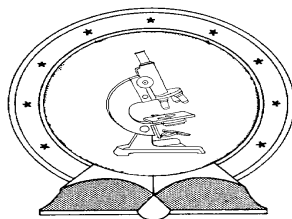


DE TTK



1949

Az ETO-jelzetek automatikus interpretálásának és elemzésének kérdései

Egyetemi doktori (PhD) értekezés

Piros Attila

Témavezető: Dr. Boda István Károly

DEBRECENI EGYETEM

Természettudományi Doktori Tanács

Matematika- és Számítástudományok Doktori Iskola

Debrecen, 2018.

Ezen értekezést a Debreceni Egyetem Természettudományi Doktori Tanács Matematika- és Számítástudományok Doktori Iskola Számítástudomány és alkalmazásai programja keretében készítettem a Debreceni Egyetem természettudományi doktori (PhD) fokozatának elnyerése céljából.

Nyilatkozom arról, hogy a tézisekben leírt eredmények nem képezik más PhD disszertáció részét.

Debrecen, 2018.

Piros Attila
doktorjelölt

Tanúsítom, hogy Piros Attila doktorjelölt 2007- 2011 között a fent megnevezett Doktori Iskola Számítástudomány és alkalmazásai programjának keretében irányításommal végezte munkáját. Az értekezésben foglalt eredményekhez a jelölt önálló alkotó tevékenységével meghatározóan hozzájárult. Nyilatkozom továbbá arról, hogy a tézisekben leírt eredmények nem képezik más PhD disszertáció részét.

Az értekezés elfogadását javasolom.

Debrecen, 2018.

Dr. Boda István Károly
témavezető

Az ETO-jelzetek automatikus interpretálásának és elemzésének kérdései

Értekezés a doktori (Ph.D.) fokozat megszerzése érdekében
a matematika- és számítástudományok
tudományágban

Írta: Piros Attila okleveles matematika tanár és informatikus könyvtáros

Készült a Debreceni Egyetem Matematika- és Számítástudományok doktori iskolája
(Számítástudomány és alkalmazásai programja) keretében

Témavezető: Dr. Boda István Károly

A doktori szigorlati bizottság:

elnök: Dr. Hajdu Lajos

tagok: Dr. Csörnyei Zoltán

Dr. Mihálydeák Tamás

A doktori szigorlat időpontja: 2017. október 18.

Az értekezés bírálói:

Dr.

Dr.

Dr.

A bírálóbizottság:

elnök: Dr.

tagok: Dr.

Dr.

Dr.

Dr.

Az értekezés védésének időpontja: 2018.

Tartalomjegyzék

Bevezetés.....	1
1 Osztályozáselméleti alapfogalmak.....	3
1.1 Az osztályozás jelentése	3
1.2 Az osztályozási rendszerek szerkezete	4
1.3 A szintaktikai relációk	6
1.3.1 A szintaktikai relációk típusai	8
1.3.2 A szintaktikai relációk jelentősége	9
1.4 Az osztályozási rendszerek típusai.....	11
1.4.1 Az enumeratív osztályozási rendszerek	11
1.4.2 A facettás osztályozási rendszerek	12
1.4.3 Az analitikus-szintetikus osztályozási rendszerek	14
1.5 Irodalomjegyzék	15
2 Az Egyetemes Tizedes Osztályozás	19
2.1 Az Egyetemes Tizedes Osztályozás története és jelentősége	19
2.2 Az ETO elterjedtsége.....	23
2.2.1 Az ETO Magyarországon	26
2.3 Irodalomjegyzék	28
3 Az ETO mint analitikus-szintetikus osztályozási rendszer	37
3.1 Az ETO jellege	37
3.2 Az ETO mint hierarchikus rendszer.....	38
3.3 Az ETO mint facettás rendszer.....	40
3.3.1 A korlátozottan közös alosztások.....	40
3.3.2 Nem korlátozottan közös alosztásokkal kifejezett facetták.....	41
3.4 Az ETO mint analitikus-szintetikus rendszer.....	42
3.4.1 Általánosan közös alosztási táblázatok	42
3.4.2 Jelzetek összekapcsolására szolgáló általánosan közös alosztási jelek.....	44
3.4.2.1 Összekötés és kiterjesztés.....	44
3.4.2.2 Viszonyítás és sorrendrögzítés.....	46
3.4.2.3 Csoportképzés.....	47
3.4.3 A jelzetek besorolására vonatkozó szabályok.....	48
3.5 A kontextus jelentősége	50
3.5.1 Az alosztási jelek jelentésmódosító hatása.....	51

3.5.2	A korlátozottan közös alosztások.....	52
3.5.3	Az általánosan közös alosztások	53
3.6	Irodalomjegyzék	53
4	Az ETO gépesítése	59
4.1	Történeti áttekintés.....	59
4.2	A programokkal szemben támasztott elvárások	65
4.3	A táblázatok gépesítése.....	66
4.3.1	Az ETO mesterfájl („UDC Master Reference File”).....	66
4.3.2	Az ETO Online elérése	68
4.3.2.1	UDC Online.....	69
4.3.2.2	UDC Summary	69
4.4	Az információ visszakeresés támogatása	70
4.4.1	A jelzetek természetes nyelvű visszakereshetősége.....	73
4.4.2	Böngészés.....	75
4.4.2.1	A jelzetek automatikus sorba rendezése	77
4.4.3	Az ETO-jelzetek szerinti keresés.....	80
4.5	Az osztályozási munka támogatása	81
4.5.1	A jelzetek felbontása.....	83
4.6	A jelzetek eltárolása.....	83
4.6.1	A jelzetek, „ahogy vannak”	84
4.6.1.1	A jelzetek karaktersorozatként való kezelése	85
4.6.1.2	A jelzetelemek KWOC-indexe	87
4.6.1.3	A jelzetelemek KWIC-indexe	88
4.6.2	A kombinált jelzetek strukturált adatként való eltárolása.....	89
4.6.2.1	Osztályozási MARC formátumok.....	90
4.6.2.1.1	MARC 21 Osztályozási Formátum	90
4.6.2.1.2	UNIMARC Osztályozási Formátum	91
4.7	Irodalomjegyzék	92
5	Az ETO számok szerkezetének leírása.....	103
5.1	Bevezetés.....	103
5.2	A formátummal szemben támasztott követelmények	104
5.3	A jelzetelemek fája	104
5.4	Az XML nyelv előnyei	107

5.5	XML séma definíció ETO-jelzetek leírására.....	108
5.5.1	Áttekintés.....	108
5.5.2	Az ETO szám leírása XML dokumentumként	110
5.5.3	Az egyszerű fogalmakat leíró típus	111
5.5.4	A táblázati számokat leíró osztályok ősosztályai	112
5.5.5	A főtblázati számokat leíró típusok.....	113
5.5.6	A korlátozottan közös alosztásokat leíró típusok	116
5.5.6.1	Jelzetszintézis.....	119
5.5.7	Az általánosan közös alosztásokat leíró típusok	120
5.5.7.1	Önálló általánosan közös alosztások.....	120
5.5.7.1.1	Nyelvi általánosan közös alosztások.....	121
5.5.7.1.2	Formai általánosan közös alosztások	125
5.5.7.1.3	Hely szerinti (földrajzi) általánosan közös alosztások	127
5.5.7.1.4	Az emberek származása, etnikai csoportok és nemzetiség szerinti (etnikai) általánosan közös alosztások.....	130
5.5.7.1.5	Idő szerinti általánosan közös alosztások.....	132
5.5.7.1.6	Külső forrásból vett jelzetek és A/Z névalosztások	136
5.5.7.2	Nem önálló általánosan közös alosztások.....	137
5.5.7.2.1	Szempontra szerinti általánosan közös alosztások.....	137
5.5.7.2.2	Általános ismérvek szerinti általánosan közös alosztások	138
5.5.7.2.2.1	Tulajdonságok általánosan közös alosztásai.....	138
5.5.7.2.2.2	Anyagok szerinti alosztások.....	140
5.5.7.2.2.3	Kapcsolatok, folyamatok és műveletek közös alosztásai	141
5.5.7.2.2.4	Személyek szerinti alosztások.....	142
5.5.7.3	Jelzetek összekapcsolására szolgáló általánosan közös alosztási jelek.....	143
5.5.7.3.1	Összekötés.....	143
5.5.7.3.2	Viszonyítás.....	144
5.5.7.3.3	Sorrendrögzítés	144
5.5.7.3.4	Csoportképzés	145
5.6	Példák az XML alkalmazására	146
5.7	Irodalomjegyzék	149
6	Az összetett ETO számok automatikus interpretálása.....	153
6.1	Bevezetés.....	153
6.2	Az ETO számok automatikus felbontása.....	154

