszerre kell diszjunktnak és lefedőnek lennie. Diszjunktnak azért, hogy biztosítsa a névtér a nevek egyértelműségét, lefedőnek pedig azért, hogy a névtér garantálhassa, hogy minden lehetséges, kiosztható névalak szerepel benne. A névtér – ilyen leszűkítő értelmezés mentén – arra lehetne használni, hogy valamilyen adatrendszerben kontrolláljuk a névelemek választásának folyamatát azzal, hogy bizonyos adatbeviteli pontokon (például családnevek megadásakor) csak adott névtérből (családnévtérből) lehessen választani. Túl sok gyakorlati haszna azonban nem lenne az ilyen névtérnek, legalábbis abban az értelemben nem, hogy a névtér segítségével elérjük azt a célt, hogy a névtér elemeivel minden esetben egyértelműen azonosítani lehessen a névtérrel leírni kívánt dolgok halmazát. Erre a célra az ilyen névtér csak akkor lenne alkalmas, ha a nevek nagyobb számban állnának rendelkezésre, mint a hivatkozni kívánt dolgok, vagy ha a neveket valamilyen hierarchikus rend szerint újra lehetne hasznosítani. Személyek, testületek, földrajzi helyek és még sok minden más dolog esetében ezt nem tudnánk biztosítani, ezért a nevekkel ilyen kontextusokban nem garantálható a homonimitásmentesség.

Visszatérve arra a kérdésre, hogy mik lehetnek a névtér – további – elemei, a válaszhoz figyelembe kell vennünk azt, hogy milyen célra tartjuk fent a névtereket. Ha a névtér elemeit valamilyen dolgok beazonosítására akarjuk használni (márpedig arra akarjuk), akkor a névtér elemei közé kell sorolnunk azokat a dolgokat is, amelyekre a nevekkel hivatkozunk. Ezeket

nevezzük **névhordozóknak**. A fájlrendszerben az állományok azok a névhordozók, amelyeket az fájlnevekkel azonosítani lehet, a domainnév-rendszerben pedig az IP-címek (illetve az oda pakolt tartalomgyűttesek) azok a névhordozók, amelyeket a domainnevekkel azonosíthatunk. Névhordozó lehet egy személy, akire utalhatunk egy vagy több névvel (de mindegyikkel azonosítani kell tudni), névhordozók lehetnek a testületek vagy a földrajzi helyek is.

Ha a névterek nemcsak neveket, de névhordozókat is tartalmaznak, akkor tartalmazniuk kell még valami mást is, mégpedig a **nevek és névhordozók közti levő kapcsolatokat**. Voltaképpen e kapcsolat kezelése a névterek alkalmazásának igazi értelme.

4.2 A tulajdonnévtér meghatározása

A névtereket a nevek és névhordozók közti – térben és időben dinamikusan változó – kapcsolatrendszer minősége alapján definiálhatjuk: **a névtér egy függvény, amely minden lehetséges megnevezésre egy és csak egy névhordozót ad vissza**. Ez a meghatározás egyszerűnek látszik, a kérdés inkább az, hogy miként lehet működtetni a gyakorlatban az így definiált névteret.

A személynevek példája kapcsán már kiderült, hogy a nevek önmagukban nem képesek a személyek (névhordozók) egyértelmű azonosítására: több személynek is lehet 'Kovács János' a neve. A kérdés az, hogy ilyen feltételek mentén, hogyan képes a névtér függvényként működni. A továbblépés érdekében pontosítani kell az eddig használt fogalmainkon. Az első pontosítás már ott van a névtér fenti meghatározásában, amelyben nem "lehetséges nevek", hanem "lehetséges megnevezések" szerepelnek. Mi a különbség a két kifejezés közt, és miért van szükség az új kategória bevezetésére?

Az eddig hivatkozott névterek mindegyikére igaz (még ha egy-két esetben talán meglepő is), hogy azok a nevek, amelyeket névtérelmként használunk, mind **tulajdonnévként** értelmezhetők. A személynevek, a földrajzi nevek, a testületi nevek esetében ez magától értetődő, a domainnevek esetében is könnyen belátható állítás, a fájlrendszer esetében viszont első hallásra meglepő lehet ez a minősítés. Pedig a fájlnevek ugyanazt a szerepet töltik be az állományok esetében, mint a személynevek a személyek, a földrajzi nevek a földrajzi helyek esetében. Egy konkrét létező dologra (fájlra, személyre, földrajzi helyre) mutatnak rá azért, hogy e megjelölés révén azt az egy konkrét dolgot meg lehessen különböztetni a többi hasonló, konkrét dologtól (az adott fájl a többi fájlától, az adott személyt a többi személytől stb.). Amikor neveket ilyen célra használunk, akkor speciális neveket, tulajdonneveket alkalmazunk.

A tulajdonneveknek az a lényegi, a köznevekhez képest megkülönböztető tulajdonsága, hogy a tulajdonnevek merev jelölők (Kripke 2010). A filozófusok azt ugyan vitatják, hogy a tulajdonneveknek van-e jelentésük, de abban egyetértés van köztük, hogy a tulajdonnevekkel egyetlen individuumra referálunk, és ez a hivatkozás merev abban az értelemben, hogy minden elképzelhető feltételek között (minden lehetséges világban) megmarad ez a jelölő, referáló kapcsolat a tulajdonnév és a hivatkozott individuum között. Ez nem így van a köznevek esetében, amikor is eltérő feltételek közepette (más és más lehetséges világokban) változhat a nevek terjedelme, extenziója. Felmerül azonban a kérdés, hogy a nevek, tulajdonnevek hogyan tölthetik

be ezt a funkciójukat, ha egy fontos filozófiai minőségükben nagyon különböznek a jelölt entitásoktól. Miről van szó?

Amikor – az azonosításra szolgáló – névterek kapcsán névhordozókról beszélünk, akkor mindig – valamilyen típusba sorolható – konkrét individuumról van szó (emberekről, emberek egyedi csoportjáról, földrajzi helyekről, könyvekről, IP-címekről stb.). Ezek létezését, konkrétságát és egyediségét nem nagyon lehet vitatni: ezek mind egyediek abban az értelemben, hogy lokalizálhatók a térben és időben.¹¹ Amikor azonban ezekre a névhordozónak minősített konkrét létezőkre, példányokra utalunk a névtér névelemeivel, akkor a nevek térhez, időhöz nem köthető minősége miatt komoly problémával kerülünk szembe. A nevek ugyanis egy időben több helyen is létezhetnek. Egy személyre tudjuk azt mondani, hogy éppen most hol tartózkodik, de a személy nevére nem mondhatjuk azt, hogy "most itt vagy ott van". Pont ebből fakad a homonimitás problémája is. Amikor egy névvel azonosítani akarunk egy embert, azzal szembesülhetünk, hogy ugyanazt a nevet kiadhatjuk több személy számára is. Akkor viszont adott a kérdés: hogyan lehet a névtereket azonosításra használni, ha a tulajdonnevek – önmagukban – nem alkalmasak erre.

4.3 A nevek egyértelműsítése

A kérdést feltehetjük kicsit másként is: hogyan lehet a (tulajdon)neveket úgy egyedivé tenni, hogy alkalmassá váljanak az individuális névhordozók azonosítására. Nos, a megoldás az lehet, hogy az azonosításra nem a (tulajdon)neveket, hanem **megnevezéseket** (megnevezési eseményeket vagy megnevezési állapotokat) használunk. Ez a megoldás Kripke tulajdonnév-értelmezésére vezethető vissza (Kripke 2010). Kripke szerint egy tulajdonnév referenciáját nem valamilyen hozzá társított leírás alapján határozhatjuk meg, hanem a névhez kapcsolódó, történeti láncba szerveződött megnevezések révén:

"... mikor kimondom a „Rubik Ernő” nevet, annak referenciája az a személy, akit a megnyilatkozásomhoz kapcsolódó névhasználatok történeti lánc (vagy inkább hálójá) meghatároz. A lánc kezdeténél a név bevezető használata áll, és ezt követik a név továbbadó használatai, amelyek révén a név végül eljutott hozzám. Nagyjából így foglalhatjuk össze a tulajdonnevek történetilánc-referenciaelméletét." (Zvolenszky 2015)

¹¹ Az individuumok definitív minősége, hogy azok mindig a térhez és időhöz kötöttek. Egy individuum esetében – elvileg – mindig meg kell tudnunk adni, hogy hol "helyezkednek el" a térben és az időben. Egy épületről, egy hegyről, egy személyről, egy könyvről, egy videokazettáról, egy CD-lemezejről mindig állítható, hogy egyértelműen lokalizálhatók a tér és az idő dimenzióiban. A földrajzi entitásokat eleve azzal tudjuk meghatározni, hogy megadjuk, hogy az összetevői hol helyezkednek el a földrajzi térben egy adott időpontban. Elvileg bármely személyt azonosíthatnánk azon a módon, hogy megadjuk, a tér melyik szegmensét foglalta el egy konkrét időpontban, de ez a megoldás információs szükségletét tekintve nem lenne gazdaságos, hiszen minden személyről tudnunk kéne a mindenkori tér- és időlokalizációjukat. Praktikusabb ehelyett nevezetes eseményeket kiragadni az emberek életéből és azokat kötni a tér- és idő koordinátáihoz. Ilyen nevezetes esemény az ember születése és halála, amit "megfoghatunk" a két esemény tér- és időkoordinátáival.

A (tulajdon)névadás és -használat tehát valamely megnevezői közösség praxisában kialakuló merev jelölő kialakítását és tartós fennmaradását jelenti. Az új elem itt a megnevezői és névhasználati gyakorlat mozzanata. "Rámutatások", névadási gesztusok révén egy közösség gyakorlatában egy névhordozóhoz tapad egy (tulajdon)név, és ez a megnevezési aktusok összefüggő hálózatán keresztül fenntartja a név és a referált dolog közti állandó kapcsolatot. Nem a (tulajdon)név azonosít tehát, hanem a névhasználó közösség megnevezési gyakorlata, a megnevezés(lánc), ami viszont mindig egyedi események láncolatát jelenti. Ez azért fontos, mert az egyedi névhordozókat az individuális megnevezéseken keresztül azonosíthatjuk (ezzel eltűnik a nevek és névhordozók közti filozófiai különbség, hiszen a megnevezések mint események és a névhordozók egyaránt individuális entitások). Természetesen ugyanazt a (tulajdon)nevet más és más megnevezéseken keresztül különböző névhordozókra lehet alkalmazni – akár egy, akár több időben. Ez az értelmezés így kezelni képes azt a homonimitási jelenséget, amiről korábban említést tettünk. A megnevezést nem egyszeri aktusként, hanem egy folyamatos megnevezési eseménysorozatként (vagy épp egy megnevezési állapotként) érdemes felfogni. Ez a megoldás azért képes megoldani a konkrét névhordozók azonosítását, mert a megnevezést (mint eseményt vagy állapotot) már hozzáköthetjük valahogyan a tér és idő koordinátáihoz. A megnevezésnek ez a térben és időben történő lokalizálása úgy is értelmezhető, hogy ez teremti meg azt a kontextust az adott (tulajdon)név számára, amely révén utóbbi alkalmassá válhat az egyedi azonosításra.

Eddig még nem volt arról szó, de most már szükséges utalni rá, hogy nevekből és megnevezésekből többféle is lehetséges, vagyis mind a neveket, mind a megnevezéseket tipizálhatjuk. A személynevek esetében beszélhetünk becenevekről vagy álnevekről, hivatalos, születési, felvett, írói stb. nevekről. Ezek mind eltérő megnevezési gyakorlatokhoz és közösségekhez kapcsolódhatnak, és persze az is előfordulhat, hogy többféle névhasználói közösség ugyanazt a nevet tartja fent valakivel kapcsolatban. A megnevezési gyakorlat mint kontextus felfogás tudja azt a jelenséget is kezelni, amikor két megnevezés esetében a névalak ugyan teljesen megegyezik, de az névhasználói kontextuson belül más típusú nevekről van szó.¹²

A hosszabb elméleti felfeztetés után térjünk rá arra a kérdésre, mi következik mindebből a névterek építésének konkrét gyakorlatára vonatkozóan. Az eddig bemutatott elméleti megfontolásoknak több fontos következménye van.

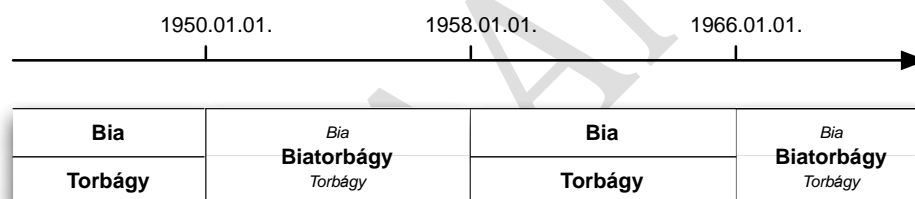
Abból a tényből, hogy a névtérnek két, egyenrangú komponense a név és a névhordozó (a jelölő és a jelölet), az következik, hogy a névtérben mindkét komponenst kezelni kell tudni. Praktikusan ez annyit jelent, hogy mind a névhordozókat, mind a neveket azonosítani kell valamilyen módon, és – talán ez a legfontosabb – harmadik komponensként fel kell venni a névtérbe a köztük levő kapcsolatot is. Annál is inkább szükség van erre, mert szinte mindegyik névhordozótípus (a személyek, a testületek, a földrajzi helyek stb.) névhasználati gyakorlatában nagyon gyakori a szinonimitás jelensége, vagyis az, hogy egy névhordozóhoz több név is tartozik. Ezt úgy kell kezelni, hogy mind a névhordozót, mind a hozzá kapcsolható összes nevet önmagában azonosítjuk,¹³ és a megnevezési esemény (gyakorlat) leképezéseként ezek

¹² Egy település hivatalos nevének megváltoztatása után a régi névhez két – olykor radikálisan eltérő, egymással versengő – névhasználói gyakorlat tartozhat: a hivatalos gyakorlat akadályozhatja, tilthatja a régi név használatát, míg a "laikus", hétköznapi praxis támogathatja azt.

¹³ Az azonosítás itt mesterséges azonosító hozzárendelését jelenti.

összekapcsolását is annyiszor elvégezzük, ahány névadási esemény (praxis) előfordul az adott névhordozó esetében (hogy erre miért van szükség, rövidesen bemutatjuk Biatorbágy példáján keresztül).

A megnevezést mint kontextust elég csak egyértelműen azonosítani, nem kell alaposabban feltárni (bár szükség esetén lehetséges). Azzal, hogy úgy kapcsoljuk össze a névhordozó (mondjuk a személy) egyedi azonosítóját (personid) és a hozzárendelt névelem egyedi azonosítóját (nameid), hogy ezt az összerendelést is azonosítjuk, magát a megnevezési kontextust ragadhatjuk meg. Nézzük meg Biatorbány példáját! Bia és Torbágy községeket 1958-ban egyesítették, az új településnek 1950 és 1958 között Biatorbágy volt a neve. 1958-ban újra szétvált a két település, és hivatalosan is visszakapták a Bia és Torbágy neveket. Végül 1966-ban újra egyesítették őket, és innentől kezdve megint a Biatorbágy nevet viseli a település. A két korábban önálló település, Bia és Torbágy az idők során kétszer is hivatalosan viselte a saját nevét, kétszer pedig a Biatorbágy hivatalos névvel jelölték az összevont települést. Ez három nevet, de hat megnevezést (névhasználati gyakorlatot) jelent. A nevek szintjén a Bia, Torbágy és Biatorbágy nevek keletkezésük óta léteztek, de a megnevezések/névhasználatok szintjén nem mindig volt így. Ez az eset arra is példát mutat, hogy amikor a megnevezésekkel hozzákapcsolunk egy névhordozóhoz egy adott nevet, akkor megtehetjük, hogy – több megnevezés felvételével – ugyanazt a nevet többször is odakötjük ugyanahhoz névhordozóhoz.



egy név, több megnevezés(i esemény)

Természetesen az emberek számára a nevek az elsődlegesek, és számukra a fenti nevek (Bia, Torbágy, Biatorbágy) többszörös használata nem kell, hogy elkülönüljön egymástól, de precíz, gépi és szaktudományi feldolgozás számára ezeket a különbségeket kezelni kell.

