

Köztaurusz a 21. században? A Köztaurusz szabványossági vizsgálata

Köztaurusz in the 21st century? Standardization exam study of the Köztaurusz

Bódog András
bodog.andras@vmk.hu
Vörösmarty Mihály Könyvtár,
Székesfehérvár
könyvtárinformatikus

Beérkezett: 2024.01.12.
Elfogadva: 2024.03.05.
Publikálva: 2024.03.26.

Magyarországon a Köztaurusz elnevezésű általános információkereső tezaurusz szolgál a nemzeti könyvtárban, illetve a közkönyvtárakban gyűjtött dokumentumok tárgyszavazására. A tárgyszórendszerként is szolgáló tezaurusz évtizedeken át épült, jelenleg is folyamatosan gyarapodik a tárgyszójegyzék, ám mára már hatályon kívül helyezett szabványok figyelembe vételével készült. A tanulmány azt vizsgálja, hogy a Köztaurusz mennyire felel meg a hatályos nemzetközi (ISO) tezauruszszabványnak. Mivel a Köztaurusz alapvető fontosságú az Országos Széchényi Könyvtár és számos közkönyvtár mindennapi munkájában és szabványos tartalmi feltárásában, a rendszer mindenkor szabványossága elengedhetetlen annak továbbfejlesztése érdekében. A Köztaurusz modernizálásának és a kor kihívásainak megfelelő szinten tartásának legelső lépése ez a szabványossági vizsgálat.

tárgyszórendszer, tezaurusz, Köztaurusz, tartalmi feltárás

In Hungary, a general thesaurus called Köztaurusz (Pubtaurus means Public Thesaurus) is used to index documents collected in the national library and public libraries. The thesaurus, which also serves as a subject heading scheme, has been built over decades and is still being updated, but it has been built according to standards that have now been withdrawn. This study examines the extent to which the Köztaurusz complies with the currently valid international (ISO) thesaurus standard. As the Köztaurusz is essential for the daily work and standardized topical indexing of the National Széchényi Library and many public libraries, the continued standardization of the system is essential for its further development. The first step in modernizing the Köztaurusz and keeping it up to date with the challenges of our times is this standardization exam study.

subject heading system, thesaurus, Köztaurusz, topical indexing

Bevezető

A múlt század második felében megkezdődött intenzív tezauruszfejlesztés sok tekintetben kihozta analóg környezetben a tartalmi feltárás akkori lehetőségeiben rejlő maximumot a tárgyszavak közötti szemantikus kapcsolatok feltüntetését illetően. A papíralapú tezauruszok messzemenően kielégítették a korabeli visszakeresés támasztotta igényeket, azonban a könyvtári informatika bevezetésével, majd fejlődésével a kontrollált tárgyszójegyzéken túlmutató funkciók nagy része általánosságban már nem került implementálásra, bár maga a tezauruszszemlé-

Cite as/Így hivatkozd: Bódog, A. (2024) Köztaurusz a 21. Században? A Köztaurusz szabványossági vizsgálata / Köztaurusz in the 21st century? Standardization exam study of the Köztaurusz, Central European Library and Information Science Review (CELISR), 1(1), p. 1-19.
<https://doi.org/10.3311/celistr.36967>



let az informatikai értelemben vett ontológiák képében vissza-visszatért a szakmai köztudatba. Oka még a teauruszok és más tartalmi feltárórendszerek látszólagos tetszalottságának az is, hogy a tömeges digitalizálás révén keletkező teljes szövegű közgyűjteményi tartalmakban való kereshetőség az átlagos felhasználó számára többnyire feleslegessé tette a korábbi, némi hozzáértést kívánó információ-visszakereső technikák alkalmazását. Manapság azonban újabb forradalom kezdetén állunk: a digitális térben olyan mértékű lett ezen állományok exponenciálisan növekvő mérete, hogy már a digitális tartalmak minimális átláthatóságát biztosító rendszerezéshez ismételten szükség lett a tartalmi feltárás olyan korszerű megoldásaira, mint a facettás osztályozás, a teauruszok és az ontológiák alkalmazása, illetve – mivel már az emberi munkaerő képtelen lépést tartani az új tartalmak feldolgozásával – a gépi tárgyszavazás. A modern kihívásokra válaszul a Finn Nemzeti Könyvtár korszerűsítette általános teauruszát, és erre alapozva korszerű kapcsolt adat (linked data) ontológiát épített belőle, amely a finn könyvtári rendszerben és az aggregált közgyűjteményi tartalmakban általánosan használt tárgyszórendszer alapját képezi, újabban pedig egy Európa-szerte széles körben adaptált gépi tárgyszavazó rendszer szótárállományaként szolgál. Felmerül a kérdés, hogy a finn általános ontológia (YSO) magyar párhuzama, a Köztaurusz is alkalmas-e egy hasonló modern feladatkör ellátására. Tanulmányomban erre keresem a választ, és ennek keretében az első részben a Köztaurusz jelen állapotát vizsgálom a jelenleg hatályos teauruszszabványnak, egyúttal a nemzetközi gyakorlat diktálta korszerű információkereső teaurusznak/tárgyszórendszernek való megfelelés szempontjából.

Fontosabb felhasznált források, kutatási módszertan

A Köztaurusz Relexen belüli vizsgálatára a MOKKA szerverére telepített szolgáltatást használtam (<https://mokka.hu/relex/guest.html>). A Közgyűjteményi Szabványosítási Iroda részére 2022 őszén a Relexben készült Gyakorló Köztaurusz néven egy másolat a Köztaurusz akkori teljes állományáról kísérleti felület céljára. Ebben az állományban vizsgáltam meg a Relex alapértelmezett teauruszkezelési beállításait, illetve a jövőben itt lehet a valódi implementációt megelőzően a gyakorlatban kipróbálni a tervezett változtatásokat. További létfontosságú információt szolgáltatott a Relex felhasználói kézikönyve, illetve a Köztaurusz régi statikus webes változatának felületén megtalálható részletes ismertető. Kihagyhatatlan információforrások a témában Ungváry Rudolf gazdag publikációs listájából azok a tanulmányok, amelyek a Köztaurusz fejlődéstörténetével, illetve a magyar teauruszkészítés és -szabványosítás elméleti és gyakorlati kérdéseivel foglalkoznak. A teauruszokat érintő relációelméletben Ungváry művei mellett Horváth Tibor, Varga Dénes és Szakadát István munkáit használtam fel.

A jelenleg hatályos teauruszszabvány az ISO 25964 szabványcsomag, amelynek 2011-ben megjelent első része a szabványos egy- és többnyelvű információkereső teaurusz fejlesztésének és fenntartásának részletekbe menő kérdéseivel foglalkozik. A 2013-ban megjelent második rész a más szótárakkal való kompatibilitást vizsgálja, meghatározva egyúttal a Köztauruszt érintő jövőbeni feladatok normatív követelményeit. A Köztaurusz jelenlegi változatának megjelenése idején hatályos szabványoknak való megfelelés ellenőrzéséhez felhasználtam a magyar teauruszszabvány (MSZ 3418) 1976-os első és 1987-es második kiadásait, illetve az ezek alapjául szolgáló ISO 2788 nemzetközi szabvány 1974-es és 1986-os kiadásait. A vizsgálat folyamán a szabványokban leírt előírások megfelelését

„A modern kihívásokra válaszul a Finn Nemzeti Könyvtár korszerűsítette általános teauruszát, és erre alapozva korszerű kapcsolt adat (linked data) ontológiát épített belőle...”

„Az információkereső teaurusz egy ún. kontrollált és strukturált szótár, amelynek szócikkei a teauruszcikkek.”

ellenőriztem le mind formai, mind logikai síkon. Az eredményeket a szabványokból, illetve a Köztauruszból vett példákkal illusztráltam. A példaként alkalmazott valós vagy illusztratív teauruszelemeket dőlt betűvel jeleztem a szövegben. A Relex MARC-adatcsere-formátumban exportálható állományának vizsgálatához egyrészt készítettem egy MARC-exportot a Gyakorló Köztauruszról, másrészt egy Köztauruszt aktívan alkalmazó megyei könyvtár katalógusában a Relexből betöltött teauruszállomány MARC-rekordjait is figyelembe vettem. Az összehasonlítás alapját a Köztaurusz és a Relex gyakorlati formátumának számító HUNMARC. *A besorolási rekordok adatcsere-formátuma* című könyvtári szabályzattervezet képezte, illetve a MARC 21 authority (besorolási rekordok) (Library of Congress 2022a), illetve osztályozási (classification) (Library of Congress 2022b) formátumának naprakész dokumentációja a Library of Congress weboldalán. Ungváry Rudolf a teauruszok MARC-kompatibilitásáról szóló publikációi is igen fontos hazai szakirodalmi források. A tanulmány szövegében félkövér kiemeléssel jelölöm a hatályos ISO-teauruszszabvány szakkifejezéseinek magyar megfelelőit.¹