6.2.1	Az ETO számok felbontásának néhány problémája	155
6.3	Az interpretálás alapelvei	157
6.4	Az interpreter működésének alapjai	158
6.5	A program kimenete.....	160
6.5.1	Példák az ETO számok elemzésére és interpretálására XML formátumban.....	160
6.5.2	KWOC index	162
6.5.3	A jelzetelési hibák felismerése	163
6.6	A szoftver tesztelése.....	165
6.6.1	Esettanulmányok	165
6.7	További fejlesztési irányok és kutatási tervek.....	166
6.7.1	A szoftver fejlesztése	166
6.7.2	A szoftver és a formátum felhasználása egyéb területeken.....	166
6.8	Irodalomjegyzék	167
	Összefoglalás	171
	Summary	173
	Köszönetnyilvánítás.....	175
	Függelék	177
	A szerző publikációi	177
	Referált publikációk.....	177
	Lektorált publikációk	177
	Egyéb publikációk.....	178
	Konferencia előadások	178

„Az ETO szép. Szép, mint minden, ami szervesen, emberek ezreinek hozzáértő, évszázados közreműködésével született.”¹ (Ungváry Rudolf)

Bevezetés

Charles Kay Ogden és Ivor Richards híres munkájukban a nyelvi szimbólumok és azok által leírt objektumok közötti kapcsolatot egy háromszöggént modellezték, melynek csúcsai az *objektumot*, az azt leíró *gondolatot*, illetve a gondolatot jelölő *szimbólumokat* jelentik. *Osztályozási rendszerek* esetében ezek a csúcsok a leírt objektumot (dokumentum, szöveg, műalkotás stb.), annak tárgyának az indexelő általi összefoglalását, végül ennek az összefoglalásnak egy osztályozási rendszer segítségével történő kifejezését jelölik. A relevanciára, illetve a dokumentum azonosítására vonatkozó *kérdések hatékony megválaszolásának alapvető feltétele a szimbólum mögötti gondolat lehető legpontosabb felismerése*. Az információkeresés során központi kérdés ennek a folyamatnak a *hatékony gépi támogatása*.

Az *Egyetemes Tizedes Osztályozás (ETO)* az egyik legnagyobb múltra visszatekintő és legelterjedtebb osztályozási rendszer. Ezt részben *analitikus-szintetikus* jellegének köszönheti: míg táblázatai tudományos mélységgel foglalják össze az egyszerű tárgyköröket, számos lehetőséget nyújt a bonyolultabb, összetettebb tartalmak szintaktikai relációk segítségével történő kifejezésére is. Egy ilyen osztályozási rendszer esetében a szintaktikai relációk sokrétűsége és központi szerepe miatt előtérbe kerül a *kompozicionalitás* elve, ami szerint a kifejezések (ebben az esetben osztályozási jelzetek) jelentését azok elemei és az elemek kapcsolódási módja együtt határozzák meg. Ennek *figyelembe vétele növeli az információ kereső rendszerek hatékonyságát*, különös tekintettel a *precízióra*, azaz a találatok keresőkérdésnek való megfelelésének mértékére. Az osztályozás hierarchikus táblázatai elsősorban az *inkluzív – az alacsonyabb hierarchiaszintekre is kiterjedő – keresés* számára biztosítanak kiváló feltételeket, ennek azonban előfeltétele a jelzetelemek pontos beazonosítása.

Az ETO gépesítésének hetvenéves története alatt *a rendszerek készítői többnyire figyelmen kívül hagyták* a kompozicionalitás elvét, lemondva a szintaktikai relációk által hordozott információról. Ennek hátterében részben a rendszer komplexitása és szintaktikai sokrétűsége húzódik meg, amely miatt a lehetőségek számbavétele meglehetősen bonyolult feladat. Másrészt a ma is használatos módszerek kialakulásakor a jelzeteknek az elemeik alapján történő visszakereshetővé tétele és a jelzetek között a priori fennálló szemantikai kapcsolatok feltárása kielégítő megoldásnak tűnt.

Kutatásom elsődleges célja annak megmutatása, hogy *a ma elterjedt leíró és programozási nyelvek már alkalmasak arra, hogy az ETO-jelzetek összeállítását, visszakeresését, elemzését kompromisszummentesen, a rendszer teljes szintaktikai gazdagságát figyelembe véve támogassuk*. Ezt a hipotézist egy olyan, számítógép és ember számára egyaránt jól olvasható **formátum kidolgozásával** kívánom bebizonyítani, amely alkalmas arra, hogy ne csak az ETO-jelzetek részeit, de azok teljes szintaktikai struktúráját eltároljuk benne.

A kutatás során abból a feltevésből indultam ki, hogy *a hatékony osztályozásnak előfeltétele az, hogy az indexelőtől csak a dokumentumok minél pontosabb leírását várjuk el*, teljes mértékben automatizálva a jelzetek felbontását és reprezentációját. Ezt az automatizálást a rendszerek és formátumok közötti átjárhatóság és a korábbi adatbázisok rekordjainak retrospektív konverziója is indokolja. Ezért a **kutatás második fázisának célja** egy olyan **interpreter alkalmazás kifejlesztése** volt,

¹ Ungváry 2000.

amely képes arra, hogy az ETO-jelzeteket emberi beavatkozás nélkül, szintaktikai alapon elemezze és struktúrájuk megtartása mellett átalakítsa őket az előzőleg bemutatott formátumba.

Az **interpreter újdonságát** nem csak az adja, hogy *a jelzetek teljes szintaktikai struktúráját kezeli* (és ennek köszönhetően képes olyan jelzetrészeket is felismerni, amelyeket a csak a részjelzetek beazonosítására törekvő jelzetelemző alkalmazások nem), hanem az is, hogy *követi a jelzetépítési szabályok változását*, és mindig annak az ETO-verzióknak a szabályai szerint elememzi a jelzetet, amelyik alapján azt összeállították.

A bemutatott formátum és interpreter nem csak állományok indexeléséhez lehet hasznos segédeszköz, hanem az osztályozástudomány területén folytatott bármely olyan elméleti és gyakorlati kutatás során, amelynek folyamán érdekes lehet az osztályozási jelzetek belső szerkezetének figyelembe vétele.

Dolgozatom első felében az ETO-val és annak gépesítésével kapcsolatos **szakirodalmat** tekintem át. Az *első fejezetben* azokat az alapvető *osztályozáselméleti ismereteket* és szakkifejezéseket ismertetem, melyek szükségesek a későbbiek megértéséhez. A *második fejezet az ETO történetét, jelentőségét és jelenlegi elterjedtségét* tekinti át. A *harmadik fejezet az ETO szerkezetét* mutatja be röviden, kiemelve azokat a speciális jellemzőket, amelyek leginkább befolyásolják a rendszer számítógépes kezelését. A *számítógépesítés kérdései* részletesebben a *negyedik fejezetben* vannak tárgyalva.

A disszertáció második része **saját eredményeket** tartalmaz. Az *ötödik fejezetben* részletesen ismertetem az **ETO-jelzetek szintakszist is megőrző reprezentálásának az alapelveit**, és az ezek alapján kidolgozott **formátumot**. A *hatodik fejezetben* a korábban említett **ETO-interpreter alapelveit és működését** mutatom be. Legvégül áttekintem, hogy az itt ismertetett eredmények milyen formában fejleszthetők tovább és milyen irányú kutatások alapjául szolgálhatnak a jövőben.

1 Osztályozáselméleti alapfogalmak

„Az osztályozás valamilyen elv vagy koncepció, cél vagy érdeklődési kör vagy ezek valamilyen kombinációja alapján felállított, valamilyen sorrend szerint rendezett osztályok sorozata vagy rendszere.”² (Henry Evelin Bliss)

1.1 Az osztályozás jelentése

Az **osztályozás** kifejezés a köznyelvben több értelemben is használatos, általában személyek, tárgyak, fogalmak csoportjainak adott célból, bizonyos szempontok szerint történő kialakítására, illetve az említett objektumoknak a kialakított csoportokba való besorolására.