A névterekben tulajdonneveket kapcsolunk névhordozókhöz megnevezéseken keresztül. Eddig a névterekről úgy beszéltünk, mint amelyekben csak tulajdonneveket kell nyilvántartanunk. A tulajdonnevek kezelésével kapcsolatos problémák elemzése után érdemes kitékinteni a köznevek kérdéskörére is. Vannak ugyanis olyan nyelvi rendszerek, amelyek közneveket tartalmaznak, és hasznos lehet tisztázni ezek névterekhez való viszonyát.

Korábban elkülönítettük a névterek szűkebb és tágabb értelmezését egymástól, és a szűkebb értelmezés szerinti névtéren a tág értelmezésű névterekben használható tulajdonnevek egyedi felsorolását értettük. Ilyen gyűjteményt összeállíthatunk köznevekből is azért, hogy valamely ismeretterület azzal jellemezzünk, hogy megadjuk a vele kapcsolatos közneveket (fogalmakat) és a nevek (fogalmak) közti levő kapcsolatokat. Ha így teszünk, akkor egyrészt más jellegű feladatot kell elvégeznünk, másrészt eltérő eredményekhez juthatunk. Amíg ugyanis a tulajdonnevek szemantikailag lapos struktúrába szervezhetők, addig a köznevekből több

relációval összefogott, olykor gazdag, összetett struktúrát lehet képezni. Olyan fogalmi struktúrát, amely alkalmas lehet a világról való tudásunk (egy részének) reprezentálására.

* * *

A tulajdonnévterek fogalmának meghatározása után érdemes egy további pontosítást tenni, és egy újabb megkülönböztetést tenni a névterek kategóriája kapcsán. Lehetséges ugyanis egy szűkebb értelmezést is adni a tulajdonnévterek fogalmára, aminek használhatósága kisebb hatókörű, mint amire nekünk szükségünk lehet, de bizonyos célok megvalósításához nagyon hasznos eszközt biztosíthatna az újonnan konstruált fogalom.

A fent megadott tulajdonnévtér-értelmezésre azt mondhatjuk, hogy az – nyelvészeti szempontok szerint – tágabb értelemben vett definíció volt. A **tágabb értelemben vett névtér** a tényleges egyedi megnevezések mint események listáját tartalmazza (a kiegészítő adatokkal együtt). Igazán pontosan nem névtérnek, hanem **megnevezéstérnek** kellene neveznünk az ilyen névtereket, de – a ritka és kivételes esetektől eltekintve – e terminus használatától tartózkodni fogunk.

Lehetőség van arra, hogy a fentihez képes szűkebb értelemben vett névtérfogalmat alkossunk. A **szűkebb értelemben vett névtér** a – megnevezési események során kiosztott vagy kiosztható – lehetséges nevek egyedi listáját tartalmazza. A számosságát tekintve utóbbi – bizonyos pontokon – lehet bővebb vagy szűkebb terjedelmű. A szűkebb értelmezés szerinti névtér a megnevezések során kiosztott neveket csak egyszer tartalmazza, ebből a szempontból kevesebb elemből áll a megnevezéstérhez képest, viszont tartalmazhat olyan egységeket is, amelyeket az adott névhasználati gyakorlat a megnevezési események során még nem vett igénybe, így azok nem szerepelnek a megnevezéstérben.

Érdemes konkrét példával szemléltetni a fenti különbségtételt. Vegyük a személynevek példáját! A dolgokat kissé leegyszerűsítve tételezzük azt, hogy a személynevek két komponensből állnak, egy családnévből és egy keresztnévből. Vegyünk pár konkrét névelemet: 'Kovács', 'Molnár', 'Tóth' mint családnév, illetve 'István', 'Pál', 'József' mint keresztnév. Ezek a nevek egyszer szerepelhetnek a szűkebb értelemben vett családnévtérben, illetve keresztnévtérben. akkor, ha külön névteret hozunk létre a család- és keresztnevéknek. Ha egyetlen – család- és keresztnevékből álló – személynév névteret definiálunk, akkor ahány változatban előfordul vagy előfordulhatnak a példának vett család- és keresztnevék egy néven belül ('Kovács István', 'Tóth Pál', 'Tóth József' stb.), annyiféle személynevet kell rögzítenünk a szűknévtérben (persze, csak egyszer). A tágabb – és praktikus szempontból számunkra hasznosabb – értelmezés szerint viszont ahányszor a 'Kovács István' nevet használták megnevezésre, annyiszor fel kell vennünk ezt a személynevet a tágabb értelmű névtérbe. Ezzel szemben ha a névtér által lefedett megnevezési kontextuson belül még egyszer sem használták a 'Tóth Pál' személynevet, akkor az nem lesz benne a megnevezéstérben, míg a szűknévtérben szerepelni fog.

5 Lokális, nemzeti és nemzetközi névtér

A névterek fogalmának elemzését el kell végezni egy eddig még nem említett szempont szerint is. A névtereket valamely névállomány elemeinek egyértelmű azonosítását biztosító rendszerként definiálva mindig valamilyen névhasználói kontextushoz kell kötnünk. Akkor viszont mindig kérdés lehet az, hogy kontextust adó névhasználói közösség milyen szinten létezik. Innen tekintve beszélhetünk lokális, nemzeti és nemzetközi névterekről. A továbbiakban ezek tulajdonságait, egymáshoz való viszonyait fogjuk elemezni.

5.1 A név fogalma, funkciója

Mielőtt a névterek fogalmát újabb elemzésnek vetnénk alá, érdemes összegezni, hogy mit tudunk mondani magáról a név fogalmáról.

A nyelvi konstrukciók konvencionális jelölések, ami azt – is – jelenti, hogy a nyelvi jelekhez önkényes módon kapcsoljuk a jelentéseket. A neveknek két típusa, a köznévi, illetve a tulajdonnévi gazdag, illetve lapos szemantikával rendelkeznek. A különbség fő oka az, hogy a tulajdonnevek merev jelölőnek számítanak, míg a köznevek nem, mert utóbbiak referenciája változhat, az előbbieké meg nem. A tulajdonnevek adott kontextusban, adott névhasználói közösség rámutatásán keresztül „létrejött” megnevezések. A megnevezés esemény, minden esemény időhöz kötött (köthető), így a megnevezéseket is időhöz kell (lehet) kötni. A nevek alkalmazása kétféle értelemben sem egyértelmű. A többértelműség létre jöhet a nevek homoním vagy szinoním kapcsolódása révén. Hononima esetén ugyanazt a névalakot, többféle jelentéssel (tulajdonnévi esetén többféle referátummal) használjuk, szinonima esetén ugyanazt a jelentést (tulajdonnévi esetén ugyanazt a referátumot) többféle névalakkal is összekapcsoljuk. A névvel jelölt, referált dolgot nevezhetjük névhordozónak. A tulajdonnévterek kontextusában a jelöllet valamilyen individuum (személy, hely, testület, film, könyv, periodika, cikk stb.). Ebben a kontextusban az – eredeti funkciója szerint azonosításra használt – név helyett ekvivalens módon szokás alkalmazni az olyan terminusokat, mint az azonosító, azonosítószám, cím, műcím, földrajzi cím, térbeli cím.

5.2 A névtér fogalma, funkciója

Először foglaljuk össze röviden azt az eddigiel alapján, mit is jelent a névtér fogalma, mi az alapvető funkciója. Kiderült, hogy kétféle névtérről kell beszélnünk, mert a köznevek és a tulajdonnevek számára egymástól részben eltérő logikájú, más felépítésű névtereket kell fenntartanunk. A két típus között sok a hasonlóság, de van különbség is köztük. Mivel a projektben a tulajdonnévtereken van a fókusz, ezért inkább ezekre koncentrálnunk. Amikor nem tüntetjük fel külön jelzők alkalmazásával, hogy milyen névterekről van szó, akkor tulajdonnévterekről beszélünk. Amennyiben szükség van arra, hogy a megállapításainkat névtértípusok szerint elkülönítve tegyük meg, akkor ezt jelezni fogjuk.

A nyelvi konstrukcióinkkal vagy egyedi dolgokra vagy dolgok halmazára akarunk utalni, és a két cél megvalósítására kétfajta nevet használhatunk. A **tulajdonnevekkel** individuumokra referálunk, a **köznevekkel** individuumok valamilyen gyűjteményére.

A nevekkel referált individuumokat (vagy azok gyűjteményét) **névhordozóknak** nevezzük.

A nevek mindig valamilyen névhasználói **kontextusban** nyerik el a jelentésüket, referenciájukat, minden nyelvi megnyilvánuláshoz szükséges valamilyen kontextus ismerete.

A nyelvhasználat során a nevek könnyen és gyakran többértelművé válnak, és az egyértelműséget csak az adott megnyilatkozás egyedi kontextusa képes biztosítani (a családon belül egyértelmű lehet, hogy a 'Jóska' név kire utal, de egy osztályban már nem. Az osztályban egyértelmű lehet az, hogy a 'Kovács Jóska' név kit azonosít, de az iskolában már nem feltétlen van így).

Minden névhasználat, minden névtér **kontextusfüggő**. A névhasználatnak, a névtérnek mindig van egy **szkópja**, amely azt fejezi ki, hogy a nevek milyen körére terjed ki a figyelem.

A tulajdonnevek másként működnek a köznevekhez képest abban az értelemben, hogy az előbbieket **merev jelölők**, míg utóbbiak nem azok. A tulajdonnevek segítségével bármilyen lehetséges világban ugyanarra az individuumra mutatunk, ez a referálás nem változik meg soha. Ez azért van, mert tulajdonnevekkel mindig olyan konkrét egyedre mutatunk, amely a térben és időben mindig elhelyezhető valahova. A köznevek jelentése, referenciája ezzel szemben mindig változhat a lehetséges világok között, így sosem lehetünk az adott kontextuson belül érvényes értelmezés, jelentés, referencia ugyanaz marad egy másik lehetséges világban, egy másik kontextuson belül.

A nevek nem lehet térhez is időhöz egyértelműen hozzáköttni, hiszen a nevek térben és időben folyamatosan újra lehet használni. Egy nevet ugyan lehet példányosítani, de a név példánya (token) és a név általános formája (type) megegyezik. Ezért önmagában a név nem alkalmas egyértelmű referálásra.

A tulajdonnevet úgy tudjuk egyértelmű referálásra használni, hogy nem a tulajdonnevet magát használjuk erre a célra, hanem azt a **megnevezést** (megnevezési eseményt, folyamatot), amelynek keretében egy névhasználói közösség megnevezi egy individuumot, és ezt a nevet (névkiosztást) valamennyi ideig fenntartja. A megnevezési eseményt már lehet példányosítani azzal, hogy az eseményt az idő valamely pontjához kötjük.

A névterek alapfunkciója a nevek és megnevezések, valamint a névhordozók **egyértelműsítése**, vagy másként a nevek és megnevezések gyakorlatban létező többértelműségének megszüntetése, feloldása, és ezzel adott névhordozó (vagy névhordozók gyűjteményének) egyértelmű beazonosítása.

A névterek **nomonima- és szinonimamentes névállományok** abban az értelemben, hogy a hononim és szinonim névalakok mindig egyértelműen fel vannak oldva.

Az egyértelmű azonosításra nyelvi jelek nem alkalmasak (sem a köznevek, sem a tulajdonnevek), ezért helyettük **mesterséges azonosítókat** (számokat) kell használni. Természetesen egy mesterséges azonosító is kontextusfüggő, és ebben az értelemben nem egyértelmű, de a névtér szkópjába tartozó névállományt, illetve a megnevezések állományát mindig lehet egyértelműsíteni mesterséges azonosítókkal.

Az egyedi mesterséges azonosítók kiosztásához és folyamatos használatához szükség van „hagyományos” leíró adatokra is annak érdekében, hogy ezek segítségével az ember is egyértelműen azonosítani tudja a neveket, névhordozókat. Az emberi beazonosításhoz szükséges, hagyományos adatokat nevezhetjük névjegynek, ami a minimálisan szükséges és elégséges adatelemek körét tartalmazza. A **névjegy** tartalma névtértípusonként változhat.

Informatikai értelemben a névtér valamely entitástípus előfordulásainak és ezek neveinek (megnevezéseinek) **normalizálását** jelenti. A tulajdonnévterek – a rájuk jellemző – névjegyekkel kiegészített terminuslisták, a köznévterek leggyakrabban taxonómiák vagy teuauruszok.

Amikor egy – neveket is kezelő – adatbázison, adatrendszeren belül normalizálják a névállományt, akkor saját névteret (adatrendszernévteret) hoznak létre, hiszen a normalizált névkezelés az egyértelmű azonosítást szolgálja az adatbázison, rendszeren belül. Ilyen esetben a névtér, az egyértelmű névkezelés kontextusát az adatbázis, adatrendszer határai adják.

Ha egy intézményen belül több névtérrel rendelkező adatrendszer is van, akkor ezek között, ezekre támaszkodva létre lehet hozni **intézményi névteret**, amely valamilyen módon egységesen kezeli az intézményen belüli, különböző adatrendszerek névtereit. Az intézményi névtér szkópja az intézmény összes, névtérrel rendelkező adatrendszerén belül kezelt névállomány egységesítése, egyértelműsítése.

Mind az adatrendszerek normalizált állományait, mind az intézményi névtereket nevezhetjük **lokális névtereknek**.

A lokális névterek nevei, megnevezései csak saját határaikon, saját szkójukon belül egyértelműek. A nemzeti kulturális adatvagyon egymástól elkülönült intézmények adatrendszereiben egymástól elkülönülten van tárolva. Még ha mindegyik kulturális intézmény fel is építi saját lokális névterét, az adatrendszerek közti semmi sem biztosítja az átjárhatóságot, interoperabilitást mindaddig, amíg a lokális névtereket nem kötik össze egymással. Ennek a kapcsolatnak a megvalósulását biztosíthatja a **nemzeti névtér**.

A nemzeti névterek szkópját a nemzeti kulturális adatrendszereken belül létező névállományok adják, de a nemzeti névtérbe csak a kulturálisan jelentős neveket és névhordozókat érdemes „betenni”. Arra kell felkészülni, hogy a közgyűjteményekben mindig maradnak, maradhatnak olyan nevek, névhordozók, amik nem kerülnek be a nemzeti névtérbe.

A nemzeti névterek építésekor és fenntartásakor a nevek és névhordozók egyértelmű azonosításának feladata mellett fontossá válik a lokális/intézményi névterek nevei, pontosabban azok lokális azonosítói közti kapcsolatok megteremtése.

A **nemzetközi névterek** a nemzeti – ma még sok esetben: az intézményi – névterek azonosítói közti kapcsolat megteremtését végzik.

A **tulajdonnévtér** adott adatkörön belül adott névtípusba tartozó megnevezések és a nevekkal jelölt névhordozók egyértelmű azonosítását végző adatrendszer (megnevezések és megnevezettek egyértelmű azonosítását végző adatrendszer).

A **köznévtér** adott tudásterületre vonatkozó általános fogalmak jelentését (ezáltal a világra vonatkozó tudásunkat), a fogalmak közti kapcsolatok struktúráján keresztül, adott nyelven kifejező adatrendszer

5.3 Lokális névterek

A lokális névterek lehetnek tulajdonnévterek és köznévterek egyaránt, utóbbiak struktúrája lehet többféle. A meghatározó minőség a névtér szójában van. Akkor beszélhetünk lokális névtérről, ha a névtér egy adatkör vagy egy intézmény névállományát kezeli. Technikai értelemben az a kérdés, hogy egyetlen névtér létezik-e, vagy sem (azonos névtípus esetén). Egy adatrendszer számára nyilván csak egy névteret lehet építeni (adott névtípus esetén). Ha egy intézményen belül több adatrendszer is létezik, amelyek névállományaiból névteret lehet építeni (adott névtípus esetén), akkor a névtér fenntartására kétféle megoldás képzelhető el.