1. A Köztaurusz jelene

A 2001-ben megjelent és azóta rendszeresen karbantartott Köztaurusz az Országos Széchényi Könyvtár és számos közkönyvtár által használt általános célú információkereső teaurusz, amely elsősorban tárgyszórendszerként áll a könyvtárak rendelkezésére. Szerkesztése és karbantartása a Relex nevű szoftverrel történik, a naprakész tárgyszavak kikeresése is ennek online felületén lehetséges. A 2011-ben, a MOKKA projekt keretében üzembe helyezett kurrens Relex előnye, hogy fejlesztése az akkor hatályos teauruszszabvány (MSZ 3418:1987) figyelembe vételével történt, ennél fogva képes a teauruszformátum széles körű relációinak kezelésére. A könyvtárszakma szempontjából további fontos előny volt a teauruszszabvány szerinti Word-formátumú szöveges bemenet és a MARC-kimenet lehetősége, utóbbi a teljes szótárra vagy az egyes lexikai egységekre vonatkoztatva, így az integrált könyvtári rendszerek és egyéb alkalmazások könnyedén átvehetik szabványos formában a lexikai egységeket, vagyis magukat a teauruszcikkek. A fejlesztés azonban az OSZK akkori igényeinek figyelembevételével történt, ezért a könyvtári szabályzatként tervezett, de hivatalosan meg nem jelent HUNMARC. A besorolási rekordok adatcsere-formátuma köré épült (Ungváry 1998). A Köztaurusz esetében az egyetlen különbség a tervezett HUNMARC-authority formátum és a MARC 21 között, hogy előbbibe tervezett módon bele kívánták építeni a magyar teauruszszabvány teljes relációkészletét. A bevezetés minden bizonnyal ösztönözte volna a hazai könyvtárinformatikai piac fejlesztését is abba az irányba, hogy a hazai integrált könyvtári rendszerek ki tudják használni a HUNMARC-ban kódolt általános **információkereső teaurusz** minden előnyét. Mivel azonban ez a könyvtári szabályzat nem jelent meg, fejlesztések sem történtek ilyen irányba. Ennek ellenére a Köztaurusz HUNMARC-kimenete az eredetileg tervezett módon kódolja a teauruszkapcsolatokat. A fő probléma abban gyökerezik, hogy nincs olyan integrált könyvtári rendszer, amely ezeket a kódokat felismerné és tárgyszókezelésébe implementálná.

1.1. Egy köztauruszos teauruszcikk felépítése

Az **információkereső teaurusz** egy ún. **kontrollált** és **strukturált szótár**, amelynek szócikkei a teauruszcikkek. A teauruszcikk **lexikai egységei**

¹ Az ISO-szabvány szakkifejezései alapján készített terminológiai szójegyzék (A melléklet: szabványterminológia) megtekinthető a folyóirat weboldalán: A szövegtörzsben dőlt betűs kiemeléssel szereplő szavak ezen terminológiai szójegyzék elemeire utalnak.



azonosítják a **fogalmat** és leírják e fogalomnak a kapcsolatrendszerét. A Köztaurusz tezaurszcikkei szabványos formában az alábbi módon épülnek fel: Vezérszavak: az egyes **fogalmakat** reprezentáló **lexikai egységek** megnevezései. A tezaurszcikk kitüntetett helyén (balra zárva az első sorban) található. A **lexikai egységek** két típusa lehet vezérszó:

- **Deszkriptor:** egy **lexikai egység** kitüntetett (előnyben részesített) névalakja, a **fogalom** azonosítására is szolgál, tehát egy adott névalak egy **információkereső tezaurszban** csak egyszer fordulhat elő.
- **Nemdeszkriptor:** egyéb névalakú **lexikai egység**, egyben **utaló** a fogalom **deszkriptorára**.

Relációk egy tezaurszcikkben (MSZ 3418 jelöléssel):

- L: LÁSD **utaló** (**nemdeszkriptoroknál** alkalmazott **utaló** az adott **deszkriptorra**).
- LV: LÁSD VAGY **utaló** (**nemdeszkriptoroknál** alkalmazott **utaló** több **deszkriptorra**).
- L&: LÁSD ÉS **utaló** (**nemdeszkriptoroknál** alkalmazott, összetettebb **fogalomra** vonatkozó **utaló**, több **deszkriptor** együttes használatát jelzi).
- H: HELYETT **utaló** (**deszkriptoroknál** alkalmazott, a rá irányuló LÁSD **utaló** vezérszavának jelzése).
- H&: HELYETT ÉS **utaló** (**deszkriptoroknál** alkalmazott, a rá és egy vagy kettő másik **deszkriptorra** irányuló LÁSD ÉS **utaló** vezérszavának jelzése).
- HV: HELYETT VAGY **utaló** (**deszkriptoroknál** alkalmazott, a rá vagy egy vagy kettő másik **deszkriptorra** irányuló LÁSD VAGY **utaló** vezérszavának jelzése).
- F: általánosabb (**főlérendelt deszkriptorra** – illetve logikailag az általa reprezentált **fogalomra** – **utaló**, a generikus hierarchia felfelé iránya).
- A: speciálisabb (**alérendelt deszkriptorra** – illetve logikailag az általa reprezentált **fogalomra** – **utaló**, a generikus hierarchia lefelé iránya).
- T: átfogóbb, egésze (arra a **deszkriptorra** – illetve logikailag az általa reprezentált **fogalomra** – **utaló**, amelynek részét képezi a mi **fogalmunk**).
- P: része, eleme (arra a **deszkriptorra** – illetve logikailag az általa reprezentált **fogalomra** – **utaló**, amely a mi **fogalmunk** részét képezi).
- R: rendeltetése, tárgya, okozata (arra a **deszkriptorra** – illetve logikailag az általa reprezentált **fogalomra** – **utaló**, amely a mi **fogalmunk** okozata, tárgya, valamilyen következménye).
- E: eszköze, kiindulása, oka (arra a **deszkriptorra** – illetve logikailag az általa reprezentált **fogalomra** – **utaló**, amely a mi **fogalmunk** oka, előfeltétele, valamilyen kiindulópontja).
- X: rokonsági kapcsolat (egy **rokonsági deszkriptorra** – illetve logikailag az általa reprezentált **fogalomra** – **utaló**, amely valamilyen egyéb típusú relációban van a mi **fogalmunkkal**).

2. A Relex és a Köztaurusz szabványossági vizsgálata a hatályos tezaurszszabvány szempontjából

2021-ben magyar nyelvű utód nélkül került visszavonásra az MSZ 3418:1987 magyar tezaurszszabvány (Magyar Szabványügyi Testület 2023). A szabványosítás szabályainak megfelelően ebben az esetben, honosított magyar szabvány híján, a hatályos nemzetközi szabvány a mérvadó dokumentum. A jelenleg hatályos ISO 25964 a korábbi ISO 2788 (egynyelvű információkereső tezaurszok és egyben

„A szabvány fontosnak tartja, hogy egy tezaurusban meghatározzák az egyes fogalmak alkalmazási területeit. ”

az MSZ 3418 forrása) és az ISO 5964 (többnyelvű információkereső tezauszok) felülvizsgálatával és kombinálásával született. (National Information Standards Organization 2023) A legfőbb különbség az 1980-as években bevezetett tezauszszabványok és az új között, hogy utóbbi már előtérbe helyezi azt a tényt, hogy maguk az **információkereső tezauszok** is a papíralapú használat és megjelenítés helyett inkább mint keresőmotorokba, valamint tartalomkezelő rendszerekbe integrálható információkereső eszközökként működnek, és ezért a legfőbb felhasználói célcsoport immáron nem az információs szakemberek, hanem a szakirányú képzésben nem részesült felhasználók köre. (International Organization for Standardization 2011 p. vi.) Az alábbiakban az ISO 25964-1:2011 egynyelvű **információkereső tezauszokra** tett megállapításainak tükrében vizsgálom a Köztauruszt, illetve a Relexben épített relációkat. A két szolgáltatást illetően az alábbi megállapítások nem kritikát közölnek, hiszen jelenlegi állapotuk létrejötté idején (2011–2012) az akkor még hatályos magyar tezauszszabvány-nak kellett megfelelniük. E jelentés sokkal inkább reflexió a tezauszépítésben és a könyvtári szabványok területén azóta életbe lépett változásokra, amelyeket majd a Köztaurusznak is követnie kell, hogy megőrizze jelenlegi funkcióját a magyar könyvtári szektor általános célú tezauszaként és **tárgyszórendszereként**.

2.1. A lexikai egységek szabványossága

Egy **információkereső tezausz** az egyes **fogalmakat** közös jellemzőik alapján kategorizálja. (International Organization for Standardization 2011 p. 18.) A fogalmakat **lexikai egységek** fejezik ki a tezaurusban. Bonyolult **fogalmak** kifejezésére **többszágú** vagy **„összetett” lexikai egységek** is alkalmazhatók. Ezek megválasztását a szabvány megengedő módon a szerkesztő szubjektivitására bízta, betartva a tezauszformátum szerkesztési szabályait (pl. a névalak gyakoriságának, az egyértelműség és a visszakeresési szempontok figyelembevételével), rugalmas módon lehetőséget teremtve arra, hogy a tezausz reflektáljon az adott nyelvre és kulturális közegre. (International Organization for Standardization 2011 p. 38–40.) A szabvány fontosnak tartja, hogy egy tezaurusban meghatározzák az egyes **fogalmak** alkalmazási területeit. Ezt a használatra vonatkozó megjegyzéssel (**megjegyzés a használatról**) lehet szabványos formában megtenni (ISO-szabvány szerinti jele: SN). Nem egyértelmű esetekben a tezausz-cikkben ez a megjegyzés ad magyarázatot és instrukciót a kifejezett **fogalomra** vonatkozóan, egyértelműsítve a jelentést, illetve bizonyos esetekben, pontosítja a használat módját. A történeti megjegyzés (jele: HN) az adott fogalom használatának a visszakeresést befolyásoló múltbeli információit rögzíti.