Az *információkeresésben* az osztályozás a felhasználók számára szükséges információk, dokumentumok csoportosítása, egyedi információk leírása, illetve a kialakított csoportok rendezése, szelektálása annak érdekében, hogy biztosítsuk azok visszakereshetőségét.³ Erre a célra speciális információkereső nyelveket hozunk létre. Az osztályozás mint tevékenység a tartalmak visszakereshető módon történő kifejezése egy ilyen nyelven, azaz valamilyen információkereső nyelvre történő fordítás. Gyakorlatban ez azt jelenti, hogy egy gyűjtemény dokumentumait⁴ egyesével megvizsgálva meghatározzuk azok tárgyát, majd leírjuk azt egy olyan mesterséges nyelven, mely célzottan arra lett kialakítva, hogy a vele kifejezett fogalmak rendszerezhetőek – lehetőleg minél több szempont szerint és minél többféle módszerrel – visszakereshetőek legyenek.⁵

Siyali Ramamrita Ranganathan az osztályozástudomány alapjait lefektető művében az „osztályozás” szónak öt jelentését definiálja, a legalapvetőbb, hétköznapi értelmezéstől az elvontabb, rendszer szintű felé haladva. Ezek a felosztás (dolgok csoportosítása), rendteremtés vagy csoportosítás (osztályokba sorolás, meglévő dolgok ismérvek szerinti csoportosítása), csoportosítás jelzetteléssel (dolgok meglévő osztályokba sorolása és sorba rendezhető azonosítóval való ellátása), csoportosítás és jelzettelés folyamatos csoportalkotással (új jelzetek származtatása az újonnan megjelenő dolgok számára), végül a dolgok helyett a mögöttük álló fogalmak jelzetekkel való ellátása és csoportosítása, folyamatos csoportalkotással, azaz osztályozási rendszer létrehozása, karbantartása és alkalmazása.⁶ Az utóbbi értelmezés az információs célú osztályozási rendszerek létrehozását és alkalmazását jelenti.

Magyarul az osztályozás szót a rendszerre, annak létrehozására és az alkalmazására egyaránt szokás alkalmazni, ezért a terminológia tisztázása érdekében érdemes megemlíteni, hogy az angol és német

² Bliss 1929: 143.

³ Pálvölgyi 1998: 58.

⁴ Bár osztályozásról leggyakrabban könyvtári dokumentumok kapcsán beszélünk, az osztályozási rendszereket gyakran használják szövegek, műtárgyak, kutatási és adminisztrációs adatok, archívumok, online források stb. feltárására is.

⁵ Valójában annak, hogy mit is jelent a „dokumentum tárgya”, gazdag irodalma és sokféle értelmezése van. Míg Cutter számára a tárgy olyan intellektuális objektumot jelentett, amely szociális konszenzus alapján nyer elnevezést, Dewey a tárgyat a saját rendszerében megjelenő osztályokon keresztül értelmezte. Ranganathan tárgyon az általa alap- és összetett tárgyaknak nevezett fogalmakat értette (megkülönböztetve az izolátumoktól). Wilson a dokumentum tárgyát annak meghatározási módján keresztül ragadja meg (Hjørland 2017a).

⁶ Ranganathan 1967: 77-78.

nyelvben külön szó van az osztályok felállítására („Klassifizieren”, „to classify”), az osztályokba sorolásra („klassieren”, „to class”) és az osztályozási rendszerre („Klassifikation”, „classification”).⁷

Tartalmi feltárás, jelzetalkotás történhet szabad indexelés vagy valamilyen természetes nyelvű információkereső vagy tudásszervező rendszer (tárgyszólista, tezaurusz, taxonómia) alapján is, de osztályozásnak általában a hierarchikus, mesterséges nyelvű információkereső nyelveken (pl. ETO, DDC, BC2, LCC) történő leírást nevezzük.⁸ A két megközelítés közötti fő különbséget az jelenti, hogy az osztályozási rendszerek a *fogalmak szisztematikus elrendezésére* törekednek és a természetes nyelvű indexekre elsősorban csak mint segédeszközökre tekintenek. Ez az elrendezés szükségessé teszi valamilyen saját **jelrendszer** alkalmazását.⁹

Egy másik megközelítés szerint, ha a tartalmi feltárást annak fogalmi szintjén tárgyaljuk, akkor osztályozásról beszélünk, ha a nyelvin, akkor indexelésről; valójában a két kifejezés ugyanannak a tevékenységnek a két szintjét jelöli.¹⁰

1.2 Az osztályozási rendszerek szerkezete

Ranganathan fogalmazta meg először azt, hogy az osztályozási tevékenységnek **három szintje** van: *fogalmi*, *nyelvi* és *jelzetelési* szint. Az első szint a fogalmak, az objektumokról kialakított gondolatok szintje, a második szinten történik a fogalmak (természetes nyelvű) szavakkal történő kifejezése, legvégül a jelzet szint, az osztályozási nyelvre való fordítás következik (amely elvileg kiküszöböli a természetes nyelv olyan problémáit, mint a homonimák és a szinonimák kérdése). A legfontosabb szint a fogalmak szintje, a valódi osztályozás, a tartalom felismerése itt történik meg; a legvégső lépés, a **koordináció** (a tartalom leírása a fogalmakat kifejező osztályozási szavak összekapcsolásával) pedig a jelzet szinten megy végbe.¹¹

Látható az analógia az osztályozás szintjei és a valóság, a fogalmak és a megnevezések problémájának leírására általában használt ún. **Ogden-Richards-féle szemantikai háromszög** között, mely a lingvisztikai szimbólumok és az általuk reprezentált objektumok közötti kapcsolatot modellezi. A háromszög csúcaiban az objektum (referens), a róla a tudatunkban kialakított gondolati kép (fogalom vagy referencia) és a fogalom megnevezése (kifejezés vagy szimbólum) állnak.¹² A modellt az osztályozás folyamatára is alkalmazhatjuk. Ekkor az objektum a leírt dokumentumot, szöveget, műalkotást stb. jelenti, a gondolat az objektum tárgyának kifejezése az indexelő által, a szimbólum pedig ennek a gondolatnak a kifejezése az osztályozási rendszer nyelvi eszközeivel.¹³ Az első csúcs az indexelés szemantikai aspektusának felel meg, az utolsó pedig a szintaktikainak. A második csúcs a fogalmak szintjén a szemantikai aspektushoz tartozik, a fogalom természetes nyelvi megfogalmazása viszont szintaktikai kérdés is.¹⁴

⁷ Dahlberg 2001: 236-237., Bliss 1929: 142-143.

⁸ Jean Martin Perrault (Perrault 1974) megfogalmazása szerint „a relevancia egyik jellegzetessége, hogy hierarchikus értékítélet” és az osztályozás „csakis a hierarchikus alapon történő feltárára (dokumentum-szurrogátum készítésre) alkalmazható joggal”. Az angol nyelvű szakirodalomban ez a megkülönböztetés terjedt el. Magyarországon az Egyetemes Tizedes Osztályozás használata annyira általános, hogy az „osztályozás” szó gyakran az ETO használatának szinonimájaként jelenik meg.

⁹ Gnoli 2018.

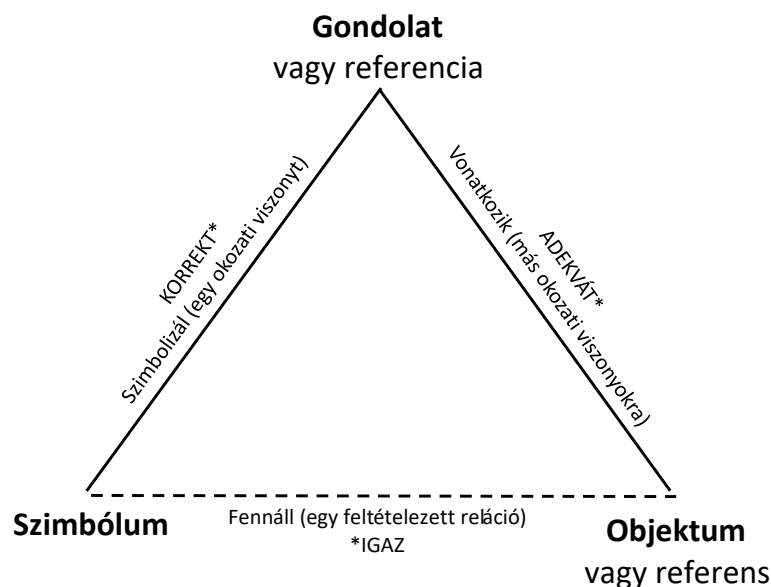
¹⁰ Ungváry és Orbán 2001: 349.