Lehetséges, hogy az egyes adatrendszerek saját névteret tartanak fent, és köztük egy összerendelési munka eredményeként teremtik meg az átjárhatóságot. Ekkor az intézményi névtér és az adatrendszernévterek különböznek egymástól, és az intézményi névtér ugyanazt a szerepet tölti be, mint a nemzeti névtér a lokális névterek között.

A másik lehetőség az, hogy a különböző adatrendszerek egyetlen névteret építenek közösen. Ebben az esetben biztosítani kell azt, hogy a névtérépítés során minden pillanatban látható legyen a névtérben tárolt névállomány minden eleme, hogy teljesíthetők legyenek az elemek egyértelműségére és egediségére vonatkozó elvárások.

A ténylegesen létező lokális tulajdonnévtereket nem nagyon lehet látni, meghivatkozni, mert ezek az intézmények falai között léteznek. A könyvtárak besorolási állományai ilyen lokális névtereknek tekinthetők. Amennyiben egy nemzetközileg látható névtér nem igényli magának azt a státust, hogy a saját azonosítóit más, alsóbb szintű névterek elvigyék és saját rendszerükbe építsék, hogy az intézményközi interoperabilitás alapjává tegyék, akkor igazából csak lokális névtérről beszélhetünk. A földrajzi neveket kezelő Getty földrajzi nevek tezauruszát vagy az OpenStreetMap térkép- és névszolgáltatását – ebből a szempontból – ilyennek, nemzetközileg látható, lokális névtérnek tekinthetnénk.

A lokális köznévterekre példa lehet minden olyan, tartalmi feltárást támogató, tárgyszó-hozzárendelő munkát támogató névállomány, amelyet valamilyen kontroll alatt tartanak.

5.4 Nemzeti névterek

A nemzeti névtér fogalmának meghatározásakor több szempontot is figyelembe venni. Egyfelől a névterek egymáshoz való viszonyára fókuszálva a nemzeti névtereket szembe lehet állítani a lokális névterekkel. Másfelől a névterek által kezelt adatkörök tartalmára fókuszálva lehet szűkebb és tágabb értelemben vett nemzeti névterekről beszélni.

5.4.1 A nemzeti névtér szűkebb értelmezése

Amikor adott adatrendszeren belül valamely névállományt egyértelműsítünk (normalizálunk), akkor ez a normalizált névállomány nem önmagában áll, hanem az adatbázis tartalmának adott

szegmensként funkcionál. Egy lokális névtér mindig valamilyen adatrendszer kiegészítője és/vagy része. A lokális névtér fő funkcióját, a szkópjába tartozó nevek egyértelműsítését azáltal biztosíthatja, hogy minden névhez és minden névhordozóhoz mesterséges azonosítót rendel. Ez az egyedi azonosító egyértelműséget teremt a lokális névtéren belül, de ha kilépünk ebből kontextusból, akkor semmi sem biztosítja ezt az egyértelműséget tovább. Más adatrendszerek, más lokális névtérek is kioszthatják ugyanazt az azonosítót, és akkor már csak azon az alapon lehetne megkülönböztetni a két – numerikusan teljesen egyező – azonosítót, ha tudjuk, melyik adatrendszerből, névtérből származik. Ha az intézményeknek, névtéreknek is adnánk egyedi azonosítót, akkor a két azonosító (a névtér- és a névazonosító) összekapcsolásával már újra egyértelműséget teremthetnénk a mesterséges azonosítók szintjén, de ezzel a megoldással semmit sem tudnánk magukról a nevekről, névhordozókról, azok különbségeiről, azonosságairól. Ez abban a pillanatban problémát jelentene, ha feltételezhetnénk, hogy a névtérek állományai közt lehetnek átfedések. Márpedig a kulturális adatrendszerek világában pont ez a helyzet. A közgyűjtemények által kezelt névállományokban gyakran komoly átfedések vannak, hiszen a kulturális élet szereplői, intézményei, helyszínei ismétlődően felbukkannak a kulturális dokumentumokat őrző, egymástól ilyen-olyan szempontok szerint elkülönített közgyűjtemények adatrendszereiben.

A kulturális értékek iránti érdeklődés nagyon sokszor intézményközi, amit az analóg világban kialakított intézményi elkülönültség keretei közt nem lehet jól kiszolgálni. A kultúra szerteágazó jellege és az intézményi elkülönültség merev határokat húzó logikája közötti inkompatibilitást a predigitális világban nem igazán lehet enyhíteni, megszüntetni. A hálózat elterjedése megteremti ugyan annak lehetőségét, hogy minél többen minél több kulturális dokumentumhoz hozzájussanak, ám még a közgyűjteményeink teljesskörű digitalizálása sem oldaná fel az adatrendszerek közti átjárhatatlanságot, amit a – predigiális korban kialakult, és akkor még szükségszerűnek mondható – intézményi elkülönültség okoz.

Mi lehet a megoldás erre a problémára? A nemzeti névtérek létrehozása, a nemzeti névtér azonosítóinak összerendelése a lokális névtér-azonosítókkal. Mi a nemzeti névtér, miért van rá szükség, hogyan működik? Nézzük végig, milyen válaszok vannak ezekre a kérdésekre.

A lokális névtérek egyértelműséget teremtenek a lokális adatrendszereken belül, de az egyértelműsítésre használt mesterséges azonosítók intézményenként különböznek, és nem segítenek a két névtér közötti kapcsolat kialakításában.

A hálózaton keresztül hozzáférhető – kulturális, közgyűjteményi – adatbázisok egymást átfedő szegmenseiben a legnagyobb részt azok a dolgok, entitások (névhordozók) adják, amelyek névtérekkel egyértelműsíthetők. A kulturális „közös részbe” tartoznak a kulturális dokumentumok (könyvek, filmek), a kulturális ágensek (személyek, testületek), a kulturális helyszínek.

Kulturális befogadói oldalról kívánatos cél, és ez alapján a közgyűjteményekkel szemben megfogalmazható elvárás az, hogy a kulturális adatbázisok legyenek átjárhatók (interoperábilisek). Az átjárhatóság legfontosabb eszköze a közös névtérek kialakítása és használata. A lokális egyértelműsítés, lokális azonosítás nem segít ebben, csak egy rávezető lépés lehet, mert a különböző forrásokból származó, lokális szinten egyértelműsített névállományok nemzeti szinten átfedhetik egymást. A megoldást a nemzeti névtérek közös használata jelenti.

A nemzeti névtér egymástól elkülönült – kulturális – adatrendszerek tartalmilag átfedő, adott névtípusba tartozó névállományainak együttes kezelésére alkalmas, nemzeti szinten

egyértelműsítő, azonosítási rendszer. A nemzeti névtér kialakítása megtörténhet a lokális névterek, a lokálisan elkülönült névállományok, a lokális azonosítók összerendelésével, összefésülésével vagy közösen kezelt online névállomány működtetésével. Hosszútávon lehetnek nemzetinévtér-használó és részlegesen összefésült lokálisnévtér-használó együttműködő partnerek. A nemzeti névterek esetében nem működőképes a hagyományos értelemben vett adatcsere művelete, mert a különböző forrásokból származó, lokális szinten egyértelműsített névállományok nemzeti szinten átfedhetik egymást. Az új nevek és névhordozók azonosítása egyelőre csak emberi közreműködéssel végezhető.

Első megfogalmazásban – szűkebb értelemben – azt mondhatjuk, hogy a **nemzeti névtér** adott adatkörön belül, adott névtípusba tartozó megnevezések és a nevekkal jelölt névhordozók egyértelmű azonosítását nemzeti szinten végző, és az adatokat publikáló adatrendszer (megnevezések és megnevezettek egyértelmű azonosítását nemzeti szinten végző adatrendszer).

Mire jó a névtér? A kulturális adatbázisok építése során felmerülő nevek, névhordozók (személyek, földrajzi helyek, testületek stb.) egyértelmű – lokális, intézményi szintű – azonosítására. Mire jó a nemzeti névtér? Egyfelől a különböző kulturális adatbázisokon belül használt nevek, megnevezések, névhordozók egyesített és egységesített azonosítására, másfelől a lokális névterek azonosítóinak összefésülésére. Ezek alapján a lokális és nemzeti azonosítókon keresztül mindkét irányban átjárhatóságot lehet teremteni azáltal, hogy a nemzeti névtér keretében a lokális névterek közt felépített kapcsolatok alapján fel lehet kínálni a lokális adatrendszerek felé mutató linkeket, ami a látogatók számára lehetővé teszi, hogy egy-egy kattintással további adatokhoz juthasson az őt érdeklő kulturális jelenséggel kapcsolatban.

5.4.2 A nemzeti névtér tágabb értelmezése (névtár)

A fentiekben a nemzeti névtér fogalmának szűkebb értelmezéséről beszéltünk. A továbbiakban bemutatunk egy tágabb értelmezést, amely kitágítja a nemzeti névtér fogalmát a szolgáltatott adatkörök tartalma alapján.

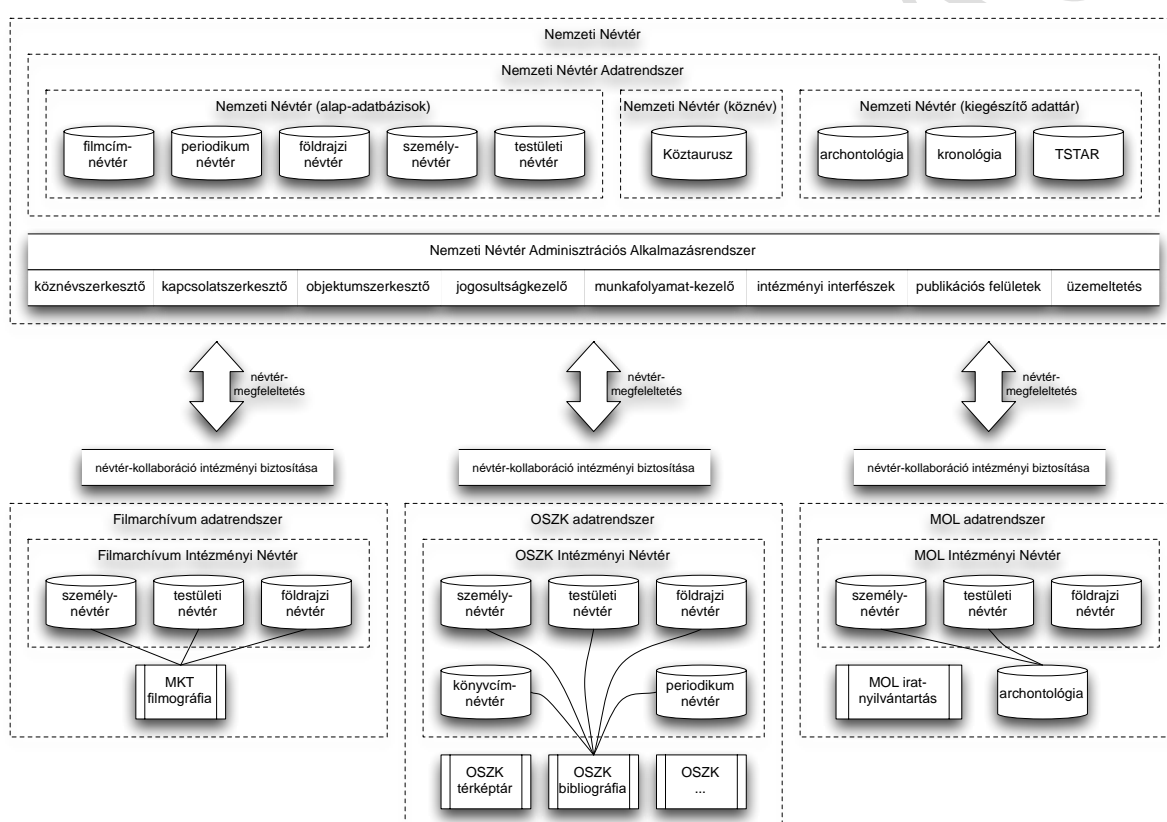
A nemzeti névterekbe – első megközelítés alapján – a kiválasztott névhordozók, névtípusok névállományai tartoznak. A nemzeti névterekben együttműködő közgyűjtemények által épített kulturális adatrendszerek tartalma ennél természetesen sokkal bővebb. A lokális és nemzeti névterek építésekor, a kulturális tudás rendezésekor olyan – kulturális érdeklődésre számot tartható – adatok is keletkeznek, amelyek a különböző névtértípusok közti kapcsolatként értelmezhetők. Ilyenek – többek között – a könyvek és személyek közti kapcsolatokat leíró személyes bibliográfiai adatok, a filmek és személyek közt létező filmográfiai adatok, a személyek és testületek közti viszonyt megragadó archontológiai adatok, a testületek (iskolák) és személyek (diákok) közti tartós viszonyokat számon tartó iskolai adatok, de említhetjük még a földrajzi helyekhez kapcsolható térstatisztikai adatköröket, a testületek (pártok), személyek (jelöltek) és földrajzi helyek közötti kapcsolatokat ábrázoló választási adatokat. A kulturális adatbázisok, a névterek építése során ezek az adatok vagy „automatikusan” előállnak vagy gyosan és könnyen előállíthatók, és így könnyen szolgáltatottá is tehetőek. Ezt a lehetőséget érdemes kihasználni és a nemzeti névtér fogalmát ebben a tágabb értelemben definiálni. Ezáltal

sokkal több értékes adat kerülhet ki a nemzeti névterek felületeire. A nemzeti névterek értelmezését a névterek közti kapcsolatokkal kibővítve jelentősen növelni lehet a névterek, közgyűjteményi adatrendszerek iránti érdeklődést.

Második megfogalmazásban – tágabb értelemben – azt mondhatjuk, hogy a **nemzeti névtér** adott adatkörön belül, adott névtípusba tartozó megnevezések és a nevekkal jelölt névhordozók, valamint a különböző névterek közti kapcsolatok egyértelmű azonosítását nemzeti szinten végző, és az adatokat publikáló adatrendszer (megnevezések és megnevezettek, valamint a köztük levő kapcsolatok egyértelmű azonosítását nemzeti szinten végző adatrendszer).

A nemzeti névtér tágabb értelmezésű fogalmára használhatjuk a **nemzeti névtár** kifejezést.

A tágabb értelemben vett névtérfogalom alapján elképzelhető névtérrendszer nagyobb szegmenseit, jól elkülöníthető funkcionális egységeit, a potenciálisan együttműködő partnerek helyét mutatja az alábbi ábra.



a szélesebb értelemben vett nemzeti névterek kapcsolatrendszerének sematikus ábrája

5.5 Globális névterek

Korábban röviden bemutatunk több globális névteret is, bár akkor nem hangsúlyoztuk azt, hogy azok nemzetközi névterek lettek volna. Az említettek között voltak dokumentumokat azonosító

rendszerek, mint az ISBN, ISSN vagy az ISAN, a dokumentumok létrehozásában, hozzáférhetővé tételében érintett szereplők, ágensek, helyszínek azonosítását végző rendszerek, mint az ISNI vagy a VIAF, a valós földrajzi helyek, földrajzi nevek világában való eligazodás segítő földrajzi névrendszerek, mint a Geonames szerver, a Getty földrajzi nevek tezaurusza, valamint a Google Maps vagy az OpenStreetMap mint térképszolgáltató. Ezek között vannak, amelyek nemzetközi szabványok is egyben (ISNI, ISBN), így egyfajta jogi támogatás is van mögöttük, de több példa is mutatja azt, hogy ez a jogi támogatás nem nélkülözhetetlen feltétele a nemzetközi elismerésnek, a nemzetközi használatbavételnek.