A **deszkriptort** ismertető egyszerű magyarázószöveg (jele: DEF, MSZ 3418-ban: M²) megadására is ad lehetőséget a szabvány (International Organization for Standardization 2011 p. 18–21.). A Köztauruszban például a **közigazgatás lexikai egységénél** jól példázható a megjegyzések szabványos használata: a magyarázat pontosan leírja a **fogalom** tezaurusban használt jelentését, a **megjegyzés a használatról** egyértelmű instrukciókat ad arról, hogy bizonyos, jól körülhatárolható esetekben melyik másik **deszkriptort** (*államigazgatás*) kell alkalmazni. A történeti megjegyzés évszám szerint leírja, hogy ezt a fogalmat Magyarországon mikor milyen formában alkalmazták. A **lexikai egységek** szabványos formában szótári formában megadott, egyes számú főnevek vagy névszói szerkezetek, ám a szabvány hangsúlyozza, hogy az egyes nyelvhasználati és kulturális sajátosságok és szabályok az irányadók. Az azonos írásmódú, de eltérő jelentéssel bíró

² A MSZ 3418:1987 csupán egyetlen szabványosított megjegyzésről tesz említést mint megjegyzésre mutató kiegészítő kapcsolatjel (M). Ez a nemzetközi szabvány szerinti definitionnek (DEF / magyarázat) és scope note-nak (SN / megjegyzés a használatról) egyaránt megfeleltethető. A Relexben definiált megjegyzések esetében az M a magyarázat jele, ezért használtam az elemzésben ebben a kontextusban.



homográf szavak esetében zárójelben értelmező adható meg: például a szabvány a daru angol nyelvű előfordulását (*cranes*) írja példaként. E konkrét fogalomnál, főleg hogy ez a nyelvi jelenség kevésbé gyakori a magyar nyelvben, a Köztaurusz szerkesztői ragaszkodtak például a *daru* és *darumadár* **deszkriptorok** használatához, így két teljesen eltérő névalak azonosítja a két fogalmat. Közöttük a MÁS ÉRTELEMBEN LÁSD relációt (jele: ≠) is feltüntették. Az *orgona* esetében láthatunk példát a hátravetett értelmezők használatára az *orgona* (*hangszer*) esetében. Érdeemes megjegyezni, hogy az ISO-szabvány mindkét értelmezési tartományban javasolja az értelmezők használatát, nem csak a fogalompár egyikenél. Ungváry Rudolf szerint azért tértek el a szabvány által javasolt alaktól, mert úgy vélték, a jövőben a hátravetett értelmezők használata megnehezíti a gépi **indexelés, osztályozás**, tárgyszavazás folyamatát, mivel a zárójeles hátravetett jelzőkkel sohasem fordulnak elő a kérdéses szavak a szövegben (Ungváry 2001a p. 26.). Az azóta megvalósult, a finn Annifhoz hasonló, gépi tárgyszavazó rendszerek azonban más módon működnek. A mesterséges intelligencia tanítása során a tanítókorpuszt, vagyis egy szakemberek által **indexelt, tárgyszavazott, osztályozott** állományrészt, az algoritmusok statisztikai és egyéb módszerek alapján elemzik, és így tanulják meg, hogy egy adott szöveghez egy szótár melyik előre kijelölt **lexikai egységét** kell **indexként, osztályozási jelzetként** vagy **tárgyszóként** hozzáilleszteni³ (az algoritmus szempontjából nincs jelentősége az **index** formai alakjának, az akár emberi olvasásra alkalmatlan gépi kód is lehetne).

A különböző névalakok közül a **deszkriptorok** kiválasztásánál a szabvány a legelterjedtebb írásmódot tekinti irányadónak. **Jövevényszavak** is lehetnek **deszkriptorok**, ha indokolt a használatuk vagy elég elterjedtek a fogadó nyelvben. (International Organization for Standardization 2011 p. 23–32.) Erre jó példa a podcast, amelynek nincs magyar nyelvű széles körben elterjedt megfelelője, ezért a Köztauruszban is így található meg, magyarázattal és történeti megjegyzéssel kiegészítve. A szabvány egyébként lehetőséget kínál a lexikai egység névalakjának további tipizálására (írásmódváltozat /SP/, elírás /MS/, rövidítés /AB/, **lexikai egység** teljes formája /FT/) (International Organization for Standardization 2011 p. 47.). Ami a kurrens ISO-szabványban újítás a régebbihez és a magyar teauruszszabványhoz képest, hogy nem javasolja a VAGY operátorok alkalmazását. Helyette a **homográf** szavaknál alkalmazott értelmezők használatára tesz javaslatot (de nem írja elő). A Köztauruszban például a *biztosítás* **utaló** a *biztonságtechnika*, *biztosítási ügylet* vagy *biztosításügy* **deszkriptorokra** utal. A szabvány javaslata szerint három, az alkalmazás módját körülhatároló értelmezővel ellátott **utaló** alkalmazása a célszerű, hogy a bonyolult **fogalmakat** kifejező **összetett lexikai egységek** között ne legyen VAGY logikai kapcsoló. (International Organization for Standardization 2011 p. 49–50.) A szabvány alkotói szerint az értelmezők használata (még a magától értetődő kifejezések esetén is) elősegíti a **lexikai egységek** olyan megfelelő rendezését, amely megkönnyíti az **információkereső teaurusz** jövőbeni esetleges bővítését (International Organization for Standardization 2011 p. 22.).

2.2. Teauruszrelációk szabványossága

Az ISO 25964-1 szerint az **információkereső teaurusz** célja, hogy az indexelőt és a keresőt egyaránt egy adott **fogalmat** kifejező adott kifejezéshez vezesse. Ezért a teauruszban egy adott tartományban meg kell adnia az **információ visszakereséséhez** használt, az adott **fogalmat** kifejező összes kifejezést,

„A különböző névalakok közül a deszkriptorok kiválasztásánál a szabvány a legelterjedtebb írásmódot tekinti irányadónak.”

³ Az Annif működéséről részletek lásd a projekt dokumentációját: <https://github.com/NatLibFi/Annif/wiki> (2023.05.26.)

„A hatályos teaurusz-szabvány a hierarchikus relációk három logikai típusát különbözteti meg: generikus (nem-faj), egész-rész (partitív), illetve előfordulás (instancia).”

majd ezekből kell kiemelni egy konzisztens módon e célra használt kitüntetett névalakot, a **deszkriptort**. A teauruszokban meglévő kapcsolatrendszer felhasználható a keresés kiterjesztésére, alternatív **keresőkifejezések** ajánlására, halmazképzésre vagy más módon a keresés pontosítására, a gyakori elírási hibák azonosítására, illetve automatikus **indexelés** támogatására. Egy adott **információ visszakeresése** során az egyes **dokumentumok** tartalmát (**tárgyát**) jelölő **indexek** (vagy mutatószavak) közötti **paradigmatikus relációk** biztosítják a leghatékonyabb keresési módot, ezek ugyanis az egyes **fogalmak** esetében magától értetődő kapcsolatokat jelentenek. A szabványban is ismertetett példa szerint az amszterdami bankok számítógépes rendszereivel foglalkozó kiadványhoz **indexként** használt **lexikai egységek** (*bank, számítógép, Amszterdam*) teauruszban rögzített **paradigmatikus relációi** kiterjesztik a keresést például a *pénzintézet, adatfeldolgozó gép* és *Hollandia* kifejezésekkel indexelt **dokumentumokra** is. Ezen relációk feltüntetése nélkül csupán az eredeti három **indexet** rögzítik a **dokumentum** leírásához, amelyek egy bizonyos fogalmi kontextusban ún. **szintagmatikus relációt** képeznek, jelentősen csökkentve a keresett témára vonatkozó találati halmazzt, mivel így a keresésben kizárólag e három mutatószó együttes használata vezet eredményre (International Organization for Standardization 2011 p. 15–17).

Az egy adott **fogalmat** kifejező **deszkriptor** és **nemdeszkriptorok** között ekvivalencia a kapcsolat típusa, köztük **utalók** biztosítják az átjárást. Nemzetközi szabvány szerinti jelölése USE (MSZ 3418 szerint: L), ha **nemdeszkriptorról deszkriptorra** utal, illetve UF (MSZ szerint: H) fordított esetben (International Organization for Standardization 2011 p. 44.). A Köztaurusz szabványos formában alkalmazza ezeket a relációkat a magyar szabvány szerinti jelöléssel.

Míg egynyelvű **információkereső teauruszban** ekvivalencia csak **lexikai egységek** között létezik, a többi teauruszreláció már **fogalmak** vagy azok az által reprezentált konkrét dolgok közötti kapcsolatokat reprezentál. A hatályos teauruszszabvány a **hierarchikus relációk** három logikai típusát különbözteti meg: generikus (nem-faj), egész-rész (partitív), illetve előfordulás (instancia). Amikor egy fogalomnak egynél több alá- vagy fölérendelt fogalompárja van, akkor beszélhetünk polihierarchikus relációról. A fogalmak közötti **asszociatív relációkat**, tehát minden egyéb nem meghatározott, magyar terminológiában rokonsági kapcsolatot, RT-vel (MSZ szerint: X) jelölnek. Fontos megemlíteni, hogy az ISO-szabvány megengedi a szemantikailag laza **hierarchikus relációk** kialakítását is a **fogalmak** között, amely rendszerezés egyaránt történhet generikus és rész-egész elvek mentén, azonban maga a reláció pontos típusa nincs megkülönböztetve. A pontos relációtípust minősítő jelző nélküli *narrower term* (NT), illetve *broader term* (BT) elvileg ezt hivatottak jelezni⁴ (International Organization for Standardization 1974 p. 7), azonban felhívjuk arra a szabványalkalmazók figyelmét, hogy a hatályos ISO-szabvány alapértelmezetten a generikus hierarchiát (megkülönböztetett jele NTG/BTG, MSZ szerint: F/A) tekinti az általános értelemben használt hierarchikus reláció (NT/BT) ekvivalensének. Ugyanakkor a szabvány használati példát ad arra is, ha valaki laza hierarchiaként kíván fogalmi kapcsolatokat kialakítani. Ekkor azonban a szabványalkotók hangsúlyozzák, hogy a teaurusz szerkesztője kerülje az információkeresésben tényleg vezető kapcsolatokat. Például gépkocsikra vonatkozóan ne adjuk meg **alárendelt deszkriptorként** a kereket, mert az az autók fogalmi csoportjához képest túlságosan tág csoportosítást képvisel, ugyanakkor a *nyíl* és *nyílhegy* viszonyában