¹¹ Ranganathan 1967: 327-328.

¹² Ogden és Richards 1946: 10-12.

¹³ Piros 2017a.

¹⁴ vö: Iyer 2012: 127-128.



1. ábra Az Ogden-Richards-féle szemiotikai háromszög

Objektum bármi lehet, amiről filozófiai szempontból ítéleteket alkothatunk, logikai szempontból pedig állítások, nyelvi szempontból kijelentések tehetők. Az objektumról kijelentés révén létrejön az adott dolog fogalmának egy része. A róla tehető igaz állítások összessége rögzíti a tulajdonságokat, amelyek a dolgot meghatározzák. Tehát *a fogalom a tárgyról tett lehetséges állítások összessége*.¹⁵

Az egyes objektumokról állítható tulajdonságok összessége különböző lehet, ennek megfelelően változik az egyes fogalmak ismertetőjegyeinek száma, azaz a **fogalmak tartalma** is, amely *fordított arányban* áll a **fogalom terjedelmével**, azaz a benne általánosított objektumok összességével: ha a fogalom tartalma bővül további lényeges és specifikus ismertetőjegyek hozzáadásával, akkor a terjedelme szűkül, és fordítva, ha lényeges ismertetőjegyet veszünk el belőle, akkor több objektumot írhatunk le vele, tehát nő a terjedelme.¹⁶ Ebben az értelemben beszélhetünk fogalmak közötti *generikus vagy integrációs hierarchikus kapcsolatok*ról: ezek azok a kapcsolatok, amelyek egy fogalom és az ismertetőjegyek hozzáadásával, illetve elvételével kapott nála szűkebb, illetve bővebb tartalmú fogalmak között állnak fenn.¹⁷ Emellett létezhetnek más, például *partitatív* (egész és része), *csoport és elemei közötti* stb. kapcsolaton alapuló hierarchiák is.

Azok a fogalmak, melyeknek a tartalma már nem csökkenthető tovább, azaz nincs értelmes hierarchikusan felső kapcsolatuk, a **kategóriák**. Másképp fogalmazva a kategóriák azok a fogalmak, amelyekről már nem tehetők tartalmi, csak formális értelemben vett állítások.¹⁸ Kategóriák az olyan

¹⁵ Dahlberg 2001: 225.

¹⁶ Pálvölgyi 1998: 51.

¹⁷ Az integrációs szintek filozófiai vizsgálata Arisztotelészig nyúlik vissza, és az osztályozási és más tudásszervező rendszerekben is a kezdetektől fontos szerepet töltenek be, ennek megfelelően az irodalmuk is jelentős (Ungváry és Orbán 2000: 271-272). A közelmúltban két átfogó tanulmány is megjelent a témában. Az integrációs szintek és hierarchiák elméletét, történetét és tudásszervezésben betöltött szerepét Kleineberg (2017), a továbbosztás formális logikai hátterét Frické (2016) vizsgálja részletesen.

¹⁸ Dahlberg 2001: 227.

*legáltalánosabb fogalmak, mint mozgás, anyag, tulajdonság, hely, idő stb. A kategóriák adják az osztályozás alapjait, hiszen ezek határozzák meg azokat a szempontokat, amelyek szerint a dolgok ismertetőjegyeit meghatározzuk.*¹⁹ Például a facettás osztályozási rendszerek is mindig valamilyen alapvető kategóriarendszer alapján teszik lehetővé a fogalmak felbontását.²⁰

A FID/CR (a Nemzetközi Dokumentációs Szövetség Osztályozástudományi Bizottsága) definíciója szerint *az osztályozási rendszer létrehozása* generikus (nem-faji) vagy egyéb relációk létrehozását jelenti egyedi szemantikai egységek között, tekintet nélkül a rendszerben megvalósuló hierarchia fokára és arra, hogy a rendszert dokumentumok keresésének hagyományos vagy gépesített módszerével összefüggésben használják.²¹ Az osztályozási rendszernek tehát nemcsak a fogalmak és fogalmi kategóriák (azaz szemantikai egységek) alapvető elemei, hanem az azok közötti szemantikai (analitikus) relációk is.

A **szemantikai relációknak** azért van kiemelt szerepe, mert ezek alkotják az osztályozási rendszer szerkezetét²², mely az indexelés logikai alapját képezi és *lehetővé teszi a tájékozódást a fogalmak között*. A szerkezet hierarchikus osztályozási rendszerekben elsősorban a rendszer által tükrözött fogalmi hierarchiát jelenti (amely a fogalmakhoz rendelt jelzetekben is kifejeződik), emellett leggyakrabban az **asszociatív**, „lásd” és „lásd még” típusú relációk jelennek meg.

Minden fogalom része egy szemantikai kontextusnak, amelyet az osztályozási rendszer definiál²³, és szintaktikai kapcsolatban állhat más fogalmakkal. A dokumentum tartalmában megjelenő fogalmak ebbe a szemantikai és szintaktikai kontextusba beágyazva adják meg a dokumentum tárgyát, melyet az indexelésnek tükröznie kell.²⁴ Ezért *az osztályozási rendszer harmadik szerkezeti egységét a szintaktikai relációk jelentik*.

1.3 A szintaktikai relációk

A **tartalmi feltárás** (indexelés, illetve osztályozás) fő célja a dokumentum - vagy más feldolgozási egység - tartalmának leírása, azaz az egyedi tartalmának olyan formában történő feltárása, mely a lehető legnagyobb mértékben támogatja *annak megállapítását, hogy egy adott szöveg releváns-e egy adott keresés során*. A relevanciára vonatkozó döntések pedig megkövetelik *az ismeretek tematikai azonosításhoz szükséges minimumát*.²⁵

Másképp fogalmazva, mivel az indexelő és a kereső számára a relevancia mást jelent, és mivel a várható keresőkérdés az indexeléskor nem látható előre, *az indexelőnek a dokumentum adottságait kell rögzítenie*²⁶ a megfelelő részletességgel. A leírásnak elégnie kell lennie a dokumentum témájának azonosításához olyan mértékben, ami alapján már eldönthető, releváns-e egy adott kérdésre. Úgy is fogalmazhatunk, hogy **a dokumentum tárgyi szurrogátumát** kell elkészíteni és lefordítani egy adott

¹⁹ Iyer 2012: 40-59.

²⁰ A legismertebb ilyen szempontrendszereket Ranganathan és Vickery állították össze (Iyer 2012: 92-93).

²¹ Ungváry és Orbán 2001: 239.

²² Dahlberg 2001: 239.

²³ Az authority állományokról, amelyek az osztályozási rendszer feletti szinten adhatnak további szemantikai kapcsolatokat a fogalmakhoz egy későbbi fejezetben lesz részletesebben szó.

²⁴ Gödert 2014.

²⁵ Horváth és Sütő 2001: 37.