A globális névterek ugyanazt a koordinációs funkciót töltik be a nemzetközi szinten a nemzeti és nemzetközileg jelentős intézményi névterek világában, mint amit a nemzeti névterek képviselnek a nemzeti szinten, az „alájuk tartozó” lokális, intézményi névterek világán belül. A nemzetközi egyértelműsítő és azonosító rendszereknek az az alapvető célja, hogy egységesítsék, „magasabb szintre” emeljék a nemzeti, regionális szintek kontextusán belül már egyértelműsített nevek azonosítását. Amilyen mértékben sikeresek e tevékenységükben, olyan mértékben válhatnak egyre jobban összekapcsolttá, átjárhatóvá a nemzeti, regionális, lokális szinten létező adatrendszerek. Az persze nagy kérdés, hogy vajon a nemzetközi névterek milyen mértékben fogadják, fogadhatják be a nemzeti, regionális, lokális szintekről érkező adatokat. Bár nem elvileg lehetetlen, de gyakorlati megfontolások miatt nem tűnik valószínű opciónak az, hogy a nemzetközi névterekbe minden névadat bekerüljön a nemzeti névterek szintjéről. Ugyanolyan érvek alapján, mint ahogy a nemzeti névterek és a lokális névterek közt is feltételeznünk érdemes egyfajta szűrő létezését, amely nem enged be minden adatot a „központba” az alacsonyabb, lokális szintekről. Ez a szűrési, kanonizálási elvárás annak köszönhető, hogy minél magasabb szinten akarunk valamilyen közösségi élményt kialakítani, fenntartani, annál biztosabban meg kell szűrniük, hogy milyen „elemeket” teszünk meg, engedünk meg közösségképző tényezőnek.

A globális névterekhez való csatlakozást, a nemzetközi névterekkel való érdemi kooperációt akkor érdemes komolyan venni, amikor a nemzeti névterek már elindultak, működnek, kialakították a lokális, intézményi névterekkel a kétirányú kapcsolataikat.

A személyneveknél és testületi neveknél az ISNI és a VIAF a potenciális partner, a földrajzi nevek esetében pedig az OSM.

6 A névtér fogalmának ontológiai elkötelezettségei

Amikor névterekről beszélünk, amikor névtereket építünk, fontos tisztázni, hogy milyen ontológiai elköteleződéseket teszünk a rendszer modelljének felállításakor. Ez a feltárási munka egyfelől segíti a fejlesztési munkát azzal, hogy értelmezhetővé, egyértelművé teszi a modellben rögzített struktúrát, másfelől láthatóvá, megérhetővé és kezelhetővé teszi azokat a következményeket is, amelyek az alapvető ontológiai elköteleződéseinkből fakadnak.

Mielőtt a névterekkel kapcsolatos fejtegetésekbe vágnánk, be kell vezetnünk egy fontos fogalom párt, amelyre a névterek értelmezésekor bizonyos pontokon hivatkoznunk kell majd.

6.1 A természeti és társadalmi tények különbsége

A társadalomtudományi diskurzusban régóta ismert az a megkülönböztetés, amely szerint a világról gondolkozva kétfajta tényről beszélhetünk, és rájuk utalva kétfajta tényítéletet fogalmazhatunk meg: vannak **természeti tények** és vannak **társadalmi tények** (Searle 1995). Mindkét típus esetében tényállításokat fogalmazhatunk meg, amiket lehet vitatni, módosítani, elutasítani vagy lehet elfogadni. Mindkét típusba tartozó tényállítást élesen el kell (el lehet) határolnunk az értékítéletektől. Mindkét fajta tény képes meghatározó erővel bírni az életünkre. A különbség köztük abban van, hogy a társadalmi tények emberi akartoktól függenek, míg a természeti tények nem. A társadalmi tények társadalmi konstrukciók, míg a természeti tények nem azok. Az a tény, hogy van valahol egy folyó, természeti tény, hiszen ha nem élnének a környékén emberek, akkor is ott folyna, ellenben az a tény, hogy a folyónak 'Duna' a neve, társadalmi tény, hiszen emberek nevezték el így. Az, hogy a folyó egyik oldalán egy hegy magasodik, míg a másikon síkság terül el, természeti tény, viszont az, hogy a helyvidékes oldal az egyik, a síkvidékes rész egy másik ország területének számít, már társadalmi tény.

A természeti, illetve társadalmi tények szembeállítását további terminuspárok segítségével is ki lehet fejezni. Vannak szakterületek, ahol erre a kettősségre a **SNAP** vs. **SPAN** minősítéseket (Grenon & Smith 2004), máshol és máskor a **bona fide** vs. **fiat** ellentétpárt alkalmazzák (Smith & Varzi 2000). Utóbbi fogalom párt például a földrajzi entitások világában található kettősségre húzták rá. Eszerint a fiat objektumok (mint például a közigazgatási határok, a műutak vagy a településnevek) emberi akaratok termékei, míg a bona fide objektumok (mint például a folyók, a hegyvidékek, a hágók vagy a tavak) nem.

A tények kettős természetét figyelembe vevő fogalmi distinkció azért fontos, mert a névtereken belül is szembesülhetünk ezzel a kettősséggel, és ügyelni kell arra, hogy kezelni tudjuk az esetleg ebből fakadó problémákat.

6.2 Névterek és tudásszervezési rendszerek

A névtér fogalmának elemzésekor kiderült, hogy a köznévtér és a tulajdonnévtér el kell választani egymástól. Ez a különbség megnyilvánul abban is, hogy más és más választ adhatunk

arra a kérdésre, hogy milyen tudásszervezési rendszernek minősíthetünk egy konkrét névteret. Amikor a tudásszervezési rendszereket bemutattuk, a rendszer elemeire építhető struktúrák különbségei alapján tipizáltuk őket. A névterek jellemzésekor is ezt a szempontot kell érvényesítenünk. A könnyebb kérdés (melyik tudásszervezési rendszer fogalma kell besorolnunk a tulajdonnévtereket) elemzésével kezdjük a mondandónkat, majd ezután kezdünk bele a második kérdés megválaszolásába (melyik KOS-típusba tartoznak a köznévtér).

6.2.1 A tulajdonnévtér struktúrája

A tulajdonnévtérre nagy általánosságban azt mondhatjuk, hogy azok a legegyszerűbb tudásszervezési rendszer, a terminuslista fogalma alá sorolhatók be, de típusonként változó módon lehet valamivel összetettebb struktúrájuk. A tulajdonnévtér elemei sokféle relációban állhatnak más tulajdonnévtér elemeivel, de ezek nem struktúráként, hanem a rendszerek közti dinamikus kapcsolatrendszerként értelmezhetők.

6.2.1.1 A személynévter struktúrája

A személynévter tiszta terminuslista. A természeti tények alapján nem igazán tudunk struktúrát érvényesíteni az elemei között. A nemi jelleg és a rassz alapján lehet tipizálni az embert, de ezen jegyek alapján nagyon lapos struktúrákat kaphatunk.

A társadalmi tények alapján többféle tipizálásra nyílik mód, és itt már előfordulhatnak olyan köznévtérstruktúrák, amelyek alá a személyek besorolhatók a példány reláció révén. A foglalkozási besorolás kiváló példa erre. A foglalkozások listájából generikus hierarchiát lehet építeni és ezek végpontjaiba el lehet helyezni a konkrét személyeket. Látni kell azonban, hogy ebben az esetben nem a személyek, személynevek halmazán építhető fel a generikus struktúra, hanem föléljük rendelhető – a foglalkozási szerepeket leíró – köznevek halmazán.

- 0. Főcsoport: Fegyveres szervek foglalkozásai
- 1. Főcsoport: Gazdasági, igazgatási, érdek-képviselési vezetők, törvényhozók
- 2. Főcsoport: Felsőfokú képzettség önálló alkalmazását igénylő foglalkozások
- 3. Főcsoport: Egyéb felsőfokú vagy középfokú képzettséget igénylő foglalkozások
- 4. Főcsoport: Irodai és ügyviteli (ügyfélkapcsolati) foglalkozások
- 5. Főcsoport: Kereskedelmi és szolgáltatási foglalkozások
- 6. Főcsoport: Mezőgazdasági és erdőgazdálkodási foglalkozások
 - 62. Csoport: Erdőgazdálkodási, vadgazdálkodási és halászati foglalkozások
 - 621 Erdőgazdálkodási foglalkozások
 - 6211 Erdészeti foglalkozású
 - 6212 Fakitermelő (favágó)
 - 622 Vadgazdálkodási foglalkozások
 - 6220 Vadgazdálkodási foglalkozású
 - 623 Halászati foglalkozások

6230 Halászati foglalkozású

7. Főcsoport: Ipari és építőipari foglalkozások

8. Főcsoport: Gépkelők, összeszerelők, járművezetők

9. Főcsoport: Szakképzettséget nem igénylő (egyszerű) foglalkozások

részlet a foglalkozások egységes osztályozási rendszeréből (FEOR)

Hasonlóképpen lehet a testületi tagság relációja mentén kapcsolatokat találni a személyek és testületek között, és a testületre felépíthető egyfajta típushierarchia, de ebben az esetben egyfelől a konkrét személyek és konkrét testületek közti viszony individuális, partikuláris lesz, másfelől megint csak nem a testületek mint individuumok, hanem a följük rendelhető testülettipológia tartalmazhat valamiféle generikus struktúrát.

A társadalmi tények önkényes konstruáltságából fakadóan olykor a személynévtér esetében is adódhatnak értelmezési problémák. Amikor valamely testületi tagságot mint társadalmi tényként kell rögzíteni a személynévtéren belül,¹⁴ akkor az emberi akaratokra való hivatkozás – mint a társadalmi tény alapja és ontológiai minősége egyben – olykor ellentmondásos eredményekhez, versengő interpretációkhoz vezethet. Szemléltető példaként hivatkozhatunk itt az akadémiai tagság intézményére. Politikai okok miatt a Magyar Tudományos Akadémia 1949-es átszervezésekor sok akadémikust megfosztottak akadémiai tagságától, akiket aztán az újabb politikai fordulat után, 1989-ben rehabilitáltak. Kérdés, vajon mit mondhatunk a kizárt, majd visszavett akadémikusok tagságáról az 1949-1989 közötti időszakokra vonatkozóan. Ebben a negyven évben biztosan nem tartották őket akadémikusoknak a kortársaik, de vajon 1989 után mit mondhattak az 1989 előtti időszakokra vonatkozóan? Lehet-e azt mondani, hogy „minden másként volt”, lehet-e a múltat átértelmezni? A válaszhoz ketté kell választanunk a kérdést.

Az természetesen mindig lehetséges, hogy a múltat átértelmezzük, amennyiben új információ birtokába jutunk. Ez mind a természeti, mind a társadalmi tények esetében elképzelhető. A most vizsgált esetben azonban nem erről van szó. Az akadémiai tagság társadalmi konstrukció, a „kinevezés” az erre hivatott emberek, testület akaratán múlik, és ebben az esetben az az igazi kérdés, hogy vajon lehet-e visszamenőlegesen akaratot érvényesíteni. A válasz nyilván az, hogy nem. Nem mondhatjuk tehát 1989 után, hogy a rehabilitált akadémikusok 1949-1989 között is akadémikusok lettek volna, mert visszamenőleg nem akarhatjuk ezt. Az akarat mindig csak a jövőre irányulhat. Az viszont már egy másik kérdés, hogy a rehabilitáció részeként visszamenőleges hatállyal is lehet rendelkezni, és – például – a szolgálati időbe bele lehet számítani a „kiesett” időszakot is. Ez azért lehetséges, mert ebben az esetben új akaratképzések mentén, „most” hozunk létre egy olyan – új – társadalmi konstrukciót, amely a múltat máshogy veszi figyelembe.

6.2.1.2 A testületi névtér struktúrája

A testületi névtér esetében egy fokkal komolyabb tipizálási lehetőség van a személynévtérhez képest, bár ez teljes mértékben a társadalmi tények világában érvényesíthető. A testületek

¹⁴ Pontosabban a személynévtér és a testületi névtér közötti kapcsolati rendszeren belül.

egyértelműen társadalmi konstrukciók, így szinte minden, velük kapcsolatos tény társadalmi ténynek tekinthető. Talán csak a tevékenységük tér- és időfüggése minősíthető természeti ténynek, minden más adat velük kapcsolatban társadalmi konstrukció. A testület telephelye, megalakulásának időpontja, a nevének változtatása stb. a tér és idő dimenzió konkrét pontjára mutat, ami természeti tény. Az a fogalmi struktúra azonban, amelyben a testületek tipizálhatók, egyértelműen a társadalmi tények világába tartozik.

Itt is igaz, amit a személyekkel kapcsolatban már említettünk: a testülettípológiát generikus hierarchiaként is felfoghatjuk, de ez a struktúra nem a individuális testületek halmazán, hanem a köznevekkel kifejezhető testülettípusok halmazán értelmezhető.

A társadalmi tények már említett másik nagy halmaza a testületek és a személyek közti kapcsolatokat írja le, amelyek egy részét az archontológiai adatok, a másikat az iskolai adatok adják. Ezek mind partikuláris adatok, a tér és idő valamely szegmenséhez köthetők.

A testületek, intézmények, szervezetek nyilvántartását testülettípusonként más helyen, más módon végzik. Más szabályozási rend vonatkozik a civil társadalom társadalmi szervezeteire, a politikai, érdekképviselői szervezetekre, a gazdasági szervezetekre, az államigazgatási szervezetekre vagy az oktatással, neveléssel, illetve kulturális értékeink megőrzésével foglalkozó szervezetekre.

Mivel a testületek lényegük szerint társadalmi konstrukciók, létezésük emberi akaratokra vezethető vissza, ontológiai elköteleződés szempontjából meghatározó, hogy milyen választ adunk arra a kérdésre, mi biztosítja a testületek identitását, mire hivatkozva lehet az azonosságukat tételezni a való életben. A testületeket jellemzi az a tevékenység, amit a tagjai végeznek, azonban mind a testület tagjai, mind a testület által végzett tevékenység változhat az időben. Meddig „él” és meddig azonos önmagával egy testület? Ha egy focicsapat, mondjuk, a Barcelona minden tagja kicserélődik, attól még nyilván ugyanarról a testületről beszélhetünk. De vajon igaz-e ez a könnyűzenei zenekarokra is? Elképzelhető-e, hogy a Beatles minden tagja más, mint az induláskor volt, és még mindig ugyanarról a zenekarról beszélünk?

Másfajta kérdés merül fel a testületek időhöz való viszonyával kapcsolatban. Ha egy testület egy ideig szünetelteti a tevékenységét, majd újakezdi, akkor mennyi idő eltelté után mondhatjuk azt, hogy már más vagy még ugyanazon testületről van szó? Magyarországon 1948-ban egyesült a Kommunista Párt és Szociáldemokrata Párt, 1949-ben beolvadtak a Hazafias Népfrontba az akkor még létező politikai pártok, és négy évtizeden keresztül nem szerepeltek a politika színpadán, majd 1989 környékén egyre-másra bukkantak fel a régi neveiket használva, a jogfolytonosságot deklarálva a régi/új politikai formációk. Megjelent a színen a Szociáldemokrata Párt vagy a Független Kisgazdapárt. Ezek a testületek ugyanazok voltak, ugyanazok maradtak annak dacára, hogy negyven éven át semmilyen életjelt nem adtak magukról? Milyen szempontok alapján, hogyan lehet ilyen kérdésekben dönteni?

Másféle problémákba ütközhetünk a testületek tevékenységi körében, hatáskörében beállt változások és a testületek azonosságának kérdésével kapcsolatban. Magyarországon 2010 után az EMMI-be olvasztották be a korábban önállóan létező Oktatási és Kulturális Minisztériumot, akkor utóbbi létezik 2010 után is, csak éppen egy nagyobb szervezet (az EMMI) része, vagy 2010 óta nincs ilyen testület? Az oktatás területéért 1848 óta sokféle elnevezésű és sokféle feladatkörű minisztérium volt felelős, ahogy azt az alábbi áttekintés mutatja.