⁴ Az ISO-szabvány első kiadásában konkrétan leírják, hogy ebben az esetben használatos az NT-BT relációpár, a kevésbé normatív értelmezés a következő kiadásoknál jelent meg.



az NT/BT reláció minden különösebb gond nélkül alkalmazható, ha a szerkesztő nem kíván élni a különböző hierarchiatípusok explicit használatával. A szabvány konkrét példájában a *koponya – csont – fej* fogalmak esetében a *koponya* poli-hierarchikus fölérendelt fogalmi szabványos módon kifejezhető úgy is, hogy egyrészt a *koponya* szűkebb értelemben alárendelődik mind a *fejnek*, mind a *csontnak*, másrészt pontosan megjelölve a hierarchia típusát úgy, hogy a *koponya* a *csont* generikus alárendeltje és a *fej* része. (1. ábra)

1. ábra

LAZA HIERARCHIA (NT/BT)	MEGKÜLÖNBÖZTETETT HIERARCHIATÍPUSOK (NTG/BTG, NTP/BTP)
<p>koponya BT csont fej</p>	<p>koponya BTG csont BTP fej</p>

A generikus hierarchia egyben mindig tranzitív is, vagyis a hierarchialánc leg-szűkebb értelmű fogalmának ismérvei igazak a lánc csúcsára is. Az egész-rész (partitív) hierarchia (jele: NTP/BTP, illetve MSZ szerint: T/P) is lehet tranzitív. Például földrajzi nevek esetében egy adott utca ugyanúgy része a világegyetemnek, mint az adott városrésznek. Ugyanakkor a partitív hierarchia, különösen alkatrészekre, részegységekre vonatkozóan, megakadhat egy ponton, lásd a *nyílhegy – nyílvevő* példáját NTP/BTP relációként kifejezve (International Organization for Standardization 2011 p. 58–62.).

A koponyás példa a Köztauruszban orvostudományi szakterületi fogalomrendszer szerinti rendezést képvisel. A *koponyának* itt is a *csont* a generikus fölérendelt **fogalma** (F), ugyanakkor nem a *fej* része, ezzel rokonsági (X) kapcsolatban áll, hanem a *tájanatómia* orvostudományi **fogalom** képezi egészfogalmát (T). A *koponya* generikus **csúcsdeszkriptora** a *sejtszövet*, az egész-rész reláció pedig már a *tájanatómiánál* megszakad, ellenkező irányban azonban sorra felvannak tüntetve a koponya részei. (2. ábra)

A **fogalmak** rendezési elvét tehát az **információkereső teaurusz** szerkesztője és persze maga a fogalom szemantikus környezete határozza meg. Ha a **fogalmak** közötti reláció megfelel a relációtípusok követelményeinek, akkor a teaurusz e szempontból szabványosnak tekinthető. A generikus és partitív relációk esetében némileg eltér az MSZ 3418, ugyanis a magyar szabvány a **deszkriptorok** közötti relációk esetében csak alá-fölé rendeltségi (nem-faj) és rokonsági kapcsolatokról tesz említést. Az egész-rész kapcsolatot már mint a rokonsági kapcsolat egyik kiemelését értelmezi, tehát a nemzetközi szabvánnyal ellentétben nem engedi meg, legalábbis alá-fölé rendeltségi kapcsolatként, a lazább hierarchikus rendezést (Magyar Szabványügyi Hivatal 1976 p. 3., Magyar Szabványügyi Hivatal 1987 p. 4.).

A magyar teauruszszabványban tehát mind a generikus, mind a partitív relációt problémamentesen adaptálták (az A/F – NTG/BTG és P/T – NTP/BTP szabványos jelölésű relációk formájában). A bizonyos **fogalmak** egyes egyedeinek előfordulásait (pl. *hegység – Alpok, főváros – Párizs*) jelölő relációt (instance relationship) azonban nem vették át. Ez az előfordulás-reláció maga az ISO 2788 1986-os revíziójának újdonsága, onnan került áttemelésre a kurrens ISO 25964-1:2011 szabványba. A Köztauruszban azonban az egyes előfordulások

„A fogalmak rendezési elvét tehát az információkereső teaurusz szerkesztője és persze maga a fogalom szemantikus környezete határozza meg.”

konzisztensen a **fogalmak** generikus alárendeltjeiként szerepelnek. Például a *Párizs a főváros* generikus alárendelt (A vagy NTG) **fogalma**, hasonlóképp az *Alpok* polihierarchikus módon az *európai hegység*, illetve a *közép-európai természeti táj* alárendeltje. Ez a megoldás alapvetően megy szembe a hatályos szabvány előírásaival, amely pontosan azzal indokolja az előfordulás-reláció létjogosultságát, hogy egy fogalmi csoportosítás egyedi előfordulása nem fajtája (generikus hierarchia), sem pedig része az adott csoportosításnak. Konkrét példával élve az *Alpok* nem az *európai hegység* fajtája, sem pedig annak része, hanem konkrétan egy *európai hegység*. (International Organization for Standardization 1986 p. 17., International Organization for Standardization 2011 p. 61.)

2. ábra

PÉLDARELÁCIÓK A KÖZTAURUSZ EGY TEZAUROSZCIKKÉBEN (AZ ELEMZÉSBEN HASZNÁLT FORMÁTUMHOZ IGAZODVA)
koponya
H
cavum cranii
cranium
regiones capitis
H&
szemgödör
F
csont
A
agykoponya
arckoponya
T
tájanatómia
P
arcüreg
ékcsontüreg
homloküreg
koponyaalap
koponyacsont
koponyatető
koponyaüreg
orrüreg
szemüreg
X
fej

Ez a relációtípus tehát nem került szabványosításra a magyar teauruszépítés gyakorlatában, ugyanakkor a hazai gyakorlat új relációval bővült a következtetési (determinációs), vagyis az előfeltétel-rezultáns (E/R) relációpár formájában. Ez a reláció már a magyar szabvány első, 1976-os változatában meghatározásra került. (Magyar Szabványügyi Hivatal 1976 p. 3., Magyar Szabványügyi Hivatal 1987 p. 4–6.) Nem meglepő a formai eltérés az MSZ- és az ISO-szabványok között, ugyanis ekkoriban a nemzetközi szabványok hazai adaptációja más folyamatot



jelentett, mint manapság. Az ISO-szabványoknak nem magyar nyelvű kiadásai jelentek meg, hanem a nemzetközi szabványok irányelvei épültek be a hazai gyakorlatot és hagyományokat megőrző állami szabványokba (MSZ). (Szente 1982 p. 146.) Jelenleg a nemzeti sajátosságok az országos és nemzetközi szabványosítás intézményrendszerében közvetlenül épülnek be a szabványalkotás folyamatába, ha az adott nemzeti és nemzetközi szabványosítási bizottságokban szerepet vállalnak a terület szakemberei. Ungváry Rudolf cikkeiben Vajda Eriknek és Horváth Tibornak tulajdonítja az akkor kétségkívül korszerűnek tekinthető magyar tezaszabvány a nemzetközínél gazdagabb szemantikus kapcsolatrendszerét. (Ungváry 2011, Ungváry 2014 p. 243.) Az természetesen elődeink előrelátását és alaposságát dicséri, hogy olyan relációtípust szabványosítottak, amelyet a nemzetközi gyakorlat ugyan nem tartott indokoltnak, de a magyar tezaszabvány építését igényelt tartott rá. A máig ható negatív következmény inkább a magyar szabvány 1987-es revíziójához kapcsolódik, amelyből hiányzik a nemzetközi gyakorlatba bevezetett előfordulás-reláció, holott az MSZ 3418-87 kijelenti: „*E szabvány kidolgozásakor figyelembe vették továbbá az ISO 2788–1986 nemzetközi szabványt is, és e szabvány műszaki tartalma az ISO szabványéval megegyezik, jelentéktelen műszaki eltérésekkel.*”⁵ (Magyar Szabványügyi Hivatal 1987 p. 10.) Az MSZ 3418 egyfajta paradoxont is létrehoz azzal, hogy a tezaszabvány összetétele és megjelenési formája kapcsán előírja, hogy egy szabványos **információkereső tezaszabványban** fel kell tüntetni a következő szöveget: „*Ez a tezaszabvány az MSZ 3418 szabványának, a KGST SZT 174-85 szabványának, valamint az ISO 2788 nemzetközi szabványának megfelelően készült.*” (Magyar Szabványügyi Hivatal 1987 p. 9.) A paradoxon abban rejlik, hogy ha akkoriban a magyar tezaszabvány szerkesztő csak a magyar szabványt vette alapul (joggal elfogadva, hogy az megfelel a nemzetközinek), akkor az 1986-tól hatályos nemzetközi szabvány értelmében munkája nem tekinthető szabványosnak, már amennyiben az egyed-előfordulásokat, például konkrét földrajzi, márka- vagy egyéb tulajdonneveket tartalmaz. Az bizonyos, hogy maga a reláció nem jelenthetett újdonságot a szakembergárdának, mivel a magyar szabványt készítők, majd revízió alá vevő szakértői bizottságban is szerepet vállaló Horváth Tibor és Varga Dénes közös kötetükben még az MSZ 3418 első megjelenése előtt a tőlük megszokott magas szakmai színvonalon értekeztek az **információkereső tezaszabványokban** létesíthető relációtípusokról, köztük az előfordulásról is, amelyet mint a fogalmak által képviselt osztályok elemeit kifejező kapcsolat-ként ismertettek. (Horváth, Varga 1977 p. 41.) Ungváry Rudolf is írja 2016-os tanulmánya jegyzetében, hogy „A földrajzi helyek besorolása névtípusokba (pl. belterületi településrész, folyóvíz) valójában a tulajdonnévvel azonosított hely és az általános névvel azonosított fogalom közötti terjedelmi eleme – fogalmi reprezentánsa reláció, melyet többnyire azonosnak vesznek a névterekben a generikus relációval, noha attól alapvetően különbözik”. (Ungváry 2016 p. 155.) És valóban, külföldi tulajdonneveket is tartalmazó tezaszabványokat (pl. a Getty Thesaurus of Geographic Names) böngészve feltűnő, hogy vagy hiányzik az 1980-as években szabványosított előfordulás-reláció, vagy eleve laza hierarchiaként (NT/BT) építették fel a fogalmi kapcsolatokat. A Getty földrajzinév-állományának esetében láthatóan szerkesztési irányelv volt, hogy az általános nevek a tulajdonnevek rendezésében kapjanak szerepet, semmint fogalmi kapcsolatot alkotva velük. A földrajzinévtípusok a deskriptorok zárójeles értelmezőjében vannak feltüntetve, maguk a földrajzi nevek egésze-része hierarchiába vannak rendezve⁶ (Harping 2010), amelynek a konkrét csúcspontja a világ (*World* – valójában