²⁶ Perrault 1974.

információkereső nyelvre. Ezt a megközelítést **tartalom központú megközelítésnek** is szokás nevezni, szembe állítva a **kérdés alapú indexeléssel**, melynek során az indexelő a célcsoport feltételezhető keresőkérdéseit veszi számba, eldöntve, melyekre releváns a leírt dokumentum, és ez alapján készíti el a leírást.²⁷

A **tárgy fogalma** és annak meghatározása az osztályozástudomány fontos kérdése, melyet a különböző teoretikusok más-más irányból igyekeztek megközelíteni.²⁸ A gyakorlatban a tárgyi leírás általában magában foglalja annak a meghatározását, *amiről a dokumentum szól* („aboutness”), illetve olyan *egyéb aspektusokat* (pl. forma, nyelv, műfaj), melyek azon kívül jelentősek lehetnek a felhasználók számára a visszakereséskor vagy a relevanciára vonatkozó döntések meghozatalakor („isness”).²⁹

A leírás során meg kell határozni a dokumentum tartalmát leíró fogalmakat.³⁰ Ez azonban kevés, mivel a jelentés nem az izolált kifejezésekben, hanem a kifejezés-komplexumokban adott, *egy objektum nem egyenlő ismérvei halmazával*.³¹ Ezért *a fogalmak egymástól független szerepeltetése zajt okozhat*. Például az „énekesmadarak” és a „madarak vándorlása” kulcsszavak együttes használata a viszony jelölése nélkül az „énekesmadarak vándorlási szokásai” és a „vándorló énekesmadarak” leírásakor egyaránt zajt fog okozni.³² Az ilyen feltárást **posztkoordináltnak** is szokták nevezni, abból kiindulva, hogy a fogalmak összekapcsolása csak a keresési fázisban történik meg.³³

A fentiek alapján a tartalmat leíró fogalmak mellett célszerű meghatározni azok kapcsolódási módjait is. *Dagobert Soergel* posztulátumában egyenesen azt mondja ki, hogy *a tartalmi szurrogátumnak tükröznie kell az eredeti szöveg fogalmi viszonyait* – úgy, és abban a sorrendben, ahogy azok a dokumentumban is megjelennek – hogy annak eredeti, egyedi, újszerű mondanivalóját is vissza tudja

²⁷ Hjørland 2017a. Valójában a kérdés alapú megközelítés is szükségessé teszi a dokumentum Perrault által leírt mértékben történő leírását, ami lehetővé teszi a tárgy beazonosítását a találati lista böngészése során.

²⁸ Hjørland 2017a.

²⁹ IFLA 2010.

³⁰ Hogy a leírásban a dokumentum tárgyát alkotó fogalmak közül melyeket kell szerepeltetni az indextételben, az általában megegyezés kérdése. Gyakori elvárás, hogy az adott tárgy ne csak említve legyen, hanem a dokumentum témájához tartozzon, valamint az, hogy az előfordulása szignifikáns legyen (Hjørland 2017a). Például a Library of Congress szabályzata szerint azokat a fogalmakat kell feltüntetni, melyek a mű 20 százalékában szerepelnek (LOC 2008, sheet H0180).

³¹ Perrault 1974.

³² Az első példa az énekesmadarak és vándorlási szokások fogalmának együttes, egyenrangú előfordulását jelenti, tehát kezelhető posztkoordinált módon, de ez éppen azért fog zajt okozni, mert az énekesmadarak – ahol a vándorlás valójában a fogalom további szűkítését jelenti egy, a madarak osztályán belül érvényes facetta alapján – is ugyanezen a módon lesznek indexelve. (Gödert 2014)

³³ A posztkoordináció jelentheti a szintaktikai relációk kifejezését is a kereső által, de a gyakorlatban az ilyen összekapcsolás általában a Boole-féle logikai műveletek (esetleg távolsági operátorokkal kiegészített) használatával történik, és többnyire a szakirodalom is ezt érti alatta. A tisztán posztkoordináltan indexelt dokumentumok esetén a keresés során is csak az „együtt megjelenés” típusú kapcsolatok érvényesíthetők. Ezt a végleletet (amikor az összetett fogalmakat mindig csupán ismérveik puszta felsorolásával írjuk le) Mortimer Taube nevével, illetve az ő UNITERM rendszerével szokták összekötni, mivel abban került először alkalmazásra (Ungváry és Orbán 2001: 328).

adni.³⁴ Ezt a megközelítést **prekoordináció**nak is szokás nevezni, mivel a fogalmak összekapcsolása az indexelés során történik.

1.3.1 A szintaktikai relációk típusai

Ranganathan után **komplex tárgykörnek** nevezzük azokat a tárgyköröket, amelyekben „két vagy több tárgykör a közöttük fennálló kapcsolat vagy e kapcsolat adott értelmezése alapján fűződik össze egésszé.”³⁵ A kapcsolatot, amely a komplex tárgykör részei között fennáll, Ranganathan **fázis reláció**nak („phase relation”) nevezte el, majd az ilyen relációkat öt további altípusra bontotta, a kapcsolat típusa, **fázisa** alapján.³⁶

Fázis relációkkal alkotott komplex tárgykörök például az alábbiak:

- Matematika – fizikusok számára
- A fizika és a kémia közötti különbség
- Geopolitika (a földrajz befolyása a politikai viszonyokra)

Ranganathan a fázisrelációk mellett **facetta reláció**nak („facet relation”) nevezte *egy fogalom és az azt pontosító facetták közötti kapcsolatot*, az ilyen kapcsolatokkal előálló fogalmakat pedig **összetett tárgyköröknek** („compound classes”).³⁷

A **facetták** olyan jellemzők, amelyek *egy adott osztályon belül tipikusan elő szoktak fordulni*: egy macskának számos tulajdonsága lehet, például, hogy „egy festővel él” vagy „szereti megrágcsálni az elektromos kábeleket”. De vannak ezek közül olyanok, amelyek jellemzőbbek és hasznosabbak a macska objektumok leírására: ilyenek például, hogy „cirmos” vagy „sterilizált”. A facetták tehát egy osztály továbbosztásai egymástól kölcsönösen független kritériumok alapján. Minden fogalom facettákba szervezhető bizonyos attribútumok alapján, de az osztályozási rendszerek általában egy adott tudományág olyan facettáira fókuszálnak, mint *módszerek vagy eljárások*.³⁸

A fázis- és facetta relációk mellett fontos megemlíteni az **agglomerációt**, amely *„egységeket nagyobb halmazokba von össze a részek kohéziója nélkül.”* Az agglomeráció megvalósulhat a leírásra használt osztályozási rendszerben közvetlenül egymás után következő vagy nem egymás után következő fogalmak között.³⁹

³⁴ Horváth 1997.

³⁵ Ranganathan 1967: 84. Egyes szerzők megkülönböztetik a fázis relációkat a szabad kombinációktól, amelyekben a tárgykörök kapcsolatát a fázis pontos meghatározása nélkül fejezik ki (ld. Gnoli 2017). Ebben a munkában a szabad relációkat is komplex tárgykörként említem, mivel bár az ETO a fázis jelölését egészen a közelmúltig nem tette lehetővé, a kezdetektől kategórikusan megkülönbözteti a fázisrelációt és a szabad kombinációt az agglomerációtól.

³⁶ Ranganathan az alábbi fázisokat határozta meg: általános, befolyás, összehasonlítás, különbség, hatás (Ranganathan 1956). A fázisokrelációknak számos más tipizálása is kidolgozásra került más kutatók által, pl. az ETO esetében a (2003-ban bevezetett) -042 alatti általánosan közös alosztások tartalmazzák a fázisokat, amelyekkel pontosítható a kapcsolat típusa.

³⁷ Ranganathan 1956.

³⁸ Gnoli 2011.

³⁹ Neelameghan 1973; idézi Binwal 1988: 254-255.

A komplex, összetett, illetve agglomerált tárgyköröket⁴⁰ részeik és a részeik kapcsolatai együttesen definiálják. Ha a relációk kapcsolatát vizsgáljuk, azt láthatjuk, hogy a fázisrelációk összeköthetnek összetett tárgyköröket is, az agglomerált tárgykörök elemei pedig lehetnek komplex tárgykörök is. Tehát a **relációk prioritási sorrendje**: facettarelációk, fázisrelációk és agglomeráció.

A dokumentumban megvalósuló és az osztályozási jelzetben megjelenő, a használt osztályozási nyelv szavai között, annak szintaktikai szabályai alapján kifejezett eseti kapcsolatokat a szakirodalom **szintaktikai relációknak** nevezi. A fenti példáknál maradva ilyen lehet a „számára”, „különbség” vagy a „befolyása” az osztályozási rendszer adta lehetőségek között kifejezve. Látható az analógia a matematikai logika predikátumszimbólumaival, melyek egy adott igazságtartományon értelmezett n-edrendű relációk ($n \geq 1$) absztrakciói.⁴¹

1.3.2 A szintaktikai relációk jelentősége

Az **információkereső nyelvek hatékonysága** sokféle mérőszámmal mérhető, ezek közül a leggyakrabban a keresés teljességét mérő **visszahívást** és a keresés céljának való megfelelést (tehát a zaj kiküszöbölésének mértékét) mérő **precíziót** használjuk.⁴² A szintaktikai relációk alkalmazása mindkettőre hatással van.