- Vallás- és közoktatásügyi Minisztérium (1848–1919)
- Közoktatásügyi Népbiztosság (1919-1919)

- Vallás- és közoktatásügyi Minisztérium (1919–1951)
- Közoktatásügyi Minisztérium (1951–1953)
- Oktatásügyi Minisztérium (1953–1957)
- Művelődésügyi Minisztérium (1957–1974)
- Oktatásügyi Minisztérium (1974–1980)
- Művelődési Minisztérium (1980–1990)
- Művelődési és Közoktatási Minisztérium (1990–1998)
- Oktatási minisztérium (1998–2006)
- Oktatási és Kulturális Minisztérium (2006–2010)
- Nemzeti Erőforrás Minisztérium (2010–2012)
- Emberi Erőforrások Minisztériuma (2012-)

Azt talán még mondhatjuk, hogy az Oktatási Minisztérium és az Oktatásügyi Minisztérium vagy a Művelődési Minisztérium és a Művelődésügyi Minisztérium csak a nevükben különböznek, de ugyanazt a testületet jelölik, de igaz-e az Oktatásügyi Minisztérium és a Művelődési Minisztérium vagy a Nemzeti Erőforrás Minisztérium esetében?

Ezekre a kérdésekre nem tudunk előzetesen, mindenre kiterjedő, általános érvényű döntéseket, elköteleződések hozni, az ilyen – szerencsére azért nem túl gyakori – kérdéses esetekben egyedileg kell majd dönteni a névtérépítés során.

6.2.1.3 A földrajzi névtér struktúrája

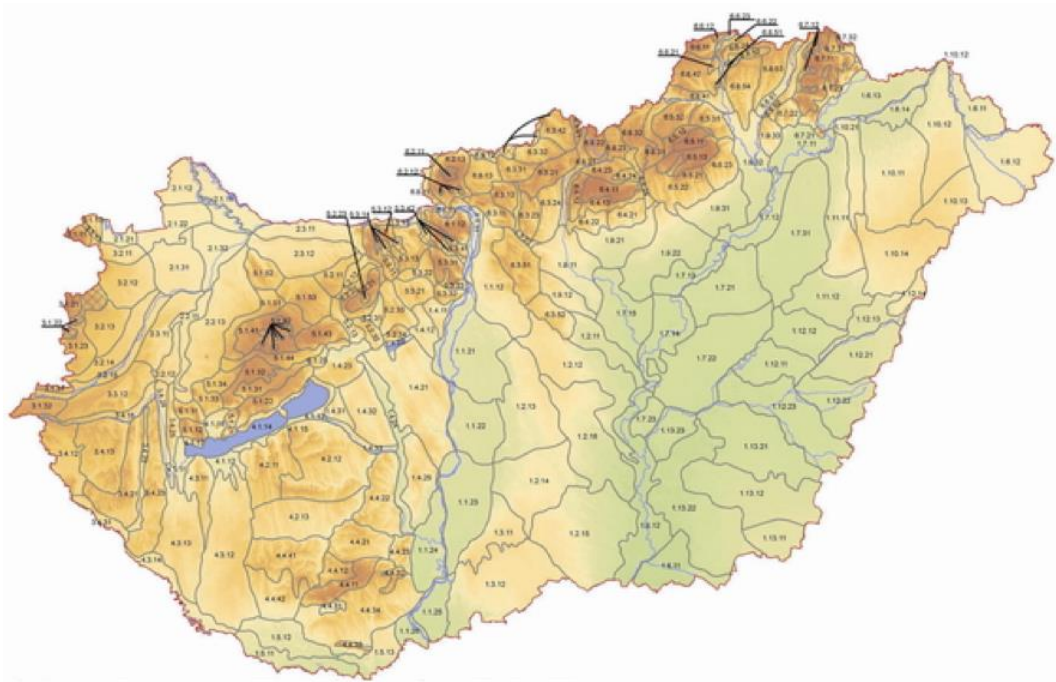
A földrajzi névtér az a tulajdonnévtér, amelynek „komolyabb” struktúrája van (lehet). Egyfelől a földrajzi helyekre többfajta tipizálási lehetőség létezik (a generikus reláció mentén). Első körben – a természeti tények szintjén – meg lehet adni, hogy milyen geometriával jellemezzük az adott földrajzi helyet (lehet pontszerű, vonalszerű, poligonszerű). Még mindig a természeti tények szintjén, de egy másik fogalmi dimenzióban jellemezve a geoindivíduumokat ezt a geometriai besorolást alá lehet bontani természetföldrajzi és vízrajzi fogalmak segítségével. Ezek között a fogalmak között fel lehet építeni egy – nem túl mély és nem túl széles, de azért létező – generikus hierarchiát.

A földrajzi helyek tipizálhatók a társadalmi tények alapján is. A geometriai jellemzők egyértelműségével szemben azonban itt megszűnik az egyértelműség. A társadalmi tények alapján definiált földrajzi helyek többértelművé teszik köztük levő kapcsolatrendszer, hogy megengedik többszörös hierarchiák létezését.

Ha veszünk egy adott geometriával leírt – tehát természeti tényekkel jellemzett – lokációt, akkor arra mondhatjuk azt, hogy önálló falu, mondhatjuk azt, hogy településrész, miután beolvadt a szomszéd városba, de lehetséges az is, hogy ugyanaz a geometria választási szavazóközteret, egyházközséget vagy filiát jelöl. Ezek a minősítések már társadalmi tényeknek felel meg. Nagyon sok szempont szerint feloszthatunk egy adott földrajzi területet, és a különböző felosztási logikák eltérő szerkezeteket eredményeznek. Vegyünk számba pár lehetséges szempontot (néhány alájuk tartozó fogalommal együtt), és szemléltessük a az eltérő szempontokból adódó felosztások területi következményeit.

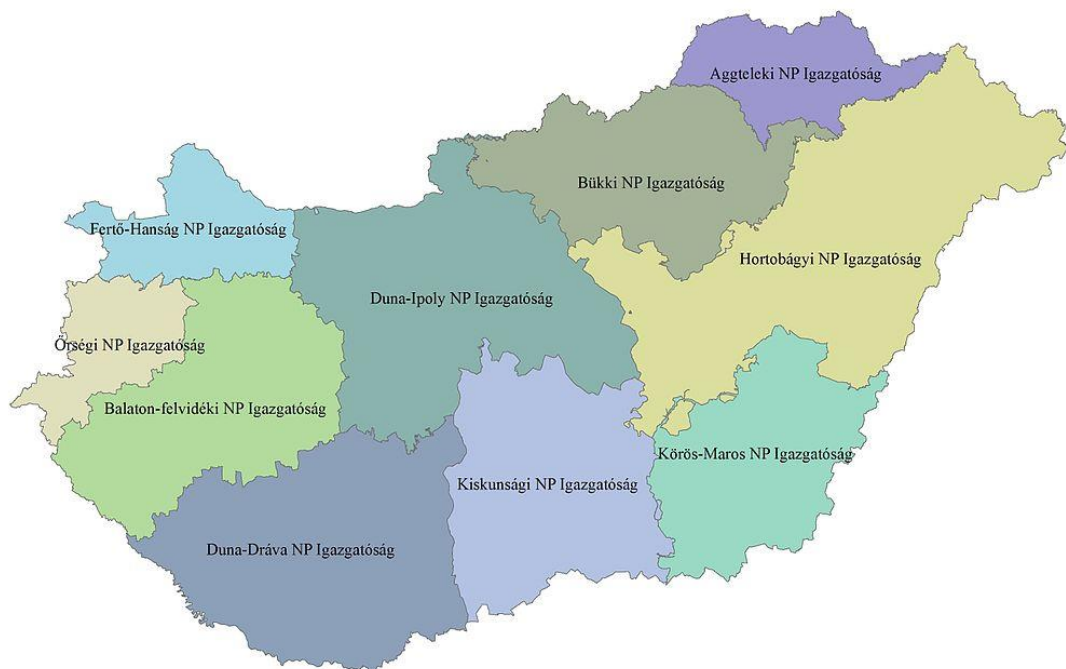
Magyarország területi felosztása:

- földrajzi tájak szerint (nagy-táj, középtáj, kistáj)



Magyarország kistájai

- természetvédelmi körzetek szerint (nemzeti park, tájvédelmi körzet, természetvédelmi terület, tanösvény)



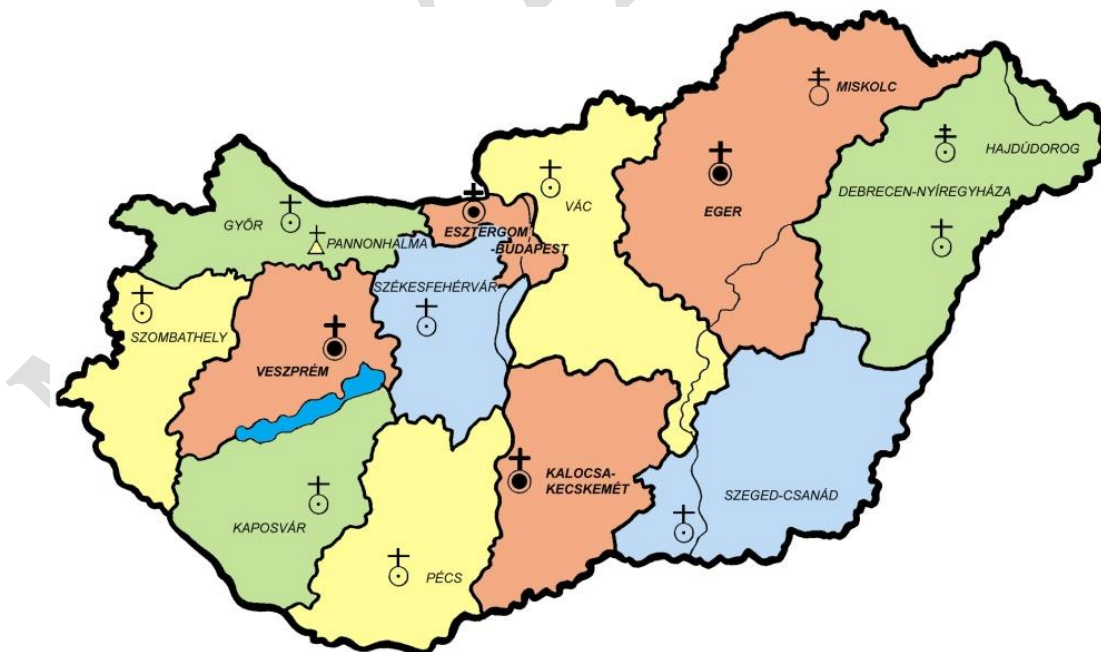
Magyarország nemzeti parkjai

- államigazgatási körzetek szerint (ország, országrész, megye, járás, település, településrész)



Magyarország megyéi

- egyházi körzetek szerint (világegyház, egyháztartomány, egyházmegye, esperesség, egyházközség/plébánia, leányegyház/fília)



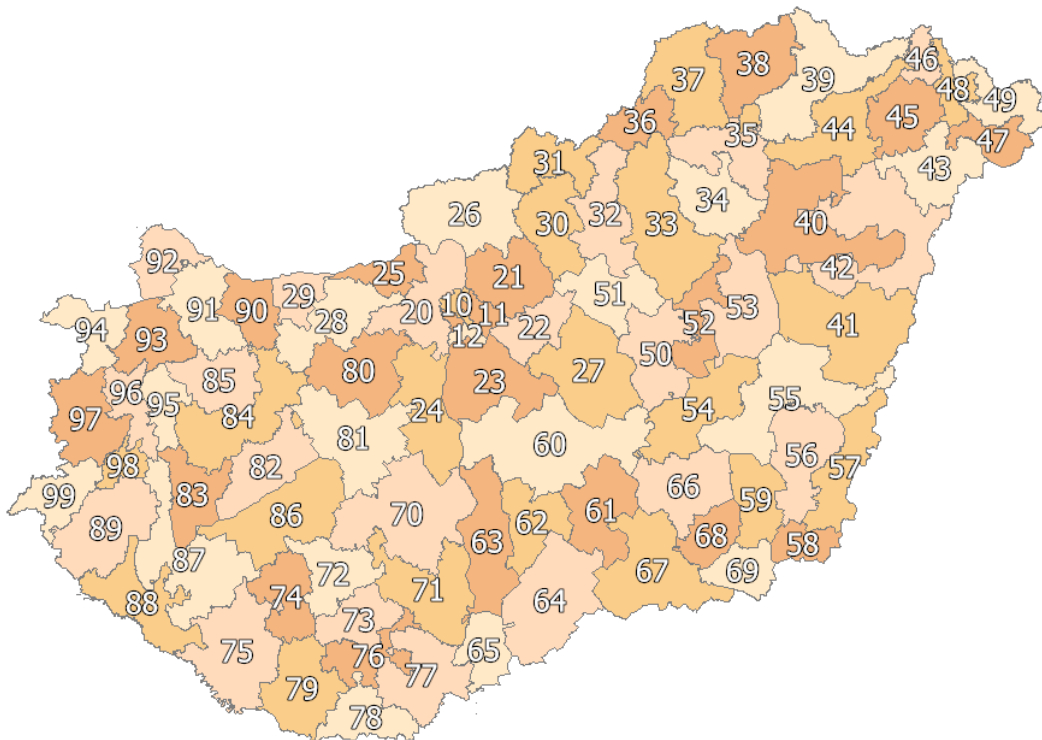
a katolikus egyház egyházmegyéi

- választási körzetek szerint (ország, megye, választókerület, település, szavazókörzet)



Magyarország választókerületei (2012-től)

- postai körzetek szerint (postai régió, két-szamos régió, budapesti kerületek, irányítószám-körzetek)



Magyarország 2-szamos postai irányítószám-körzetei

Mindegyik dimenzióban egymás alá sorolhatók be a használt fogalmak, de nem a generikus, hanem a partitív reláció mentén. Ebből következően a kiválasztott szempont szerint egybetartozó fogalmak között nem generikus, hanem partitív hierarchiákat lehet építeni.

Miért nem a generikus reláció köti össze az egy szempont szerint definiált fogalmakat? Nézzük meg ezt az egyházkormányzati fogalmakat példján keresztül! A generikus kapcsolat egyik fontos minősége az, hogy érvényes rá az IS-A reláció. Amikor azt mondjuk:

kutya IS-A állat

akkor minden tulajdonság, amit az állatra érvényesnek tartunk, érvényes lesz a kutyára is. Ezt fejezzük ki azzal, hogy a kutya fogalmát besoroljuk az állat fogalma alá. Az egyházkormányzati (és az összes többi, itt tárgyalt) fogalmak esetében azonban ez nincs így. Nem mondhatjuk ugyanis azt, hogy:

egyházközség IS-A egyházmegye

A közigazgatási fogalmak esetében sem igaz, hogy

település IS-A járás IS-A megye IS-A ország

Ezeket a fogalmakat nem a generikus reláció kapcsolja össze, hanem a partitív reláció¹⁵ (IS-PART-OF), ami egyfelől tranzitív láncba, másfelől – legtöbb esetben – monohierarchiába köti az azonos szempont szerint definiált fogalmakat.

egyházközség IS-PART-OF egyházmegye

település IS-PART-OF járás IS-PART-OF megye IS-PART-OF ország

Fontos azt is rögzítünk itt, hogy ezek a relációk univerzálék közti relációk, amelyeket figyelmesen el kell választanunk majd a partikulárek közt létező individuális partitív relációktól. Az az állítás, hogy

település IS-PART-OF¹ megye

két univerzálé közti kapcsolatot ír le, míg az a kijelentés, hogy

Dunaújváros IS-PART-OF² Fejér megye

két partikuláre kapcsolatáról szól. Ezt a megkülönböztetést a földrajzi névtérben érvényesíteni érdemes azzal, hogy kétféle partitív relációt engedünk meg.

Említést kell még tennünk azon ritka esetekről, amikor a partitív reláció mentén alárendelt földrajzi entitások nem monohierarchiába rendeződnek, hanem polihierarchiába. A társadalmilag konstruált (bona fide) földrajzi helyek esetében előfordulhat, hogy a geometriájuk

¹⁵ Ezek a partitív tartalmazási viszonyok nem-atomos mereológianak felelnek meg.

(tehát a természeti tények) alapján több szülőjük is van. Ennek legnyilvánvalóbb példája az, amikor egy település nem egyetlen, összefüggő poligonból áll, hanem több részből. A település bona fide objektum (társadalmi konstrukció), míg a határát jelző geometria fiat objektum (természeti tény). Ha egy település területe diszjunkt poligonokból áll, akkor valamelyik része egy másik település poligonjába ágyazódik be, és fizikai-geometriai értelemben a beágyazott poligon a befogadó – másik – település része (ilyen magyarországi település Géderlak, Foktő, Fajsz).