⁵ Szigorúan vizsgálva már a generikus és partitív relációk „lazaságának” kérdése is megkérdőjelezi a nemzetközi és a magyar szabvány ekvivalenciájának érvényességét, de az a megkötés kétségkívül inkább jellemezhető „jelentéktelen műszaki eltérésként”, sőt a mai napig előfordulnak apróbb, a szabványok nemzeti előszavában jelzett, eltérések, amelyek az alkalmazó ország eltérő alkalmazási környezetére, kultúrájára reflektál.

⁶ A hivatkozott kézikönyv is bemutatja a szabványos hierarchiatípusokat, azonban az előfordulásra nem hoz példát a Getty gyakorlatából, ezzel is jelezve, hogy e relációt ők maguk sem alkalmazták.

„A Köztaurusz földrajzinév-állományának alapja a Geotaurusz és Geohistaurusz volt.”

pedig egy *hierarchy root*nak nevezett **csúcsheszriptor**). A háttéradatbázisban valószínűleg egyed-kapcsolat szinten vannak összekötve a keresési szempontként is használható földrajzi névtípusok, országnevek és földrajzi nevek. Tehát az ISO-szabványban meghatározott NTI/BTI reláció explicit módon nincs kifejezve itt sem, pedig az egyik legjelentősebb tulajdonneveket tartalmazó tezauszról beszélünk. (Getty Thesaurus of Geographic Names Online.) Tehát valamiért – nem kizárt, hogy az 1986-os bevezetésekor már régen létező tezauszok nagyon munkaigényes átszerkesztését kívánták megelőzni – általánosan a tezauszépítés gyakorlatában ez a relációtípus nem érvényesült.

A Köztaurusz földrajzinév-állományának alapja a Geotaurusz és Geohistaurusz volt. Ez egészült ki a GEONAMES mai magyarországi névállományával, az Egyetemes Tizedes Osztályozás általános és konkrét földrajzi neveivel, legutóbb pedig a FÖMI (Földmérési és Távérzékelési Intézet) által vezetett „Magyarország földrajzinév-tára” állományával.⁷ Az utóbbi egyben tartalmazza a KSH által vezetett földrajzinév-nyilvántartás neveit. Idővel pedig a Köztaurusz folyamatosan bővült és bővül a kurrens feldolgozás által igényelt újabb nevekkel.

A Köztaurusz földrajzinév-állományának magját képező Geotaurusz és Geohistaurusz is generikus hierarchiával fejezi ki a földrajzinévtípusok és az egyes földrajzi nevek közötti kapcsolatot (Cserbák 1995.). Nem nehéz megérteni, hogy a fogalom és egyede kapcsolatának kifejezésére miért bevett gyakorlat a generikus hierarchia szerinti rendezés. Ezt a fajta relációt a nyelvészetben hiponimának nevezik, amely során egy fogalom magába foglalja (inklúzió) a másik fogalmat. A hiponima esetén a nyelvészetben az „is_a” („az_egy”) nyelvi teszttel állapítható meg a hiponima esete (Szakadát 2005 p. 45–46.). A Köztaurusz példájára alkalmazva a tesztet nincs logikai ellentmondás, hibátlan a logikai generikus láncolat: az *Alpok* az egy *európai hegység*, az *európai hegység* az egy *hegység*. Ennélfogva az *Alpok* is egy *hegység*, sőt pár hierarchiaszinttel feljebb lépve is igaz a következtetés, miszerint az *Alpok* az egy *hely*, és az is hogy az *Alpok* az egy *valami* (ez a Köztaurusz **csúcsheszriptora**). Ungváry Rudolf is egyik tanulmányában e logikát követve vezeti le a fogalmak hierarchikus rendezésének elvét a specifikus **fogalomtól** az elvont csúcsheszriptomig. Bár itt nem kerül említésre a tezauszban megjelenő egyed-előfordulások problematikája, a gondolatmenet jól példázza a tezausz alapvető hierarchikus felépítése mögötti logikai elveket. Ugyanakkor Ungváry felhívja a figyelmet, hogy a generikus reláció azt a tényt jelöli, miszerint az egyik fogalom fajtája a másiknak (Ungváry 2007 p. 445–446.). Itt viszont filozófiai probléma annak értelmezése, hogy az *Alpok* a *valami* és az *európai hegység* egy fajtája lenne. A hatályos tezauszszabvány nem mond erre az állításra, és a tudásreprezentációs és informatikai gyakorlatban, például ontológiáknál, objektumorientált programozási nyelvekben és adatbázisok szerkesztésében, szintén megkülönböztetett jelentést kap az „is_a” viszonyítás. Ott ugyanis a fajtáját jelölő reláció egyértelműen többek között osztály-alosztály (class-subclass), illetve típus-altípus (type-subtype), egyedtípus és az egyedtípus altípusát jelenti (Halassy 1994 p. 102–103.). Az **információkereső tezauszokban** rendszerezett **fogalmak** informatikai értelemben (relációs adatbázisoknál) az egyedeknek (entity) nevezett konkrét dolgok közös tulajdonságait magába foglaló absztrakt csoportosításoknak, vagyis az egyedtípusoknak (entity type) feleltethetők meg. Az egyedtípusok egyes egyedeit, előfordulásait, instanciáit az informatika sajátos logikával az egyed-előfordulással (entity occurrence vagy instance) fejezi ki (Halassy 1994 p. 33., Wikipédia 2022, 2023). Tehát informatikai, adatbázis-építői

⁷ Az állományt az intézményen belül kezelik, külső felhasználó nem veheti igénybe.

szemmel nézve az *Alpok* egy egyed, amely közös tulajdonságok révén (földrajzi pozíció: Európa, illetve földrajzi tájegység típus: hegység) az *európai hegység* nevű egyedtípus egyik egyed-előfordulása. Ugyanez a logika érvényes az *Alpok* mint objektum és az *európai hegység* mint osztály esetében. Az ISO 25964-ben is meghatározott NTI/BTI relációpár ezt a gyakorlati megközelítést vezeti be az **információkereső teauruszok** adatmodelljébe. A Köztaurusz mindenkori szabványosságát szem előtt tartva a jövőbeni fenntartási és szerkesztési munkában érdemes bevezetni majd ezt a relációtípust, különös tekintettel a tervezett kapcsolattal-alapú teauruszformátum létrehozásának és a más szótárakkal, névterekkel való kompatibilitás igényére. Ezekkel a kérdésekkel foglalkozik részletesen az ISO 25964-2:2013 szabvány, amelynek normáit szem előtt kell majd tartani a jövőbeni, a Köztauruszt érintő fejlesztések során.

A magyar teauruszszabványban bevezetett következtetési (E/R) kapcsolatra visszatérve megjegyzendő, hogy ez mára már nem teljesen idegen a hatályos ISO-szabványban meghatározott relációktól. Az ISO 25964-1:2011 az egyedi kapcsolatok meghatározásánál a következtetési relációra (CAUSE/EFFECT jelöléssel) hoz példát a hierarchikus relációk tipizálásának analógiájával az **asszociatív relációk** szükség esetén hasonló módon típusokra bontásával (International Organization for Standardization 2011 p. 67).

3. A Relex és a Köztaurusz szabványossági vizsgálata a MARC 21 szempontjából⁸

A korabeli magyar tervezettel ellentétben natív módon a MARC 21 csupán kétféle hierarchia kódolását teszi lehetővé. Az egyszerűbb a kétirányú rokonsági reláció, amelyet az X50-es mező \$a almezője ad meg jelöletlen relációtípust vezérlő almezővel (\$w). A másik a generikus hierarchia (vagy a pontos szabályok nélkül meghatározott általánosabb laza hierarchia) az X50-es mezők \$w relációtípus-vezérlő (nulladik érték) g és h pozícióján, amely meg is egyezik a HUNMARC-tervezetével, így elvileg a Köztaurusz jelen formátumában is képes MARC-ban legalább ezt a hierarchiatípust kezelni. Elvileg – és ezt az elképzelést a gyakorlatban is ki kívánjuk próbálni – az X50-es mezők \$w relációtípus-vezérlőjében az r pozícióval beállítható, hogy az X50 \$i almezőbe (relációinformáció) szövegesen beírt módon, szabványos formában kifejezhető legyenek további teauruszrelációk, így szabványos formában MARC 21 szerint lehetne kódolni a magyar teauruszhasználati sajátosságokat.