A **kompozicionalitás elve** azt mondja ki, hogy *egy összetett kifejezés jelentését a struktúrája és a részeinek jelentése együttesen határozza meg*.⁴³ Tehát ugyanazok a fogalmak együttesen, más módon összekapcsolva más tartalmat fejeznek ki. A fentebb idézett Soergel-posztulátum azt mondja ki, hogy az indexelés során ezt figyelembe kell venni. Ennek a feltárásnak az elvárt mértéke viszont vitatott⁴⁴, elsősorban azért, mert a kapcsolatok száma (a rájuk vonatkozó egyéb információval) túl magas ahhoz, hogy valamennyit beillesszük az indexelési és visszakeresési fájlokba, ráadásul a rendszereket is fel kellene készíteni a kezelésükre.⁴⁵ Ugyanakkor a fenti példákból látható, hogy az indextelben való teljes mellőzésük a relevanciadöntéseket is megnehezíti, a keresés során való figyelmen kívül

⁴⁰ Jelen a munkában az egységesség miatt a Ranganathan-féle terminológiát használom. Fontos azonban megjegyezni, hogy a szakirodalomban a fenti kifejezések ettől különböző használata is előfordul, gyakran a facetta- és fázisrelációkkal, illetve agglomerációkkal alkotott jelzeteket egységesen komplex („complex”) vagy összetett („compound”) jelzetként vannak említve, de az is előfordul, hogy csak a fázisrelációkkal alkotott jelzeteket említik összetettként.

⁴¹ Horváth és Sütő 2001: 47. Az így alkotott fogalmak tanulmányozhatók az elsőrendű predikátumlogikán keresztül, a kalkulus vagy a halmazelmélet jelölésrendszerét eszközül véve (Frické 2017).

⁴² A visszahívást a visszakeresett releváns elemek és az összes releváns elem számának arányaként a precíziót pedig a visszakeresett releváns elemek és a visszahívott elemek számának arányaként szokás meghatározni (ld. Soergel 1994).

⁴³ Gendler-Szabó 2004.

⁴⁴ A szintaktikai relációk elmélete és alkalmazásuknak kutatása az ötvenes és a hatvanas években érte el a tetőpontját; ekkor születtek olyan rendszerek, illetve elméletek, mint a Gardin által alkotott SYNTOL, a PRECIS, illetve Farradine relációs indexelése. Ebben az időszakban az ETO relációs lehetőségeinek finomítására is történtek kísérletek, például Ranganathan (1956) és de Grolier (1962: 20-21) is fogalmazott meg ilyen jellegű javaslatokat. A legtöbbet hivatkozott ilyen kezdeményezés valószínűleg Perrault relátorrendszere (ld. Perrault 2001). Később az érdeklődés inkább a posztkoordináció alkalmazása felé fordult és a szintaktikus relációk kutatása visszább szorult. (ld. Ungváry és Orbán 2001: 194-198)

⁴⁵ Gödert 2014.

hagyásuk pedig zajt⁴⁶ okozhat (alkalmazásuk a keresés során ezzel szemben csökkenti a visszahívást, szélsőséges esetben csendet is eredményezhet, azaz azt, hogy a keresés során egyetlen releváns dokumentumot sem találunk⁴⁷).

A fogalmi kapcsolatok megfelelő feltárására szükség van a kérdés-alapú megközelítés során, és akkor is, ha a dokumentumban szereplő fogalmak mindegyike visszakereshetővé van téve. Ez ugyanis szükséges a fellelt tartalmak megfelelő megjelenítéséhez éppen úgy, mint a kifinomult relevanciadöntések meghozatalához.

Az előbb felsorolt negatív hatások a koordináció (a fogalmak összekapcsolása) „ellentétes” alkalmazásán alapultak. A kapcsolatok hiányából eredő zajt és a relevanciára vonatkozó döntések nehézkességét az váltja ki, hogy már az indexelési („pre-”) fázisban sem voltak a kellő módon ábrázolva a fogalmi kapcsolatok, mert a keresés egyszerűbbé tétele határozta meg a leírást. A prekoordinációból adódó esetleges adatvesztés pedig abból ered, hogy a keresési („poszt-”) fázisban is érvényesítve voltak. Valójában *a tiszta poszt- és prekoordináció ritka*: kifejezéseket előzetesen is össze lehet kapcsolni, ahogyan a kereséskor is elképzelhető a prekoordinált kifejezések, illetve azok részeinek összekapcsolása.⁴⁸ Az mindenesetre általánosan elmondható, hogy *a kapcsolatok nem megfelelő kezelése prekoordinált rendszerek esetén egyszerre csökkentheti a visszahívást és a precíziót*.

Fentebb említettük, hogy az indexelési és kereső rendszereket fel kell készíteni a szintaktikai kapcsolatok kezelésére. Ebből a szempontból jelentős különbség van az „a priori” (már a szövegfeldolgozás előtt is adott) és az „a posteriori” (kontextusfüggő, tehát csak az indexelendő dokumentumban fennálló) kapcsolatokat leíró jelzetek között.⁴⁹ A szakirodalomban az „a priori” relációkra a **paradigmatikus**, az „a posteriori” relációkra pedig a **szintagmatikus** jelző terjedt el.

Úgy is fogalmazhatunk, hogy *a szintagmatikus kombinációk eseti, ad hoc kapcsolatokat írnak le*. A természetes nyelvben egy mondat vagy szintagma fejezi ki őket, pl. „feminizmus és sport” vagy a „marketing hatása az adományozásra”. *A paradigmaticus kifejezések ezzel szemben tartósan fennállók, (a keresés során is) gyakran előforduló, bevett terminológiát használó kapcsolatokat* fejeznek ki, mint pl. „kultúrpolitika” vagy „vallásszociológia”.⁵⁰

Egy reláció lehet paradigmaticus az egyik dokumentációs nyelvben és szintagmatikus a másikban, vagy akár felváltva paradigmaticus vagy szintagmatikus ugyanabban a nyelvben.⁵¹ „A lényeges ismertetőjegyek egyik sajátossága, hogy mindig valamilyen konkrét vonatkozásban lényegesek, vagyis *a tárgyakkal, jelenségeknek különböző vonatkozásokban más-más ismertetőjegyei fontosak*”.⁵² Ez vonatkozik a fogalmi kapcsolatokra is: az, hogy melyik kapcsolatok tekintendők paradigmaticusnak, illetve szintagmatikusnak függhet a gyűjteménytől, az osztályozási nyelvtől vagy az azon alapuló

⁴⁶ Perrault olyan kutatási jelentést is idéz, amely szerint a pusztán a szintaktikai kapcsolatok figyelmen kívül hagyása miatti zaj mértéke a 24-60 százalékot is elérte (Perrault 2001).

⁴⁷ Az, hogy a prekoordináció csökkenti a visszahívást és növeli a precíziót, illetve, hogy az egyik mérőszám erősítésére szolgáló módszerek alkalmazása általában a másik csökkenésével jár szemléletesen is nyilvánvalónak tűnik, a hatvanas években pedig tudományos igényű kísérletekkel is alátámasztották (Horváth 1968).

⁴⁸ Horváth és Sütő 2001: 43.

⁴⁹ Horváth és Sütő 2001: 48.

⁵⁰ Vukadin 2015.

⁵¹ Hutchins 2001: 435.