Géderlak: több poligonból álló település

A partitív reláció pontosítását azonban nem csak az itt jelzett ontológiai különbségek miatt érdemes elvégezni, hanem azért is, mert ha a többféle fogalomképzési szempont miatt egy-egy földrajzi hely több „felettesel” rendelkezhet, akkor valahogyan érdemes jelölni azt, hogy a szóban forgó konkrét földrajzi hely konkrét felettese milyen szempont mentén van kapcsolatban vele. Fentebb bemutatuk, hogy egy települést sokféle szempont szerint része lehet egy másik földrajzi helynek (járásnak, megyének, egyházmegyének, tájvédelmi körzetnek stb.). Ha ezeket a rész-egész viszonyokat egyetlen (ugyanazzal a) relációval reprezentáljuk, akkor az egyik adott szempont szerint felépíthető partitív hierarchiát (például a közigazgatási hierarchiát) csak úgy tudjuk összeállítani, hogy a felettesek (partitív fölrendelt kategóriák) minősítését is mindig figyelembe vesszük (lekérdezzük). Ez a művelet elvégezhető, csak sokkal több erőforrást (számítási kapacitást) köt le. Ezen a ponton tehát nincs szükség új ontológiai elköteleződésre, de a rendszer működtetésével kapcsolatos – hatékonysági, performancia – szempontok miatt elképzelhető új – egyébként redundáns – entitások (adatmezők) felvétele a rendszerbe.

6.2.1.4 Az írásműcímtér struktúrája

A nemzeti névterek körébe olyan név- vagy címtereket is be lehet sorolni, amelyek adott típusú kulturális dokumentumok azonosítására alkalmasak a dokumentumok címei és további adattípusok alapján. A filmes világ számára ilyen névtérszegmens lehet a **filmcímtér**, a zene műveknek a **zeneműcímtér**, a könyvtári világ számára pedig az írásos dokumentumok címtere, az **írásműcímtér**. Utóbbira nehéz jó címet adni, és nehéz a határait is megvonni, vagyis azt a kérdést megválaszolni, hogy milyen írásos dokumentumok kezelését kell megoldani egy ilyen írásműcímtér keretein belül és milyenkét nem. Nem minden írásos dokumentum tartozik a könyvtári világ fennhatósága alá, hiszen a levéltárakban is alapvetően írásos dokumentumokat őriznek. Ha megteesszük azt a leegyszerűsítést, hogy a könyvtári kompetenciába – legalábbis első körben – a sokszorosított írásos dokumentumokat soroljuk be, akkor talán kicsivel könnyebb dolgunk lehet. Persze akkor olyan komoly filozófiai, modellezési problémákkal szembesülünk, amelyekkel nem találkozánk, ha a másolás jelenségétől eltekintenénk. Az írásbeliség keretein belül létrehozunk **írásos dokumentumokat**. Az írásos dokumentum adott gondolati **tartalom** rögzítése valamilyen írástechnikával valamilyen **hordozóra**.

Hogy milyen és mennyi tartalmat szükséges rögzíteni ahhoz, hogy dokumentumról beszéljünk, illetve milyen hordozó esetén beszélhetünk egyáltalán írástechnikáról, nehéz kérdés, nincs egyértelmű válasz rá. Még inkább bizonytalanok lehetünk annak a kérdésnek a megválaszolásakor, hogy a dokumentumok közül miket tarthatunk (egyéni jellegű, valamilyen kreativitást megjelenítő) műveknek és miket nem. A dokumentumok közül valamilyen társadalmi/értékelési mechanizmus révén kiválogatódnak azok a **művek**, amelyekhez szerzőket, kreatív alkotókat rendelhetünk, és szerzői jogokat tulajdonítunk nekik, illetve a művek azonosítása végett címeteket adunk nekik. Egy következő lépésben pedig, annak érdekében, hogy ezek a művek minél több befogadóhoz eljuthassanak, sokszorosítjuk őket. Mivel ezek a sokszorosított írásművek bekerülnek abba a kulturális térbe, amelyekben a kulturális dokumentumok megőrzését, fenntartását közgyűjtemények (is) végzik, ezért ezen a ponton megjelenik az az igény, hogy az írásműveket is azonosítani tudjuk valamilyen írásműcímtér segítségével. A kérdés már csak annyi, hogy pontosan mi is az, amit itt azonosítanunk kell.

Ha csak egyedi művek lennének a könyvtárakban (ahogy – kis túlzással – mondhatjuk, hogy a levéltárakban ez a helyzet), akkor könnyebb lenne megválaszolni azt a kérdést, miket és hogyan kell azonosítanunk az írásos dokumentumok világában. Minden egyes dokumentumot külön-külön kellene leírni és egyedi azonosítóval ellátni. A könyvtárak azonban sokszorosított műveket kell, hogy feldolgozzanak, leírjanak, amikor mondhatjuk, hogy a – gépi úton – sokszorosított írásművek (könyvek) nem különböztethetők meg egymástól, ezért ezeket egy csomagban kellene leírni valahogyan, nem pedig külön-külön. Természetesen a könyvtári világ már rég kitalálta a megoldást erre a problémára, hiszen a könyvtárak mindig is elkülönítették a könyvek kiadását, illetve az egyes kiadások konkrét, individuális példányait egymástól. Ilyen módon a könyvtárban mindig is elég volt egyszer felvenni a könyv egy adott kiadásának adatait: a könyv címét, szerzőjét, kiadóját, kiadási idejét stb. Ha egy művet több kiadásban is kiadtak, akkor az nem jelentett különösebb pluszmunkát, ha az új kiadás esetében újra fel kellett venni az alapadatokat, hiszen ezekből nem volt olyan sok. Az új kiadások esetében voltak adatok, amelyek nem változtak az új kiadással (például a mű szerzője vagy címe), de a könyvtári világ számára vállalható "fölösleges teher" volt a már ismert adatok újbóli rögzítése. Azt a veszteséget is

vállalhatónak tartotta a könyvtári világ hosszú időn keresztül, hogy a kiadvány-példány alapú modellben nem tudták igazán jól kifejezni a különböző kiadványok közti kapcsolatot.

Már az 1990-es évek végén megszületett az az új elgondolás, az FRBR-modell, amely a dokumentumok többszintű leírását javasolta (FRBR 2008). Az FRBR négy szint elkülönítését javasolja, hogy egy egységes modellben lehessen kezelni a **példányok (instance)** és kiadványok vagy **megjelenési formák (manifestation)** mellett a **műveket (work)** és a művek különböző **kifejeződései formáit (expression)**. Utóbbi szint behozatalával lehet remélni azt, hogy össze tudjuk kapcsolni egy írásos mű (mondjuk: egy regény) hagyományos könyvkiadványait a mű más formátumú, más médiumokon keresztüli (például hangoskönyv vagy képregény alakban való) megjelenéseitől.

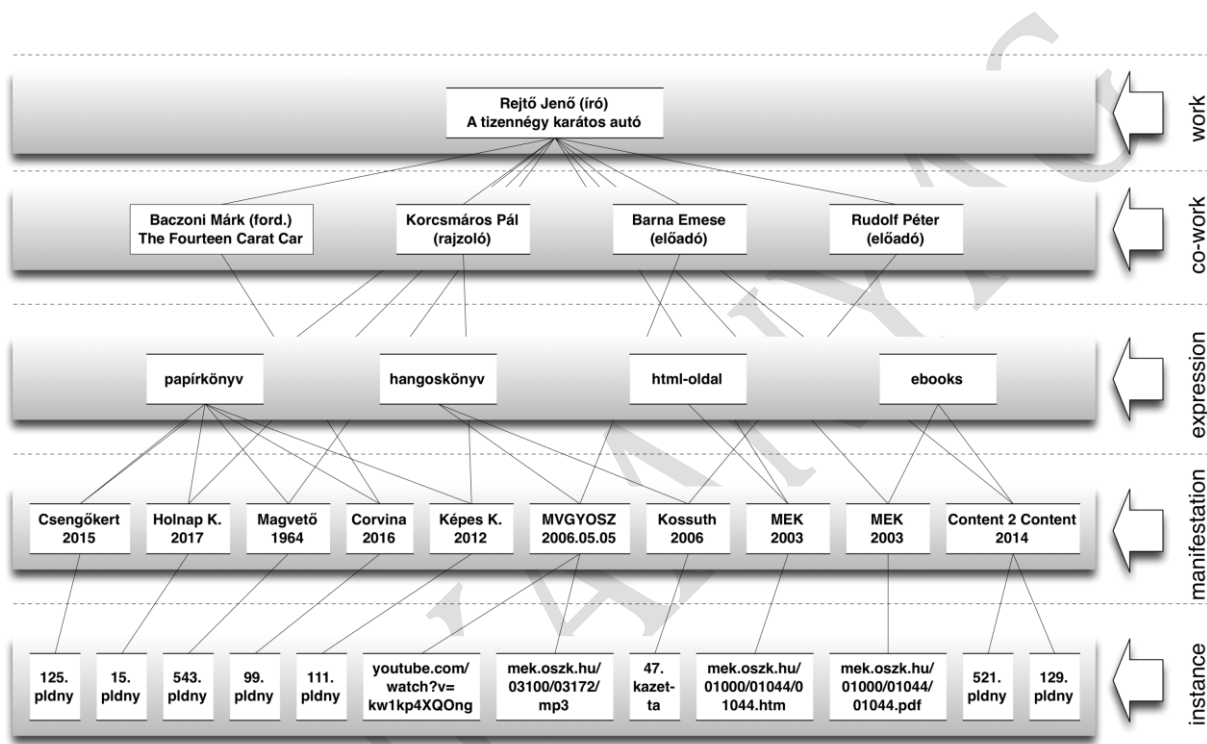
A négy szintű FRBR-modell sokkal jobban alkalmas arra, hogy megragadjuk vele a művek (tartalmak) többféle médiumon keresztüli nyilvánosságához közvetítéseket, illetve a különböző sokszorosítási eseményeken keresztüli megjelentetéseket, de azzal, hogy a modell elméletileg képes lekezelni a világ bonyolultságát, az adatrögzítés gyakorlati feladatát is bonyolultabbá teszi. Sok más tényező mellett ez a nehézség is hozzájárulhatott ahhoz, hogy a könyvtári világ sehol sem fogadta be és nem implementálta azonnal a modellt a hétköznapi munkájába. Az új, feltörekvő szabvány, a Bibframe ugyan elfogadta az FRBR-t a dokumentumok modelljeként, de csak három szintet (a work, a manifestation és az instance fogalmait) alkalmazza.

Pedig az igazán pontos modell felállításához nemhogy elvenni kéne belőle, hanem inkább még egy szintet bele kéne tenni a modellbe. Az FRBR ugyanis nem tudja elég finoman kezelni azt, hogy az eredeti műveket olykor – ugyancsak kreatív alkotói tevékenység révén – átalakítják, valamilyen értelemben más formátumra hozzák. A legnyilvánvalóbb példa erre az írásművek más nyelvre történő lefordítása és kiadása. Ilyen esetben azt sem mondhatjuk igazán, hogy az eredeti mű jelent volna meg egy új kiadványban, hiszen könnyen lehet, hogy a mű szerzője nem is tudná elolvasni, nem értené az idegen nyelvre lefordított művét, de nyilván azt sem mondhatjuk, hogy a fordítással egy új mű keletkezett volna. Ebbe a problémahalmazba tartozik az a jelenség is, amikor egy művet több fordító is lefordít egy másik nyelvre, és ilyenkor kellene tudni jelezni a kiadványok esetében, hogy melyik fordító fordításáról van szó. Ez a megkülönböztetési igény világosan jelzi a fenti problémát. Nyilván ugyanarról a műről van szó, hiszen mindkét fordítás/kiadvány esetében az eredeti mű megegyezik, de a két kiadvány mégis kétféle fordításban (valamilyen értelemben eltérő tartalommal, akár eltérő címeikkel) jelenteti meg az eredeti művet, kétféle szerzői jogi igény keletkezik a kétféle fordításért. Másik példa lehet a hangoskönyvek esete, amikor a szöveget különböző előadók olvassák fel, és ezzel másféle hozzáadott érték keletkezik. Az így kiadott hangoskönyvek nyilván eltérnek egymástól (bizonyos értelemben tartalmilag is), hiszen más színész (felolvasó) hangján szólalnak meg. Nem elég tehát azt mondani, hogy ezeket hangoskönyvként elválasztjuk a papírkönyvektől (az expression szinten), és megengedjük, hogy a hangoskönyveknek legyenek egymást követő kiadásai, mert így nem tudjuk elválasztani egymástól a felolvasás révén keletkezett kreatív hozzájárulásokat, az eltérő színészi munka révén keletkező, eltérő hozzáadott értékeket.

A megoldás az lehet, hogy a mű (work) szintjéhez kapcsolódóan bevezetünk még egy második, **társ-mű (co-work)** szintet, ahol már a származtatott, kiegészítő jellegű alkotói tevékenység eredményét is meg tudjuk ragadni (az eredeti műhöz hozzáragasztva, de egyben el is választva tőle). Javaslatunk szerint az FRBR eredeti négy szintje helyett helyesebb lenne öt szintet elkülöníteni egymástól (ez a megoldás pontosabb modellt eredményezne).

- work
- co-work
- expression
- manifestation
- instance

Az öt szintű modell értelmét (értelmességét) egy olyan ábrával szemléltetjük, ahol a megfelelő szinteken tüntetjük fel egy konkrét mű (Rejtő Jenő 'A tizennégy karátos autó' című regényének) legfontosabb adatait.



az ötszintű dokumentummodell szerkezete egy Rejtő-mű példáján keresztül

Az FRBR modell ahhoz nyújt segítséget, hogy milyen szinten kell megragadni a dokumentum fogalmát. Az írásos dokumentumokkal kapcsolatos másik fontos kérdés annak tisztázása, hogy milyen **szerkezetű dokumentumról** van szó. Szerkezetét tekintve a dokumentum lehet **összetett**, illetve **egyszerű** (vagy gyűjteményes), ami annyit jelent, hogy egy dokumentum (egy könyv, egy periodika adott száma) vagy tartalmaz több dokumentumot vagy nem, tehát vagy többedmagában vagy egymagában áll. Előbbi esetben összetett vagy gyűjteményes dokumentumról beszélhetünk, utóbbi esetben pedig egyszerű dokumentumról. Ha ezt a megkülönböztetést megtesszük, akkor összetett dokumentumként értelmezhetjük egy szerző tanulmánygyűjteményét, egy költő verskötetét és egyszerű dokumentumként az egyedi tanulmányokat vagy verseket, és természetesen meg kell engednünk (és a modellnek kezelnie kell tudni) az olyan eseteket, amikor egy tanulmány vagy egy vers több gyűjteményes dokumentumban is megjelenik. Ebből a szempontból tekintve a periodikumok mind gyűjteményes dokumentumoknak tekinthetők.

A harmadik dokumentumtipizálási szempont szerint arra is érdemes figyelni, hogy bizonyos dokumentumok úgy is összekapcsolhatók egymással, hogy nem egy kiadványban válnak egy

összetett dokumentum részeivé, hanem időben, **sorozatként** kapcsolódnak össze egymással. A könyvsorozatokba foglalt egyedi kiadványok ilyenek, de a periodikumoknak is fontos tulajdonsága ez a sorozati jelleg. Fontos megjegyezni, hogy a sorozatiság is egyfajta gyűjteményességet jelent, csak itt az idő tengelyen kapcsolunk össze dokumentumokat egymással, míg az összetett dokumentumokban egy időben, egyszerre megjelent műveket fogunk össze. Mondhatjuk, hogy a gyűjteményesség kétféle értelme abban a különbségben ragadható meg, hogy a sorozatiság elve mentén az idő eltérő pontjain megjelent dokumentumokat kapcsolunk össze, míg az összetett dokumentumok esetében az idő ugyanazon pontján kapcsolunk össze különböző dokumentumokat. A világot így értelmezve a periodikumok egyszerre összetett és sorozati dokumentumok, a gyűjteményes kötetet összetett dokumentumok, és a könyvsorozatok tipikusan sorozati dokumentumok.