Mindezeket figyelembe véve pl. az *államigazgatás* teauruszcikk esetében így alakulna a MARC 21 szerinti szabványos leírás. A bal oldali és a középső a jelenlegi „HUNMARC” szerinti export, egyrészt a Relexből lekérdezve, másrészt egy lokális HUNTEKA-alkalmazás⁹ rekordjából kimásolva (kék kiemeléssel jelöltem a MARC 21-gyel nem kompatibilis kódolást). (3. ábra)

Szembeötlő a különbség a közvetlen Relex-export, illetve a lokális alkalmazás között. A közvetlen exportnál a relációtípus-vezérlő almezők nulladik karakterpozícióján megadott kódok megfelelnek a HUNMARC-tervezetben javasolt relációtípusoknak, azonban a Huntékában már ugyanezen a pozíción valamiért teljesen más kódok szerepelnek a rekordokban. A kétirányú rokonsági reláció sem került át a lokális rekordba. Ha ez utóbbi jelenség általános, akkor a MARC 21 szerinti majdani exportban a kétirányú rokonsági relációt is dedikáltan kell jelölni a \$w és \$i almezők kombinációjával. Itt érdemes megjegyezni, hogy például az alárendelt *rendészet* teauruszcikk esetében a rokonsági relációkat a Huntéka \$wm almezővel kódolta.

„...így szabványos formában MARC 21 szerint lehetne kódolni a magyar teauruszhasználati sajátosságokat.”

⁸ A Köztaurusz/ Relex MARC-export-lehetőségeinek és a MARC 21 authority formátumának az összevetésével készült táblázat megtekinthető a folyóirat weboldalán.

⁹ E célra a kaposvári Takáts Gyula Könyvtár OPAC-ját használtam.

3. ábra

„HUNMARC” (RELEX)	„HUNMARC” (HUN- TÉKÁBÓL)	MARC 21
150 ##\$a államigazgatás	150 ##\$a államigazgatás	150 ##\$a államigazgatás
450 ##\$wd \$a felsőfokú igazgatás	450 ##\$wy \$a felsőfokú igazgatás	450 ##\$wr \$i helyett \$a felsőfokú igazgatás
450 ##\$wy \$a adminiszt- ráció	450 ##\$wv \$a adminiszt- ráció	450 ##\$wr \$i helyett VAGY \$a adminisztráció
450 ##\$wy \$a igazgatás	450 ##\$wv \$a igazgatás	450 ##\$wr \$i helyett VAGY \$a igazgatás
550 ##\$wg \$a irányítási-ve- zetési tevékenység	550 ##\$wg \$a irányítási-ve- zetési tevékenység	550 ##\$wg \$a irányítási-ve- zetési tevékenység
550 ##\$wg \$a ismeretterület	550 ##\$wg \$a ismeretterület	550 ##\$wg \$a ismeretterület
550 ##\$wh \$a hadügy	550 ##\$wh \$a hadügy	550 ##\$wh \$a hadügy
550 ##\$wh \$a közigazgatás	550 ##\$wh \$a közigazgatás	550 ##\$wh \$a közigazgatás
550 ##\$wh \$a külügy	550 ##\$wh \$a külügy	550 ##\$wh \$a külügy
550 ##\$wh \$a rendészet	550 ##\$wh \$a rendészet	550 ##\$wh \$a rendészet
550 ##\$wt \$a államjog	550 ##\$wj \$a államjog	550 ##\$wr \$i egésze \$a államjog
550 ##\$wt \$a civilizáció	550 ##\$wj \$a civilizáció	550 ##\$wr \$i egésze \$a civilizáció
550 ##\$wu \$a államigazga- tás és igazságszolgáltatás elleni bűncselekmény	550 ##\$wk \$a államigazga- tás és igazságszolgáltatás elleni bűncselekmény	550 ##\$wr \$i része \$a állam- igazgatás és igazságszol- gáltatás elleni bűncselek- mény
550 ##\$wu \$a gazdasági irányítás	550 ##\$wk \$a gazdasági irányítás	550 ##\$wr \$i része \$a gazda- sági irányítás
550 ##\$wu \$a közigazgatási ellenőrzés	550 ##\$wk \$a közigazgatási ellenőrzés	550 ##\$wr \$i része \$a köz- igazgatási ellenőrzés
550 ##\$wu \$a pénzügyi támogatás	550 ##\$wk \$a pénzügyi támogatás	550 ##\$wr \$i része \$a pénz- ügyi támogatás
550 ##\$wu \$a ügyvitel	550 ##\$wk \$a ügyvitel	550 ##\$wr \$i része \$a ügyvitel
550 ##\$wa \$a állami szerv	550 ##\$wr \$a állami szerv	550 ##\$wr \$i eszköze, kiindu- lása, előfeltétele \$a állami szerv
550 ##\$a államigazgatási előmenetel	nem került át a lokális alkál- mazásba!!!	550 ##\$a államigazgatási előmenetel

3.1. Megjegyzések rögzítése a MARC-ban

A megjegyzésmezők esetében nagyrészt megegyezik a HUNMARC-tervezet szerinti export és a MARC 21 formátuma. Egyedül a használatra vonatkozó megjegyzésre (**megjegyzés a használatra / scope note**) használt 691-es mező hiányzik



a MARC 21-ből. Mivel a nemzetközi teauruszszabvány a 2011-es megújítás előtt szerényebb megjegyzéskészletet szabványosított, mint az MSZ 3418, egyedül a **megjegyzés a használatra** volt szabványos módon deklarálva az ISO-szabványban, a 680-as általános megjegyzésmező rugalmas módon láthatta el az egyes teauruszokra vonatkozó megjegyzések szabványos rögzítését MARC 21-ben. Mára a helyzet megoldódott, ugyanis az ISO 25964 teauruszszabványban meghatározott megjegyzéstípusok mindegyikéhez van megfelelő MARC-mező. A **megjegyzés a használatra** feladatát átveheti a MARC 21-ben a 680-as mező, ugyanis a deszkriptorok magyarázószövegéhez már létezik a 677-es meghatározás (*Definition*) mező, amely megfelel az ISO 25964-ben szabványosított „definition” (jele: DEF) megjegyzéstípusnak. Utóbbi, mivel a **deszkriptorral** azonosított **fogalmat** magyarázza, megfeleltethető az MSZ 3418-ban szabványosított magyarázatnak (M).

Fontos megjegyezni, hogy az Ungváry Rudolf tanulmányaiban leírtaktól eltérően a Relexben jelenleg alapértelmezett módon a 681-es mező alkalmas a **megjegyzés a használatra** mező funkciójának betöltésére. Ez a megjegyzésmező viszont mind a HUNMARC-tervezetben, mind a MARC 21-ben arról tájékoztat, hogy az 1XX mező tételfejét példaként idézik összetett magyarázatos tárgyi hivatkozásban és/vagy az általános megjegyzésmezőben. Érdekes módon a HUNTEKA authorityrekordjában már a 691-es mező szerepel (piros kiemeléssel, mivel a MARC 21 authority formátumában nem alkalmazott hívójel).

667 ##\$a A mai magyarországi dokumentumok esetén legfeljebb a felsőfokú kormányzati igazgatással foglalkozó dokumentumok esetében használható

670 ##\$a Se az EUROVOC, se a Kongresszusi Könyvtár teaurusza nem tartalmazza

678 ##\$a Magyarországon 1947-ig a „közigazgatás”, 1947-1990 között az „államigazgatás”, 1990 után megint a „közigazgatás” a hivatalos kifejezés

680 ##\$i Az államhatalom felsőfokú rendelkező, végrehajtó, szervező, irányító tevékenysége

691 ##\$a Az állami irányításon alapuló közigazgatás és a pártállami igazgatás esetén használandó

691 ##\$a A kormányzati szervezetek és tevékenységük esetén („state administration”) is használható

691 ##\$a Az önkormányzatiságon alapuló igazgatás, továbbá a közép- és helyi igazgatás esetén a „közigazgatás” használandó

3.2. Osztályozási jelzetek és azonosítók csatolása a teauruszcikkekhez MARC-ban

A Relex gyakorlata szerint a **lexikai egység** reprezentálta **fogalomnak** megfelelő ETO-jelzetet csak tárgyszókapcsolatként rögzítik, mivel a Relexen belül épül egy ETO-jelzeteket tartalmazó részteaurusz (a MARC 21 Classification formátumnak megfelelően a 153-as mezőkben rögzített vezérszavakkal), amelyet a következő módon kapcsolnak össze a Köztaurusz **lexikai egységeivel**:

750 ## \$a 35

A forrásmegjelölés, illetve a platformfüggetlen információadás igénye (ha az ETO-jelzeteket nem a Relexen belül, hanem például egy dedikált ETO-szótárból kinyerve szeretnénk összekötni a teauruszsal) miatt érdemes a MARC 21 szerinti szabványos módon is felvinni a jövőben az egyes **lexikai egységeknek** megfelelő ETO-jelzeteket. Például az alábbi módon, a \$2 almezővel azonosítva az alkalmazott ETO kiadását (a P057 a 2005-ös magyar teljes kiadás UDC Publication Numbere):

080 0# \$a 35 \$2 P057

„A Relex gyakorlata szerint a lexikai egység reprezentálta fogalomnak megfelelő ETO-jelzetet csak tárgyszókapcsolatként rögzítik...”

Előre el kell dönteni, hogy az egyes tárgyszavakhoz milyen mélységben rendeljünk hozzá **osztályozási jelzeteket**. Jelen példában (35 Államigazgatás. Közigazgatás. Hadtudomány) jelentősen tág csoportosításról beszélünk, amely a közigazgatás mellett a hadtudományt is magában foglalja, így jelentős számú **tárgyszóhoz, deszkriptorhoz** hozzárendelhető. Az alkalmazó könyvtár gyűjteményétől függ, hogy milyen mélységben kapcsolják össze a **tárgyszórendszert** az **osztályozási rendszerekkel**. Például ha egy könyvtár gyűjteményében jelentős mennyiségben vannak **dokumentumok** mind a közigazgatás, mind a hadtudomány témájában, érdemes mindkét témához szűkebb értelmű **osztályozási jelzeteket** csatolni.