⁵² Barátné 2011.

kontrollált szótártól (authority állománytól) és egyéb – objektív és szubjektív - tényezőktől. Ahogy Wersig fogalmazott: a „különbség a fogalmi relációk között nem alapvető (a belső tudati világban egyformán vannak reprezentálva), csak a valószínűségeik különböznek; csak szemléleti különbségről van szó”.⁵³

A legegyszerűbb esetben az **authority control** a keresés során *hozzáférési pontként alkalmazható terminusok meghatározását* jelenti, illetve annak biztosítását, hogy a leírt objektumhoz minden esetben a megfelelő terminus van hozzárendelve. Ennek érdekében az alkalmazott **authority fájl** rekordjai tartalmazzák a *preferált kifejezési módokat*, melyek a leírás során alkalmazandók, valamint az azokhoz tartozó gyakori *szinonimák és kereszthivatkozások* listáját. Mivel az authority fájlok általában *központilag, gyűjteményektől függetlenül vannak karban tartva*, emellett még más olyan információt is tartalmazhatnak, amely a bibliográfiai rekordokban nem található meg, de felhasználható az információkeresés során. Tárgyi authority control esetén ilyenek lehetnek a fogalmak közötti hierarchikus és mellérendelő *szemantikai kapcsolatok* vagy, osztályozási rendszer esetén az – akár több nyelvű – *természetes nyelvű leírások*, kulcsszavak.⁵⁴

A szakirodalomban – és az alkalmazás során is – konszenzus uralkodik arról, hogy fenti formájában az *authority control kielégítő megoldást nyújt a szemantikai és a paradigmatis kapcsolatok automatizálására*. Azonban a *szintagmatikus szintaktikai relációk számítógépes kezelése csak korlátozottan oldható meg* ezzel a módszerrel.⁵⁵

1.4 Az osztályozási rendszerek típusai

Az osztályozási rendszereket sok szempontból lehet csoportosítani, például a jelkészletük, a jelzetek összekapcsolásának helye és ideje (pre- illetve posztkoordinált) stb. szerint. A legelterjedtebb tipizálás azon alapul, hogy a nyelv milyen filozófiát követ a komplex fogalmak leírása során. Ez a rendszer szintjén azt tükrözi, hogy az adott nyelv milyen fogalomkészletet definiál, illetve milyen formában teszi lehetővé – és várja el – a szintaktikai kapcsolatok leírását. A fenti szempontok alapján enumeratív, facettás, illetve analitikus-szintetikus rendszereket szokás elkülöníteni.⁵⁶

1.4.1 Az enumeratív osztályozási rendszerek

Az enumeratív osztályozási rendszerek valamilyen szisztematikus *struktúrában, felsorolásszerűen tartalmazzanak valamennyi várható fogalmat*, vagy egyáltalán nem, vagy csak igen kis mértékben adva lehetőséget új jelzetek kialakítására. Ezért egy ilyen séma nem alkalmas az eseti szintaktikai kapcsolatok megjelenítésére, viszont általában nagy részletességgel jeleníti meg a szemantikai, főleg a hierarchikus kapcsolatokat; utóbbiak gyakran a jelzetekben is tükröződnek, pl. a tizedes továbbosztás elvén. Az ilyen rendszerek előnye, hogy *könnyű tájékozódni és navigálni az osztályok között*, és tudástérképek kialakítására is nagyszerűen alkalmasak. Hátrányuk, hogy az *eseti kapcsolatok leírására tulajdonképpen alkalmatlanok*, továbbá az, hogy tudomány- és tudásterületek változásai nagyon *sűrű revíziót* tehetnek szükségessé. Az ilyen revízió magas tudományos

⁵³ Wersig leírása szerint egyik esetben a nyelvhasználat által is szentesített vagy egy adott vonatkozásból, szubjektíve adott, a másik esetben valamilyen művelet eredményeképpen, a konvencionizált kapcsolat valószínűsége nélkül keletkezett relációkról van szó (Wersig 2001: 381-382)

⁵⁴ O'Neill, Bennett és Kammerer 2014.

⁵⁵ Tartaglia 2004.

⁵⁶ Broughton 2015: 33-40.

felkészültséget igényel, és nagy körültekintést is, hiszen a már meglévő katalógusokra is folyamatosan tekintettel kell lenni. Másik gyakran említett probléma a hierarchikus enumerikus rendszerekkel kapcsolatban, hogy *egy fogalom a hierarchia több helyén is megjelenhet*, ezért a hatékony kereséshez elméletben minden lehetséges megjelenési helyét tekintetbe kellene venni a minél teljesebb visszahívás érdekében.⁵⁷ Ugyanakkor a hierarchikus enumerikus rendszerek nagy előnye, hogy *kiváló eszközt adnak az inkluzív kereséshez*, legalábbis a saját (mono-)hierarchiájuk keretein belül.⁵⁸ A legelterjedtebb enumerikus osztályozási rendszerek a **Dewey Tizedes Osztályozás (DDC)**⁵⁹ és a **Library of Congress osztályozási rendszere (LCC)**.

1.4.2 A facettás osztályozási rendszerek

Bár a facettás osztályozás alapkonceptiója egészen a XVII. századig, Cyprian Kinner, George Dalgarno és John Wilkins munkásságáig nyúlik vissza⁶⁰, és egyes elemeit Paul Otlet az Egyetemes Tizedes Osztályozás legelső kiadásában a gyakorlatban is megvalósította⁶¹, ennek az osztályozási típusnak a megjelenését elsősorban **Ranganathan**hoz szokás kötni, mivel ő dolgozta ki a facettaanalízis elméleti hátterét és ő alkotta meg az első teljes egészében ezen alapuló általános osztályozási rendszert.⁶²

Ranganathan osztályozási elméletében – ahogy azt az általa alkotott osztályozásdefiníciókban is láthattuk - abból indult ki, hogy *lehetetlen táblázatokban feltüntetni minden szakjelzetet és fogalmat*, melyekre szükség lehet a dokumentumok leírásához, mivel nem lehet előre meghatározni a tudományok, ismeretek, illetve jelenségek és tárgyak sokoldalú és fejlődő kapcsolatait.⁶³ Ezért olyan osztályozási rendszert állított össze, melyben *csak kis számú osztályt definiált*, amelyek a természettudományok és technika, a humán- és a társadalomtudományok főosztályai alá besorolva tartalmazzák az általa alaptárgyaknak nevezett fogalmakat. Ezek azok a fogalmak, amelyeknek nem eleme egyetlen izolátum sem. Izolátumnak Ranganathan azokat a fogalmakat nevezte, melyek összetevői lehetnek valamely összetett tárgykörnek, de önmagukban nem alkothatnak tárgykört. Az *összetett tárgyak pedig az alaptárgyból és egy vagy több izolátumból álló tárgyak*.⁶⁴ Minden alaptárgy esetén meg vannak határozva a használható izolátumok, mégpedig öt alapvető szempont⁶⁵ alapján (más szavakkal az izolátumoknak alapvetően öt fajtája van). Az egyes osztályok esetén tehát pontosan meg van adva, milyen izolátumokat lehet és kell használni, és minden esetben a szempontrendszerből levezetett, jól definiált szabályok vonatkoznak a sorrendre, illetve a felhasználandó formulára. Például az irodalom osztály (O) csak egy facettát (P) tartalmaz, négy különböző formában: nyelv (P1), forma

⁵⁷ Ezt a jelenséget szokták párhuzamos hierarchiának vagy tudományágak közötti („cross-discipline”) kapcsolatnak is nevezni (Gnoli, De Santis és Pusterla 2015).

⁵⁸ Inkluzív keresésről akkor beszélünk, amikor egy kulcsszóra rákeresve a keresőmotor nem csak az azzal, hanem az abból hierarchikus leszármaztatott kulcsszavakkal indexelt szövegeket is visszakeresi. (Soergel 1994)

⁵⁹ A Dewey Tizedes Osztályozásra (Dewey Decimal Classification) magyarul általában a tizedes osztályozás rövidítésével, TO-ként szoktak hivatkozni. Itt az egyértelműség kedvéért mindenütt a világszerte elterjedt angol rövidítést (DDC) használjuk.

⁶⁰ Schulte-Albert 2009.

⁶¹ vö. Otlet 1990.

⁶² Raghavan 2015.

⁶³ Babiczky és Barátné 1998: 113.

⁶⁴ Ranganathan 1967: 83-84.

⁶⁵ Személyiség, anyag, energia, idő és tér, melyekre együttesen az angol rövidítésük alapján PMETS-ként szokás hivatkozni.

(P2), szerző (P3) és mű (P4), ezek felsorolásával kell meghatározni az adott művet. A facettákat az elkészült jelzetben kettőspont választja el egymástól, az osztályozási rendszer a nevét⁶⁶ is erről kapta.⁶⁷

Egy adott tudásterület facettáinak a meghatározását, illetve egy új osztályozás (vagy más tudásszervező rendszer) létrehozását nevezzük **facettaanalízisnek**.⁶⁸ A fentiek alapján az analízis során meghatározott facetták egyfelől a saját osztályukon belül határoznak meg egy továbbosztást, másfelől kapcsolódnak egy általános kategóriához. Az előző példában a „szerző” az irodalom osztály egy facettája, amely a „személyiség” kategóriához kapcsolódik. A kategóriák meghatározása rendszerenként változik, Ranganathan esetében az öt említett szempontot tartalmazza.⁶⁹

A fentiekből jól láthatók a *facettás osztályozási rendszerek fő jellemzői*: a szigorú *strukturáltság*; hogy *csak egyszerű fogalmakat sorolnak fel*; hogy amikor és ahogyan szükséges, *új jelzetek építését* tesz szükségessé. Ezek a rendszerek sokkal nagyobb teret engednek a jelzetek kombinálásának a klasszikus enumerikus osztályozási rendszerekénél.