Az írásművek névterében maga az írásmű mint dokumentum áll a középpontban. Az eddigi terminológiánkat használva az írásmű felel meg a névhordozó fogalmának. Az FRRB-modell alapján a dokumentum fogalmát több ontológiai szintre „széttérítve” tudjuk csak jól megragadni. Ezen a ponton elég sok új entitást kell a modellünkbe beemelni. Ezzel szemben az **írásműcím** a nyelvi konstrukciók világába tartozik, amelyek kezelése végett nincs szükségünk további ontológiai elköteleződésekre, hiszen a nyelvi konstrukciókat minden névtér esetében egységesen kezeljük. Ügyelni kell viszont arra, hogy a modell biztosítsa azt, hogy a címeket mindig a megfelelő dokumentumszinthez rendeljük hozzá.

Az írásművek metaadatainak leírásához szükség van arra, hogy rögzítsük az írásművek és a létrehozásokban, terjesztésükben, megőrzésükben stb. szerepet vállaló ágensek (személyek és testületek), valamint a kulturális termelés és fogyasztás helyszínei közti kapcsolatokat, illetve a tartalom hordozójára vonatkozó adatokat. Az írásműcímterén belül azonban ezekre a metaadatokra (pontosabban a metaadatok egy kisebb körére) csak azért van szükség, hogy segítségükkel egyértelműen azonosítani lehessen az írásműveket. Ha van két 'Alkonyat' című könyv, akkor abban dönteni, hogy ugyanarról a műről van-e szó vagy sem, nagy segítséget jelenthet, ha tudjuk, ki a szerzője, kiadója, mikor jelent meg. Az írásműcímterén belül definiálni lehet azt a minimálisan szükséges (és elégséges) adatelemkészletet, amely az egyértelmű azonosítást segíti, és a többi metaadatot nem szükséges „beengedni” ebbe a névtérbe.

Mivel az írásművek azonosításához szükségesnek látszanak a dokumentum-ágens viszonyt leíró metaadatok, ezeket be kell emelni a névtérbe, ami a személynévtér és a testületi névtér felé mutató kapcsolatok kezelését igényli. Ennek a kapcsolatnak az inverze viszont lehetővé teszi azt, hogy mind a személynévtérbe, mind a testületi névtérbe elhelyezhessük a személyes, illetve intézményi bibliográfiákat.

6.2.1.5 Az audiovizuálismű-címterek struktúrája

Az audiovizuális művek esetében nagyon hasonlóak, még ha nem is ugyanazok a modellezési kérdések, a meghozandó döntések, a szükséges ontológiai elköteleződés, mint amelyeket az írásműcímterével kapcsolatban bemutatunk. Mivel mind a tartalom, mind a hordozó minősége más az írásos és audiovizuális művek esetében, ennek nyilvánvalóan vannak következményei a modell egészét tekintve, de nagyon fontos felismerni, hogy az FRBR-modellt alkalmazhatjuk az audiovizuális dokumentumok esetében is. Itt is ugyanúgy hasznos az ötödik szint (co-work) felvétele az FRBR-modellbe (a zeneművek előadói esetében kiváltképp fontos lehet ez a kiegészítés). Vannak sorozatok ebben a világban is (gondoljunk csak a rádió- és

tévésorozatokra), és előfordulhat a gyűjteményes jellegű egymás mellé rendelés is, bár kisebb mértékben, mint az írásos művek esetében. A művek azonosításához itt is szükség van a fontosabb dokumentum-agens kapcsolatok ismeretére, és itt is értékes „melléktermék” lehet a személyes filmográfiák, diszkográfiák létrejötte a személynévtérben, testületi névtérben.

Mivel a kulturális érdeklődések igen nagy hányada az audiovizuális művekre irányul, ezért a **zeneműcímter** és a **filmcímter** létezése valós közönségigényt elégíthetne ki. Ezeket az íráscímter alapján igen nagy részben fel lehetne építeni, és csak kisebb részben lenne arra szükség, hogy az audiovizuális dokumentumok specialitásainak kezelésére további ontológiai elköteleződések tegyünk.

6.2.2 A köznévterek struktúrája

Arra a kérdésre, hogy milyen struktúrája van a köznévtérnek, melyik tudásszervezési rendszer alá sorolható be, nincs egyértelmű válasz. Sokféle köznévteret hozhatunk létre, és hogy melyet, az jelentős mértékben függhet attól, hogy milyen célokra akarjuk használni. A tulajdonnévterek célja a névhordozók és nevek egyértelmű azonosítása, és ezáltal annak a lehetőségnek a megteremtése, hogy a nemzeti névterek azonosítóit átvevő kulturális adatszolgáltatások megnyílnak egymás és a közönség előtt abban az értelemben, hogy a közös ID-kel azonosított személyek, testületek, földrajzi helyek, művek révén sok irányban, sok szálon keresztül bejárhatókká, átjárhatókká válnak ezek az adatrendszerek. Elméletileg a nemzeti köznévterek is betölthetik ezt a funkciót, amennyiben a kulturális adatszolgáltatók ugyanazt a köznévsztruktúrát veszik igénybe az általuk kezelt tartalmak formai és tartalmi jellemzésére. Erre azonban kevésbé számíthatunk, mert a különböző archívumok dokumentumai eltérő adatelemkészlettel írhatók le pontosan, ami megmutatkozik a nemzetközi szabványok különbségeiben is. A tartalmi leírás területén nem kell eltérésekre számítanunk az egyes szakmák, intézmények, archívumok, műtípusok között, hiszen tartalom rövid jellemzésében nincsenek a dokumentumok típusától függő szempontok, ennek ellenére ebben a tevékenységi körben sem alakultak ki közös archívumi gyakorlatok.

A köznévterek modelljét úgy kell megtervezni, és a köznévteret működtető rendszert úgy kell felépíteni, hogy mindenféle köznévteret kezelni, építeni, reprezentálni lehessen benne. Ez nem tűnik igazán nagy kihívásnak, hiszen ha belegondolunk, hogy milyen tudásszervezési rendszereket képzelhetnünk el egyáltalán, ezek a tudásszervezési rendszerek milyen struktúrával rendelkeznek, és mindezek fölé milyen általánosított modellt és struktúrát lehet megtervezni és elhelyezni, akkor elfogadhatjuk azt az álláspontot, miszerint egy általánosított, tetszőleges relációt kezelni képes köznévteret megtervezve, felépítve és működtetve minden köznévtérkezelési igényt ki lehet elégíteni.

Az ETO egy taxonómia, a Köztaurusz egy tezaurusz, a személynévtér, a testületi névtér, a földrajzi névtér, az íráscímter mint tulajdonnévtér egyben egy terminuslista. Korábban bemutattuk, hogy ezek a tudásszervezési rendszerek egységes modellben kezelhetők, és „csak” a rájuk jellemző relációk típusában és számoosságában különböznek egymástól. Ha a köznévteret úgy építjük fel, hogy az általános tudásszervezési rendszer modelljében felsorolt összes relációt felvesszük a modellbe és implementáljuk az adatrendszerbe, akkor az összes tudásszervezési

rendszer igényét képesek lehetünk kiszolgálni. Praktikusan, a fejlesztés első időszakára nézve mondhatjuk, hogy a köznévtérnek valahol a teauruszok és a formális ontológiák között van a helye. Kezelnie kell tudni minden teauruszrelációt, sőt, fel kell tudni venni benne bármilyen új relációt, ebben az értelemben el kell indul a teljesen formális és általános ontológia felé, de arra még nincs szükség, hogy a formális ontológiák következtető képességét is elvárjuk a rendszertől.

A tulajdonnévterekben szükség lehet arra, hogy a köznévtérből listákat, olykor hierarchikus listákat vegyenek igénybe akkor, amikor a névhordozókat, a neveket, a névhordozók, nevek közti kapcsolatokat kell tipizálni, de ezek a köznévstruktúrák sosem lesznek bonyolultak.

A dokumentumleíró rendszerek más adatelemeket, más relációkat is megkövetelhetnek, ám ezeknek az igényeknek már nem kell megfelelnünk a névterek tervezésekor, fejlesztésekor.

7 A névtérdefiníció fejlesztésre gyakorolt hatásai

A névtereken belül megkülönböztetjük a névhordozókat (akiket vagy amiket megnevezünk, akikre vagy amikre a nevekkal referálunk) és a neveket (amelyekkel a referálást végrehajtjuk). E kettősségből fakadóan a névterekben nem csak a névhordozókat, de a neveket is – egyedileg – azonosítani kell.

Névtérről névtérre változhat az, hogy névhordozót vagy nevet azonosít-e a névtér. A KSH Helységnévtára nevet azonosít (külön azonosítóval szerepel benne 'Dunaújváros', Sztálinváros' és 'Dunapentele', holott ezek ugyanazon település különböző nevei). A FÖMI Földrajzi Névtere viszont névhordozót azonosít, és nem azonosítja külön-külön a neveket.

A tulajdonnévtereken belül az eseményeknek kulcsszerepe van. Egyrészt azért, mert a tulajdonnévterekben kezelt névhordozók mindig individuális entitások (partikulárék), amelyeket – elvileg – mindig el kell tudnunk helyezni a tér és idő valamely pontján. Ebben segítenek az események: a névhordozók azonosításához, változásainak nyomon követéséhez mindig eseményeket kell felvennünk az adatbázisban. Az események tulajdonnévtereken belüli központi jelentőségének másik oka az, hogy itt nem neveket, hanem megnevezéseket, megnevezési eseményeket kell kezelnünk. A nevek mint tértől és időtől független entitások (univerzálék) sosem lehetnek alkalmasak önmagukban az egyedi azonosításra. Az individuális névhordozókhöz mint a típusuk instanciáihoz csak úgy tudunk névpéldányokat rendelni, ha a megnevezési aktuson (gyakorlaton), azaz egy eseményen keresztül hozzákötjük a referáló nevet az időhöz (és természetesen a jelölt entitáshoz).

A névtér modelljében minden időfüggő adatot „eseményesíteni” kell, vagyis nem valamely entitás attribútumaként kell felvenni, hanem „önálló” eseményként, és az eseményt kell összekapcsolni az adott entitással (névhordozóval, névvel stb).

A megnevezések kontextusfüggősége, megnevezőiközösség-függése miatt a nevekhez kapcsolt tipológiában valahogyan jelölni kell, hogy milyen megnevezési kontextusa, közösségi státusa van a névnek.

Az események központi kezelése lehetővé teszi azt is, hogy egyfelől egy helyen, egységesen oldjunk meg minden időkezeléssel kapcsolatos problémát (az időpontosság, az összetett események vagy a balról vagy jobbról nyitott végű események kérdését), másfelől egységes módon, egy helyről építhessük fel az idővonal-szolgáltatásokat, amelyek az egyes névterek részeként is, de önállóan is működhetnek.

Minden nyelvi adat az elkülönült nyelvi modulban (és csak ott) van. A köznévterek, a szűkebb értelemben vett tulajdonnévterek minden adata a nyelvi sémán keresztül érhető el.

Szükségesnek látszik, hogy a partitív relációkat, vagy általában a hierarchikus relációkat kifinomultabban kezeljük. Egyfelől el kell különíteni az univerzálék (köznevek) és partikulárék (tulajdonnevek) közti partitív relációkat egyástól, másfelől külön kell kezelni az atomos és nem-atomos partitív relációkat. E két szempont figyelembe vétele már azt követeli meg, hogy legalább négyféle partitív relációt tudjunk kezelni a rendszeren belül. Arra is fel kell készülni, hogy – szükség esetén – különbséget tudjunk tenni a természeti, illetve társadalmi tényeken alapuló (fiat, illetve bona fide) partitív relációk között is.

A földrajzi helyek világában egy-egy entitást több entitással is összekapcsolhatunk a partitív reláció segítségével (egy hely többféle „körzetnek” lehet a része). Ilyen esetekben a partitív reláció ontológiai minősége lehet ugyanaz, és csak a partitív fölérendelt entitás eltérő minősége mentén tudjuk elkülöníteni a kapcsolatokat egymástól: ha egy település része egy megyének is, meg egy egyházmegyének is, akkor ugyanarról a partitív relációról beszélünk, és a két kapcsolat különbsége a megye és az egyházmegye eltérő minőségéből vezethető le. Megfontolandó, hogy az ilyen különbségek jelzésére is önálló relációt vezessünk be (különben olykor nagy és talán fölösleges számításokra lehet szükség az adatok megjelenítésekor).

A geoentitásokat – elvileg – egyértelműen azonosítani lehet a földrajzi térben való pontos elhelyezkedésükkel. A gyakorlatban ezt a lehetőséget akkor tudjuk kihasználni, ha rendelkezésre állnak a megfelelő pontosságú geometriaitartalmazási adatok. Amíg nem ez a helyzet, addig szükségünk van másfajta azonosítási megoldásokra is. Már csak azért is, mert az ember számára nem ez a legbefogadhatóbb azonosítási megoldás. A geoentitások „földrajzi térben való elhelyezésében” kiemelt szerepet játszik az az útvonal (path), amely az „origo”-tól, a Földtől kiindulva megmutatja a kiválasztott geoentitás helyét egy partitív tranzitív lánc segítségével. Ennek során meg kell oldani azt, hogy vannak több felettséggel rendelkező geoentitások (a folyók szolgáltathatnak példát erre, hiszen azok több településen, akár több megyén, több országon keresztül folyhatnak, és így több geoentitásnak is részét képezik).

A partitív relációval mindig a közvetlen fölérendeltje/alárendeltje viszonyt rögzítjük, de a tranzitivitás kihasználásával természetesen mindig meg tudjuk mondani, hogy egy adott geoentitás milyen más entitást tartalmaz a többszörös partitív kapcsolódások mentén. A fejlesztés során meg kell oldani, hogy valamely pont összes alatta levő egyedét egyetlen függvénnyel ki lehessen listázni (klánlistázás).

A tágabb értelemben vett nemzeti névterek, a névtárak tartalmaznak olyan adatszegmenseket, amelyek az egyes névterek entitásai közti kapcsolatokat fejezik ki (archontológiák, bibliográfiák stb.). Ezek egy része olyan egyedi kapcsolat, amelyet egyedileg lehet csak felvenni, ezért az ilyen adatokat csak a nemzeti névtér szerkesztőségi rendszerén belül menedzselni. Vannak olyan adatkapcsolattípusok is, amelyek kezeléséhez nincs szükség szerkesztőségi rendszerre, anélkül is nagy tömegben menedzselhetők (választási, térstatistikai adatok stb.). Az egyedi kapcsolatok szerkesztésére az adminrendszeren belül kell megfelelő felületeket biztosítani, a nagytömegű adatkapcsolatok menedzselését SQL-konzolon keresztül is biztosítani lehet. A fejlesztés során viszont az adminrendszert fel kell készíteni arra, hogy lehetővé tegye azt, hogy a nagytömegű adatokat megjelenítő oldalakat be lehessen fűzni a publikus szolgáltatás menüjébe.

A tulajdonnevek egyszerű terminuslisták. A szűkebb értelemben vett tulajdonnévterek a lehetséges tulajdonnevek listáját tartalmazzák. Ezek a listák a rendszeren belül elkülönített nyelvi modulban vannak. A tágabb értelemben vett tulajdonnévterek, azaz a megnevezések az individuális névhordozókat kapcsolják össze a tulajdonnevekkel, és a nevek „példányosítását” a megnevezési események hozzárendelésével végezhetjük el.

A névtér általános modelljébe a tudásszervezési rendszer (KOS) általános modelljét kell beemelni, és ezt kell implementálni a fejlesztés során. Ha a névtér-projekt bármely más, későbbi fázisában új igény jelenik meg egy speciális tudásszervezési rendszer alkalmazására, akkor a köznévtéren belül ezt mindig meg kell tudni oldani a modell, az adminrendszer bárminemű változtatása nélkül.