A **tárgyszavakhoz, deszkriptorokhoz** társítható egyéb azonosítók rögzítéséhez a Relexben résztezauruszokat kell létrehozni, amelyek az adott azonosítók listáját tartalmazzák. A HUNMARC-tervezet szerinti ezekhez kapcsolódó relációkat már létrehozták a szerkesztőfelületen (a részletekért és az apróbb HUNMARC-MARC 21 eltérésekért lásd a B mellékletet: a Relex MARC-exportjának lehetőségei), tehát amennyiben a jövőben felmerül ilyen irányú adatgazdagítás, akkor erre a Relexen belül is lehetőség kínálkozik.

3.3. A rekordfej és a 008-as mező beállításai

A rekordfej és a meghatározott jellemzők és információs adatokat tartalmazó 008-as mező beállítása is korrekcióra szorul a MARC 21-re való megfelelő konverzió érdekében. A HUNMARC-tervezet és a MARC 21 rekordfeje közötti egyetlen javasolt különbség a 05-ös karakterpozíción az *n* érték *c*-re cserélése, ezzel jelezve a rekord státuszát új helyett javítottként.

Tezauruszcikk rekordfeje MARC 21-ben:

```
00000cz###2200000n##4500
```

A 008-as mezőben a rekord számos tulajdonságát lehet előre kódolni. A MARC 21 esetében az alkalmazott **tárgyszórendszert/információkereső tezauruszt** azonosító 11. pozíció az első, ami eltér a HUNMARC-tervezettől, ahol a Köztaurusznak saját kódot javasoltak. Ehelyett az egyéb rendszert jelölő *z*-t kell alkalmazni, közösen a 040-es mező \$f almezőjével (illetve az intézményt jelölő \$a és az alkalmazott nyelvet jelölő \$d almezőket), amelyben megadható a tezaurusz neve vagy MARC-kódja (utóbbit, ha igényeltünk a MARC 21 fejlesztőtől). A HUNMARC-tervezetben nem alkalmazott 12-es és 13-as pozíciókon a MARC 21-ben a tételfejben esetlegesen rögzített sorozatokra vonatkozó információkat lehet kódolni. A Köztauruszban ez nem releváns, tehát mindkettő érték *n*. A 16-os pozíció szintén nem alkalmazott a HUNMARC-tervezetben, a MARC 21-ben ezzel azt lehet jelölni, hogy a rekordon belül az 1XX mező tartalma alkalmas-e arra, hogy egy bibliográfiai rekordban betöltse a sorozati címre vonatkozó hozzáférési pont szerepét. Mivel ez is irreleváns a Köztaurusz tekintetében, az érték *b*, mint „nem alkalmas”. A 28-as pozíció a tételfejben szereplő egyes testületi és földrajzi nevek hatásköri (kormányzati) szintjét jelöli. Többnyire az érték az irrelevanciát jelölő #, azonban egyes esetekben, a majdani tartalmi revízió és karbantartás során egyedileg kódolhatók az önálló vagy félig önálló területi testületek (*a*), a multilokális (*c*), országos (*f*), nemzetközi-kormányközi (*i*), helyi (*l*), nagyrégiós (*m*), hivatalos testületi és földrajzi nevek, illetve a meghatározatlan joghatóságú kormányzati szervek (*o*), állami és regionális kormányzati szervek (*s*). Ha nem lehet meghatározni, hogy a tételfej (vagyis a **deszkriptor**) kormányzati szerv-e, akkor ezt az *u* érték kódolja. Az ezekkel nem kódolható joghatóságú egyéb kormányzati szervek jelölése *z*.



Az utalók és a tételfejben szereplő névalak szabványosságát jelölő utalóértékű 29-es pozíció nem szerkeszthető a Relexben, ahol alapértelmezett módon a # karaktert kódolja a szoftver, amely azonban sem a HUNMARC-tervezet, sem a MARC 21 tekintetében nem szabványos. A majdani tartalmi revízió alkalmával, az utalók felülvizsgálata során a kóddal jelölhető, hogy a **vezérszó** és az **utaló** összhangban vannak egymással.

Meghatározott jellemzők és információs adatok a HUNMARC-tervezetben:

008 011231nn#ano##ba#n#####ana#####

MARC 21-ben:

008 230512nn#anznbnbn#####a#ana#####

040 ## \$a hubposzk \$b hun \$f koztaurusz

Összegzés

A Köztaurusz megfelel az építése idején érvényes, azóta azonban visszavont magyar szabványnak (MSZ 3418:1987). A jelenleg hatályos ISO-szabványnak (ISO 25964-1:2011) való megfelelés nagyrészt megállapítást nyert, ám egy fontos relációtípust, az egyes fogalmak egyedeit jelölő előfordulás-relációt nem szabványosították Magyarországon, ebből eredően a Köztauruszban sincsenek ezek kifejezve, sőt a jelenlegi gyakorlat, bizonyos szemszögből azonban logikailag megengedhető módon, generikus alárendeltként rögzíti az egyedeket, többnyire földrajzi nevek esetében. Előíró szabványosság tekintetében, szabványos információkereső teauruszként ennek az egyetlen feltételnek nem felel meg jelen formájában a Köztaurusz. Ezenkívül néhány, az ISO-szabvány szerint ajánlott (tehát nem előíró, változtatást tulajdonképpen nem igénylő) megoldásban tér még el a Köztaurusz a szabványtól, amely szerint:

- a **homográf** szavak esetében minden névalaknál hátravetett zárójeles értelmezők használata ajánlott,
- a VAGY logikai operátorok használata helyett zárójeles értelmezőkkel megkülönböztetett **utalók** használata ajánlott (ez a LÁSD VAGY és HELYETT VAGY **utalókkal** kapcsolt **lexikai egységeket** érinti).

A magyar szabvány a nemzetközi gyakorlaton túlmutatóan szabványosította a rezultáns-előfeltétel (R/E) reláció formájában a rokonsági kapcsolat egyik altípusát (következtetési reláció). Ez a megoldás sem mond ellen a hatályos teaurusz-szabványnak, amely szintén megadja a lehetőséget további relációk szabványos kifejezésére. A Köztaurusz a Magyar Nemzeti Névtérrel való összekapcsolását, illetve magának a szótárnak kapcsoltadat-alapú teauruszá (a teauruszszabvány terminológiája szerint **ontológiává**) való továbbfejlesztésében a jelen vizsgálat szempontrendszerét adó ISO 25964-1:2011 teauruszszabvány, illetve az ISO 25964-2:2013 szabvány normáit kell követni. Utóbbi a teauruszszabvány első részében leírt **információkereső teaurusz** kompatibilitását tárgyalja más típusú szótárakkal, illetve leírja az egyes szótárakban reprezentált **fogalmak** közötti megfeleltetés szabályait és módszereit.

Ami a MARC21-formátumnak való megfelelést illeti, alapvető kihívást jelent, hogy a Köztauruszt és a Relexet egyaránt a tervezett, de könyvtári szabályzatként be nem vezetett „HUNMARC besorolási rekordformátumra” alapozták. A Relex a teauruszcikkek MARC-rekordjait ténylegesen a tervezet szerint kódolja annak ellenére, hogy gyakorlatilag nincs olyan integrált könyvtári rendszer,

„A magyar szabvány a nemzetközi gyakorlaton túlmutatóan szabványosította... a rokonsági kapcsolat egyik altípusát (következtetési reláció).”

amely alkalmas lenne a magyar formátumtervezetben kialakított MSZ 3418:1987 szerinti tezaurusrelációk könyvtári katalógusokon belüli megfelelő kezelésére. A tezaurusrelációk közül a MARC 21 alapértelmezett módon csak a szimmetrikus rokonsági reláció (kétirányú lásd még **utaló**), illetve a generikus (fajta-nem) alá-fölérendeltség kifejezésére képes. Elvileg a MARC 21 lehetőséget kínál a tezauruszcikkek rekordjai közötti többi kapcsolattípus szabványos kifejezésére, ennek gyakorlati kipróbálása lesz a Köztaurusz továbbfejlesztését szolgáló munka következő fázisa. A Köztaurusz MARC 21-gyel való kompatibilitása még az alábbi apróbb változtatásokat igényli:

- a címekre vonatkozó deskriptoroknál (ha vannak) a kronológiai kiegészítő almező javítása 130 #0\$d-re;
- címeknél a művet nyelvét jelölő új almező 130 #0\$;
- címeknél a zeneművek számozási adatainak almezőjét javítani 130 #0\$n-re;
- címeknél a részmű címét jelölő almezőt javítani 130 #0\$p-re;
- címeknél a változatot jelölő almezőt javítani 130 #0\$s-re;
- a magyarázat megjegyzéstípust átállítani a 677 #a almezőre;
- a megjegyzés a használatról megjegyzéstípust átállítani 680 ##\$i almezőre;
- életrajzi vagy igazgatástörténeti megjegyzésnél az indikátorral jelölni az igazgatástörténetet: 678 1#a;
- 016-024 mezőknél az azonosítóknál az első indikátort 7-re állítani, \$2-ben jelölni az országos bibliográfiai központot;
- földrajziterület-azonosító ISO igazgatássegység-kódjánál a 043 ##\$c almező használata;
- 700, 711, 730 mezőknél (ha vannak) az első indikátort egyedileg beállítani a megfelelő értékre, a második indikátor 7, ekkor \$2-ben HuBpOSZK kód jelzi az OSZK-t.

A következő részben a Finnországban alkalmazott, feladatkörét tekintve hasonló tezausból továbbfejlesztett ontológiát és tárgyszórendszert mutatom be, illetve mindazon követelményeket ismertetem, amelyeket a Köztaurusz fejlesztése során teljesíteni kell annak érdekében, hogy egy, a finnhez hasonló, már gépi tárgyszavazásra is felhasználható, korszerű, 21. századi tárgyszórendszerré válhasson a Köztaurusz.

Felhasznált források

Cserbák A. (1995) *Geohistaurus, a Kárpát-medence humángeográfiai tézaurusza, 1723-1983*, Elérhető: <http://www.mek.iif.hu/porta/szint/egyeb/szotar/cserbak/> (Utolsó elérés: 2023.05.23.)

Getty Thesaurus of Geographic Names Online, Elérhető: <https://www.getty.edu/research/tools/vocabularies/tgn/> (Utolsó elérés: 2023.05.23.)

Halassy, B. (1994) *Az adatbázistervezés alapjai és titkai – Avagy az út az adattól az adatbázison át az információig*, [online], Elérhető: <https://mek.oszk.hu/11100/11123/11123.pdf> (Utolsó elérés: 2023.05.24.)

Harping, P. (2010) *Introduction to Controlled Vocabularies: Terminology for Art, Architecture, and Other Cultural Works*, Elérhető: https://www.getty.edu/research/publications/electronic_publications/intro_controlled_vocab_relationships.html (Utolsó elérés: 2023.05.23.)



Horváth, T., Varga, D. (1977) *Információs tezauszok*, Budapest, Népművelési Propaganda Iroda

International Organization for Standardization (1974) *ISO 2788–1974 Documentation – Guidelines for the establishment and development of monolingual thesauri*, Geneva

International Organization for Standardization (1986) *ISO 2788–1986 Documentation. Guidelines for the establishment and development of monolingual thesauri*, Geneva

International Organization for Standardization (2011) *ISO 25964–1:2011 Information and Documentation. Thesauri and interoperability with other vocabularies Part 1: Thesauri for information retrieval*, Geneva

International Organization for Standardization (2013) *ISO 25964–2:2013 Information and Documentation. Thesauri and interoperability with other vocabularies. Part 2: Interoperability with other vocabularies*, Geneva

Library of Congress (2022a) *MARC 21 Format for Authority Data*, 1999 Edition, update No. 1 (October 2000) through Update No. 35. (December 2022), Elérhető: <https://www.loc.gov/marc/authority/> (Utolsó elérés: 2023.05.25.)

Library of Congress (2022b) *MARC 21 Format for Classification Data*, 2000 Edition Update No. 1 (October 2000) through Update No. 37 (December 2023), Elérhető: <https://www.loc.gov/marc/classification/> (Utolsó elérés: 2023.05.25.)

Magyar Szabványügyi Hivatal (1976) MSZ 3418–76. *Magyar nyelvű információkereső tezauszok szerkezete, részei és formái*, Budapest

Magyar Szabványügyi Hivatal (1987) MSZ 3418–87. *Magyar nyelvű információkereső tezauszok szerkezete, részei és formái*, Budapest

Magyar Szabványügyi Testület (2023) *Szabványkereső, MSZ 3418:1987*, Elérhető: <https://ugyintezes.msz.hu/webaruhaz/szabvany-adatok?standard=88864> (Utolsó elérése: 2023.05.09.)

Molnár, I. (1973) *Az információkezelés fogalomrendszere (Az információs szaktezausz fogalma, rendszere és felépítése)*, Budapest, Közgazdasági és Jogi Könyvkiadó

National Information Standards Organization (2023) *ISO 25964 – the international standard for thesauri and interoperability with other vocabularies*, Elérhető: <https://www.niso.org/schemas/iso25964> (Utolsó elérés: 2023.05.09.)

Országos Széchényi Könyvtár (2009) *KSZ 4.2 HUNMARC a besorolási rekordok adatcseré formátuma* (Tervezet) [belső anyag] Budapest

Országos Széchényi Könyvtár, Könyvtári Intézet (2005) *Egyetemes Tizedes Osztályozás*. 1. kötet Táblázatok. 1. rész, Budapest, UDC Publ. No. P057.

Országos Széchényi Könyvtár, Magyar Könyvtárosok Egyesülete (2002) *OSZK TEZAURUSZ/KÖZTAURUSZ. Az Országos Széchényi Könyvtár és a Közművelődési könyvtárak átfogó tezauszja*. 3.0 változat, Budapest, Elérhető: <https://mek.oszk.hu/00700/00769/html/#222> (Utolsó elérés: 2023.05.24.)

RELEX webes felület (benne a naprakész Köztauruszsal), Elérhető: <https://mokka.hu/relex/index.html> (Utolsó elérés: 2023.05.25.)

Szakadát, I. (2005) Réteges struktúra, alaprelációk, In: *III. Magyar Számítógépes Nyelvészeti Konferencia*, Szeged. 2005. 12. 8-9, p. 43–55, Elérhető: http://acta.bibl.u-szeged.hu/58563/1/msznykonf_003_043-055.pdf (Utolsó elérés: 2023.05.24.)

Szente, P. (1982) *Néhány szó a szabványosításról – Beszélgetés Fügedi Péterné főosztályvezetővel*, OSZK Híradó, 25(07–08), p. 145–147.

Takáts Gyula Könyvtár (2023) *Központi könyvtár katalógusa*, Elérhető: <http://opac3.kaposvar.monguz.hu/> (Utolsó elérés: 2023.05.25.)

- Ungváry R. (2001a) *Az OSZK tezaszusa és a KÖZTAURUSZ*, Könyvtári Figyelő, 47(01), p. 11–40, Elérhető: <http://ki2.oszk.hu/kf/kfarchiv/2001/1/ungvary.html> (Utolsó elérés: 2024.03.19.)
- Ungváry R. (2001b) *Tezaszusa a felhasználói felületen. Az optimális tezaszusz megjelenítés problémái*, Tudományos és műszaki tájékoztatás, 48(3), p. 99–108, Elérhető: <https://journals.bme.hu/tmt/article/view/32034/17736> (Utolsó elérés: 2024.03.19.)
- Ungváry R. (2001c) *Tezaszuszkezelő program és a RELEX*, Tudományos és műszaki tájékoztatás, 48(1), p. 4–15, Elérhető: <https://journals.bme.hu/tmt/article/view/32052/17754> (Utolsó elérés: 2024.03.19.)
- Ungváry R. (2007) *Ontológiai és tezaszuszok csúcspogalmi*, Tudományos és műszaki tájékoztatás, 54(10), p. 441–459, Elérhető: <https://journals.bme.hu/tmt/article/view/33443/19158> (Utolsó elérés: 2024.03.19.)
- Ungváry R. (2009) *Diplomatezaszuszok hibái és tanulságai*, Tudományos és műszaki tájékoztatás, 56(5), p. 211–228, Elérhető: <https://journals.bme.hu/tmt/article/view/33208/18923> (Utolsó elérés: 2024.03.19.)
- Ungváry R. (2010) *MARC21/HUNMARC: a besorolási adatok metaadat-formátuma. Főbb jellemzők, fejlődés és problémák*, Könyvtári Figyelő, 56(1), p. 9–70.
- Ungváry R. (2011) *A magyar tezaszuszszabvány és Horváth Tibor nemzedéke*, Könyv és nevelés, [online] 4. sz. melléklet, Elérhető: <https://folyoiratok.oh.gov.hu/konyv-es-neveles/a-magyar-tezaszusz-szabvany-es-horvath-tibor-nemzedeke> (Utolsó elérés: 2023.05.16.)
- Ungváry R. (2012) *RELEX névtér- és tezaszusz-szerkesztő rendszer: Több felhasználó számára hozzáférhető kliens-szerver alkalmazás névterek (tulajdonnév- és köznévterek), tezaszuszok, taxonómiák és osztályozási rendszerek (például ETO) osztott kezelésére egy és többnyelvű környezetben az Interneten keresztül: Felhasználói kézikönyv: 1.0 változat*, Budapest
- Ungváry R. (2014) *Vajda Erik elméleti munkássága*, Tudományos és műszaki tájékoztatás, 61(6), p. 242–244, Elérhető: <https://journals.bme.hu/tmt/article/view/34281/19996> (Utolsó elérés: 2024.03.19.)
- Ungváry R. (2016) *Névterek és földrajzinév-tárak*, Tudományos és műszaki tájékoztatás, 63(4), p. 135–157, Elérhető: <https://journals.bme.hu/tmt/article/view/34212/19927> (Utolsó elérés: 2024.03.19.)
- Ungváry R. (2018a) *A névterek értelme. Filozófiai–szerkezeti jellemzők*, Tudományos és műszaki tájékoztatás, 65(1), p. 1–14, Elérhető: <https://journals.bme.hu/tmt/article/view/35218/20931> (Utolsó elérés: 2024.03.19.)
- Ungváry R. (2018b) *Névtérkezelő készítésének nehézségei*, Tudományos és műszaki tájékoztatás, 65(10), p. 502–517, Elérhető: <https://journals.bme.hu/tmt/article/view/35042/20755> (Utolsó elérés: 2024.03.19.)
- Ungváry R. (összeáll.) (1998) *HUNMARC, a besorolási rekordok adatcsere formátuma: tervezet*, [kézirat] [Rövidített – „CONCISE” – formátum; kész. Sipos Márta]
- Ungváry R. (szerk.) (2001) *Geotaurusz és geohistaurusz: földrajzi nevek és humáangeográfiai nevek tezaszusa*, 3.0 változat, Elérhető: <https://mek.oszk.hu/00000/00070/html/> (Utolsó elérés: 2024.02.29)
- Wikipédia (2022) *is-a*, Elérhető: <https://hu.wikipedia.org/wiki/Is-a> (Utolsó elérés: 2023.05.24.)
- Wikipédia (2023) *is-a*, Elérhető: <https://en.wikipedia.org/wiki/Is-a> (Utolsó elérés: 2023.05.24.)