Az adminrendszer használata szempontjából nagyon fontos biztosítani, hogy a rendszer bármely pontján a köznévtér elemeire támaszkodva új, szakmai kontextusokat biztosító nézeteket lehessen definiálni és a rendszerbe felvenni. Ennek példái lehetnek az olyan nézetek, amelyeket annak érdekében kell definiálni, hogy általuk biztosítani tudjuk az együttműködő intézmények szakmai elfogultságait (kinek mi számít kiemelt névalaknak, névváltozatnak stb.) vagy, hogy csak az általuk használni kívánt adatszegmenseket láthassák, ne kelljen „kerülgetni” az összes adatot.

8 Hivatkozások

8.1 Irodalom

- Alexiev, Vladimir, Isaac, Antoine, Lindenthal, Jutta (2016). On the composition of ISO 25964 hierarchical relations (BTG, BTP, BTI). *Int J Digit Libr.* 17:39–48
- Bittner, T., Donnelly, M., Smith, B. (2004). Individuals, universals, collections: On the foundational relations of ontology. In Achille C. Varzi – Laure Vieu (eds.): *Formal Ontology in Information Systems. Proceedings of the Third International Conference.* Amsterdam, Berlin, Oxford, Tokyo, Washington DC, IOS Press, 37–48. p.
- Bittner, T., Donnelly, M., Smith, B. (2009). A Spatio-Temporal Ontology for Geographic Information Integration. *International Journal of Geographical Information Science.* 23(6), pp.765-798.
- Clarke, Stella G. Dextre, Smedt, Johan De (2011). *ISO 25964 ISO 25964-1: a new standard for 1: a new standard for development of thesauri and exchange of thesaurus data*, at: https://at-web1.comp.glam.ac.uk/pages/research/hypermedia/nkos/nkos2011/presentations/DextreClarke_DeSmedt_ISO25964.pdf
- Corcho, O., Fernández-López, M., Gómez-Pérez, A. (2003). Methodologies, tools and languages for building ontologies. Where is their meeting point?, in: *Data & Knowledge Engineering*, Vol. 46, 2003, pp.41-64.
- Cruse, D.A. (1986). *Lexical Semantics*, Cambridge University Press
- Dancs Szabolcs (2017). Bemutatkozik az ISNI, a nemzetközi szabványos névazonosító. *TMT*, vol. 64, No. 4, 199-2007.o.
- Davidson, D. (1967), 'The Logical Form of Action Sentences', in N. Rescher (ed.), *The Logic of Decision and Action*, Pittsburgh: University of Pittsburgh Press, pp. 81–95; reprinted in *Events*, pp. 3–17, and in Davidson 1980, pp. 105–122.
- Davidson, D. (1969). 'The Individuation of Events', in N. Rescher (ed.), *Essays in Honor of Carl G. Hempel*, Dordrecht: Reidel, pp. 216–34; reprinted in *Events*, pp. 265–283, and in Davidson 1980, pp. 163–180.
- Doctorow, C. (2001). *Metacrap: Putting the torch to seven straw-men of the meta-utopia*, at: 2001.08.26. at: <http://www.well.com/~doctorow/metacrap.htm>
- Egenhofer, M. J., Mark, D. M. (1995). Naive geography. In A. U. Frank –W. Kuhn (eds.). *Spatial Information Theory - A Theoretical Basis for GIS (COSIT'95)*. Berlin, Heidelberg, 1995, Springer, 1–15. p.
- Feilmayr, Ch., Wöß, W. (2016). An analysis of ontologies and their success factors for application to business. *Data & Knowledge Engineering*, Vol. 101, pp.1-23.

- Furnas, G. W., Landauer, T. K., Gomez, L. M., and Dumais, S. T. (1987). The vocabulary problem in human-system communication. in: *Commun. ACM* 30, 11
- Gangemi, A., Guarino, N., Masolo, C., Oltramari, A. (2001). Understanding top-level ontological distinctions. In *Proc. of IJCAI 2001 workshop on Ontologies and Information Sharing*.
- Green, Rebecca, Bean, C.A., Myaeng, Hyon, S. (2001). *The Semantics of Relationships: An Interdisciplinary Perspective*, Dordrecht: Kluwer
- Grenon, P., Smith, B. (2004). SNAP and SPAN: Towards Dynamic Spatial Ontology. *Spatial Cognition and Computation*, 4(1), pp.69-103.
- Gruber, T. (1993). Towards principles for the design of ontologies used for knowledge sharing. In N. Guarino – R. Poli (eds.): *Formal Ontology in Conceptual Analysis and Knowledge Representation*. Deventer, The Netherlands, Kluwer Academic Publishers
- Gruber, T.R. (1993). A translation approach to portable ontology specification. 5. évf. (1993), *Knowledge Acquisition*, 199–220. p.
- Guarino, N., Welty, C. (2000). A Formal Ontology of Properties. *Proceedings of 12th Int. Conf. on Knowledge Engineering and Knowledge Management Lecture Notes of Computer Science*, Springer Verlag
- Guarino, N., Welty, C. (2001). Supporting ontological analysis of taxonomic relationships. *Data & Knowledge Engineering* 39 pp 51-74.
- Guarino, N., Welty, Ch. (2002). Evaluating ontological decisions with ontoclean. *Communications of the ACM*, Vol.45. No.2., 61–65. p.
- Guy, Marieke, Tonkin, Emma (2006). Folksonomies. Tidying up Tags?, in: *D-Lib Magazine*, January 2006, Vol.12., No.1, at: <http://www.dlib.org/january06/guy/01guy.html>
- Habibzadeh, Farrokh; Yadollahie, Mahboobeh (2009). The problem of "Who". *The International Information & Library Review*. 41 (2): 61–62.
- Hill, Linda L. (2000). Core elements of digital gazetteers: Placenames, categories, and footprints. In *ECDL '00: Proceedings of the 4th European Conference on Research and Advanced Technology for Digital Libraries*. London, Springer-Verlag, 280–290. p.
- Kripke, S. (2010). *Megnevezés és szükségszerűség*. Budapest: Akadémiai Kiadó
- Kuti Judit, Varasdi Károly, Cziczelszki Judit, Gyarmati Ágnes, Nagy Anikó, Tóth Marianna, Vajda Péter (2006). Igei wordnet és igei eseményszerkezet ábrázolása, *IV. Magyar Számítógépes Nyelvészeti Konferencia*, Szeged, 2006, 97–108. at: http://www.nytud.hu/oszt/korpusz/resources/kuti_et_al_WN_2006.pdf
- Lappalainen, Mikko, Frosterus, Matias, Nykyri, Susanna (2014). Reuse of library thesaurus data as ontologies for the public sector. In *IFLA WLIC 2014*, 16-22.p.
- Lowe, E. J. (1998). *The possibility of metaphysics*. Oxford: Clarendon Press.
- Lyons, J. (1977). *Semantics*, Cambridge University Press

- Marlow, Cameron , Naaman, Mor , Boyd, Danah and Davis, Marc. (2006). Ht06, tagging paper, taxonomy, flickr, academic article, to read. In *HYPertext '06: Proceedings of the seventeenth conference on Hypertext and hypermedia*, pages 31–40, New York, ACM Press
- Mathes, Adam (2004). *Folksonomies - Cooperative Classification and Communication Through Shared Metadata*, UIC Technical Report, at: <http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html>
- Scott A. Golder, Bernardo A. Huberman, Usage Patterns of Collaborative Tagging Systems, in: *Journal of Information Science*, 2006, 32(2). 198-208. at: <http://www.hpl.hp.com/research/idl/papers/tags/tags.pdf>
- Searle, J.R. (1995). *The Construction of Social Reality*. The Free Press
- Shirky, C. (2005). *Ontology is Overrated: Categories, Links, and Tags*, at: http://shirky.com/writings/herecomeseverybody/ontology_overrated.html
- Simons, Peter (2000). *Parts: A Study in Ontology*. Oxford University Press
- Smith, B., Varzi, A. (2000). Fiat and Bona Fide Boundaries. *Philosophy and Phenomenological Research*, 60, pp.401–420.
- Sowa, J.F. (2000). *Knowledge Representation : Logical, Philosophical, and Computational Foundations*. Brooks Cole Publishing Co.
- Speroni, Pietro (2006). *On Tag Clouds, Metric, Tag Sets and Power Laws*, at: <http://blog.pietrosperoni.it/2005/05/25/tag-clouds-metric/>
- Staab, S., Studer, R. (2004). *Handbook of Ontologies*, Springer Verlag
- Star, Susan Leigh (1996). *Slouching toward Infrastructure*, at: <http://is.gseis.ucla.edu/research/dl/star.html>
- Strawson, P. F. (1959). *Individuals. An Essay in Descriptive Metaphysics*. London and New York: Routledge
- Syi (2007). *Egyben az egész. Egytől egyig*. Typotex Kiadó
- Syi (2011). Navigáció. In: Kangyal András, Laufer László (szerk.). *Gépéret: Interfész, interakció, navigáció*. Budapest: L'Harmattan Kiadó, pp. 196-294.
- Syi (2014). syi.hu/cse. L'Harmattan – Könyvpont Kiadó
- Szabó Dóra. *Felhasználói segédlet a Web of Science adatbázis és a ResearcherID használatához*. Szent István Egyetem Kosáry Domokos Könyvtár és Levéltár. at: http://lib.szie.hu/sites/default/files/files/SZIE_KDKL_WoS_ResearcherID_utmutato.pdf
- Szabó Orsolya, Tóth Eszter (2008). *ISMN útmutató*. Országos Széchényi Könyvtár Magyar ISBN és ISMN Iroda, at: http://www.ismn-international.org/download/Manual_translations/HU_ISMN_utmutato.pdf
- Szakadát István (2005). Réteges struktúra, alaprelációk. In: *III. Számítógépes Nyelvészeti Konferencia*. Szeged, Magyarország, Szeged, pp. 20-28.

- Szakadát István, Szóts Miklós, Gyepesi György, Varasdi Károly, Ungváry Rudolf, Simonyi András, Gyarmathy Zsófia, Szaszko Sándor, Szeredi Dániel (2006). MEO ontológiamodel. In: Alexin Zoltán, Csentes Dóra (szerk.) *IV. Magyar Számítógépes Nyelvészeti Konferencia*, Szeged, pp. 377-383.
- Szakadát, I., Szóts, M., Gyepesi, Gy. (2006). MEO - Ontology Infrastructure. In Gabor Magyar – Gabor Knapp – Wita Wojtkowski – Gregory Wojtkowski – Joze Zupancic – Stanislaw Wrycza (eds.): *Advances in Information Systems Development : New Methods and Practice for the Networked Society, Proceedings Information Systems Development*. Springer
- Szóts Miklós, Lévay Ákos (2005) Szerepfogalmak az ontológiákban - az OntoClean metodológia továbbfejlesztése, in: Alexin Z., Csentes D., *III. Magyar Számítógépes Nyelvészeti Konferencia*, Szeged
- Thomasson, A.L. (2004). Methods of categorization. In Achille C. Varzi – Laure Vieu (eds.): *Formal Ontology in Information Systems. Proceedings of the Third International Conference*. Amsterdam, Berlin, Oxford, Tokyo, Washington DC, IOS Press, 3-16. p.
- Toffler, Alvin (2001). *Harmadik hullám*, Budapest: Typotex
- Tózsér János (2009). *Metafizika*, Budapest: Akadémiai Kiadó
- Ungváry Rudolf, Orbán Éva (2001). *Osztályozás és információkeresés, I-II.* Budapest: OSZK
- Vander Wal, Thomas (2005a) *Off the Top: Explaining and Showing Broad and Narrow Folksonomies*, 2005.02.21, at: <http://www.vanderwal.net/random/category.php?cat=153>
- Vander Wal, Thomas (2005b). *Folksonomy Definition and Wikipedia*, 2005.11.05, at: <http://www.vanderwal.net/random/entrysel.php?blog=1750>
- Vander Wal, Thomas (2007). *Folksonomy*, 2007.02.02, at: <http://vanderwal.net/folksonomy.html>
- Vander Wal, Thomas (2008). *Getting to know collective and collaborative*, 2008.03.29, at: <http://www.personalinfocloud.com/2008/03/getting-to-know.html>
- Varzi, A.C. (1996). Parts, wholes, and part-whole relations: The prospects of mereotopology. In *Data and Knowledge Engineering*, 20. évf. 177-198. o.
- Vendler, Z. (1957). 'Verbs and Times', *Philosophical Review*, 66, pp.143-60.
- Will, Leonard (2012). The ISO 25964 Data Model for the Structure of an Information Retrieval Thesaurus, *Bulletin of the American Society for Information Science and Technology* – April/May 2012 – Volume 38, Number 4, at: https://www.asis.org/Bulletin/Apr-12/AprMay12_Will.pdf
- Zemach, E.M. (1970). Four Ontologies. *The Journal of Philosophy*, 47(8), pp.231-247.
- Zeng, Marcia Lei. *Update: The state of KOS in the Linked Data movement*, at: http://www.getty.edu/research/tools/vocabularies/7_itwg_kos_lod_update_zeng.pdf

- Zvolenszky Zsófia (2015). Pofonegyszerű, kedves Kripkém? *Műút*, <http://www.muut.hu/?p=15361>

8.2 Szabványok, szolgáltatások

- Academia.edu, at: <https://www.academia.edu/>
- BIBFRAME, *Bibliographic Framework Initiative*, at: <https://www.loc.gov/bibframe/>
- BNF, at: <http://catalogue.bnf.fr/index.do>
- DOI, at: <https://www.doi.org/>
- FRBR (2008). *Functional Requirements for Bibliographic Records, Final Report*, IFLA, at: https://www.ifla.org/files/assets/cataloguing/frbr/frbr_2008.pdf
- FRISAD (2010). *Functional Requirements for Subject Authority Data (FRISAD) A Conceptual Model*, at: <https://www.ifla.org/files/assets/classification-and-indexing/functional-requirements-for-subject-authority-data/frisad-final-report.pdf>
- GDN, at: <http://www.dnb.de/EN/Standardisierung/GND/gnd.html>
- *Getty Thesaurus of Geographic Names*, at: <http://www.getty.edu/research/tools/vocabularies/tgn/about.html>
- *Getty Vocabulary Program ontology*, at: <http://vocab.getty.edu/ontology>
- Google Maps, at: <https://www.google.hu/maps>
- Google Scholar, at: <https://scholar.google.hu/>
- IMDb, at: <http://www.imdb.com/>
- ISAN, at: <http://www.isan.org/>
- ISBN, at: http://www.oszk.hu/isbn_ismn
- ISMN, at: http://www.oszk.hu/isbn_ismn
- ISNI, at: <http://www.isni.org/>
- ISO 25964 – the international standard for thesauri and interoperability with other vocabularies, at: <http://www.niso.org/schemas/iso25964>
- ISRC, at: <http://isrc.ifpi.org/en/>
- ISSN, at: <http://www.oszk.hu/issn>
- ISTC, at: <http://www.istc-international.org/>
- ISWC, at: <http://www.iswc.org/>
- Magyar Egységes Ontológia (MEO) projekt dokumentumai, at: <http://ontologia.hu>

- MSZ 3418-87 Magyar nyelvű információkereső teauruszok szerkezete, részei és formái. Budapest, Magyar Szabványügyi Hivatal, 1987 [Hungarian Standard of Thesauri]
- MTMT, at: <https://www.mtmt.hu/>
- NBN, at: <http://nbn.depositolegale.it/>
- ODT, at: <https://doktori.hu/>
- OpenStreetMap, at: <https://www.openstreetmap.org/>
- ORCID, at: <https://orcid.org/>
- PIM, at: <https://opac-nevter.pim.hu/en/>
- ResearcherID, at: <http://www.researcherid.com/>
- ResearchGate, at: <https://www.researchgate.net/>
- Scopus, at: <https://www.scopus.com/home.uri>
- ULAN, at: <http://www.getty.edu/research/tools/vocabularies/ulan/>
- VIAF, at: <https://viaf.org/>
- Web of Science, at: <http://webofknowledge.com/>
- Wikidata, at: https://www.wikidata.org/wiki/Wikidata:Main_Page