A **kettőspontos osztályozás (Colon Classification, CC)** volt az első tudományos alapon megtervezett facettás osztályozás, és mint ilyen, nagy hatással volt valamennyi később létrejött vagy felújított rendszerre.⁷⁰ Maga az elnevezés is Ranganathan-tól származik, aki az összetett tárgykörök összetevőit jelölő fogalmak (alaptárgyak és az izolátumok) összefoglaló nevéként alkalmazta a „facetta” szót.⁷¹ Az 1950-es évek elején megalakuló *Classification Research Group* is a facetta analízisben találta meg azt az elméleti hátteret, amely meglátásuk szerint lehetővé teszi egy „minden információkeresés alapjaként” szolgáló osztályozási rendszer kialakítását. A csoport egyik fő célkitűzése egy ilyen, Ranganathan elméleti munkásságán alapuló, de a brit kultúrához jobban illeszkedő facettás rendszer kialakítása volt. Ennek érdekében nagyon jelentős elméleti tevékenységet végeztek, de a lefektetett elveket nem sikerült maradéktalanul átültetniük a gyakorlatba. Ehhez legközelebb a Henry Evelyn Bliss osztályozási rendszerének Jack Mills és Vanda Broughton által továbbfejlesztett verziója (**Bliss-Classification 2, BC2**) áll, amely a *kettőspontos osztályozás mellett máig az egyetlen általános facettás osztályozási rendszer*.⁷²

⁶⁶ „Kettőspontos osztályozás”, angolul „Colon Classification” (röviden CC)

⁶⁷ Iyer 2012: 109-112.

⁶⁸ A facettaanalízis lépései: (1.) a feldolgozott tudásterület meghatározása, (2.) a szóba jöhető terminusok meghatározása a terület irodalmának és kommunikációjának vizsgálata alapján, (3.) facetták strukturálása (általában hierarchikus elrendezés szerint), (4.) fogalmi leírások („scope notes”) elkészítése, (5.) facetták elrendezése a későbbi felhasználáshoz igazodva (pl. poszt- vagy prekoordinált használatot feltételezve), (6.) jelölések hozzáadása a facettákhoz (Hjørland 2013).

⁶⁹ Gnoli 2009.

⁷⁰ A gyakorlatban viszont nem tudott elterjedni Indián kívül (ahol az első számú osztályozási rendszernek számít könyvtárakban, bár bizonyos könyvtártípusok a DDC-t és az ETO-t is használják, ld. Singh 2008). A kettőspontos osztályozás hetedik kiadását még Ranganathan készítette elő, de csak jóval a halála után jelent meg egy – nem teljes – kiadása, mely hiányosságai miatt sem az oktatásban sem a gyakorlatban nem terjedt el, még Indiában sem (Raghavan 2015).

⁷¹ Ranganathan 1967: 88.

⁷² Broughton 2017, BCA 2015.

A jelenlegi tendencia szerint a facettaanalízisre egyre inkább az osztályozás egyetlen lehetséges megközelítési módjaként szokás tekinteni, figyelmen kívül hagyva az egyéb megközelítéseket.⁷³ Jellemző a meglévő osztályozási és más tudásszervezési rendszerek facettás formára való átalakítása.⁷⁴ A módszer központi témája a tudományos diskurzusnak: az osztályozás és információkeresés területén és azon kívül is számos kutatás és publikáció vizsgálja annak elméleti kérdéseit és azok megjelenését a megvalósított rendszerekben. A téma aktualitását mutatja az is, hogy a 2017. évi ETO Szeminárium témája is a facettás osztályozás jelenlegi helyzete volt.⁷⁵

A facetta listát és a facetták felhasználására vonatkozó szabályokat definiáló hagyományos facettás rendszerek mellett, azok továbbfejlesztéseként kialakult az úgynevezett **szabad facettás** („freely faceted”) modell, mely szerint elég csak követelményeket és elveket definiálni a facetták meghatározására. Ez a facettás rendszerek merev formulái helyett jóval *rugalmasabb jelzetszerkesztést* enged meg. Az ilyen rendszerek létrejöttét a számítógépesítés tette lehetővé, és részben a számítógépesítés, részben az interdiszciplinaritás elterjedése indokolja.⁷⁶ A legismertebb szabad facettás rendszerek *Ranganathan kettőspontos osztályozásának hatodik kiadása (CC6)*⁷⁷, illetve *Derek Austin* osztályozása, melynek alapelveit később a *PRECIS*-ben fejlesztette tovább.⁷⁸ A közelmúltban megalkotott rendszerek közül ezt a szemléletet követi a *PRECIS*-hez hasonló szerkezetű *Integrative Level Classification (ILC)* is.⁷⁹

1.4.3 Az analitikus-szintetikus osztályozási rendszerek

Az analitikus-szintetikus⁸⁰ osztályozási rendszerek *félúton vannak az enumerikus és a facettás rendszerek között*.⁸¹ Általában rendelkeznek egy fő táblázati résszel, illetve az ezt kiegészítő segéd táblázatokkal és lehetőséget adnak az osztályok kombinálására. Ezért a belőlük képzett jelzetek egyszerű táblázati számok és komplex, prekoordinált kifejezések egyaránt lehetnek. Ezeknek a rendszereknek az alapkonceptiója a leírandó tartalmak részekre bontása, a részek leírása táblázati számokkal és facettákkal, majd a végleges jelzet kialakítása az így kapott részjelzetek

⁷³ Hjørland 2017b.

⁷⁴ Ahogy a későbbi fejezetekben látjuk majd, az ETO átszervezéseinek is ez az egyik fő iránya.

⁷⁵ Piros 2017b, „International UDC Seminar 2017”.

⁷⁶ Iyer 2012: 112, Gnoli et. al. 2011.

⁷⁷ Satija 2017.

⁷⁸ Gnoli 2017.

⁷⁹ Gnoli et. al. 2011., Gnoli 2017.

⁸⁰ Az „analitikus-szintetikus” kifejezés is Ranganathantól származik (Ungváry és Orbán 2001: 75). A magyar szakirodalomban néha előfordul az „elemző-összerakó” osztályozás kifejezés is.

⁸¹ Bár az általánosan elfogadott, hogy az ETO félúton helyezkedik el az enumerikus és a tisztán facettás rendszerek között, illetve tartalmaz olyan elemeket is, amelyek azokban nem fordulnak elő, a szakirodalom ezt nem minden esetben jelöli önálló típusként. Jelen munkában az idézett „klasszikus” tagolást használom, mert ez világít rá legjobban az ETO-nak azokra a sajátosságaira, amelyek megkülönböztetik az enumerikus és a teljesen facettás rendszerektől. De meg kell jegyezni, hogy az „analitikus-szintetikus” kifejezést (mely a tárgyak részekre bontását, majd az azoknak megfelelő fogalmak szintetizálását, összeépítését jelenti a megfelelő szintaktikai relációk segítségével) az információkereső szakirodalom gyakran egységesen használja az összes, ilyen típusú jelzetalkotást lehetővé tévő rendszerre, a „facettás” kifejezést azok számára tartva fenn, melyek olyan elméleti posztulátumok alapján épülnek fel, mint a Ranganathan vagy a CRG által megfogalmazott szempontrendszerek (Slavić és Cordeiro 2004).

összekapcsolásával.⁸² Így az ilyen rendszerek az enumerikus rendszerekénél *rövidebb táblázatok* összeállítását igénylik, miközben *jóval nagyobb rugalmasságot* tesznek lehetővé a szintaktikai kapcsolatok leírásában, ami *növelheti a precíziót* (tehát csökkentheti a zajt a keresés során) és megkönnyíti a helyes relevanciadöntéseket minél jobban támogató jelzetek elkészítését. Hátrányuk, hogy a komplex szabályrendszer (és a még így is elég nagy táblázatok) alkalmazása *komoly szakmai felkészültséget igényel*, illetve, hogy *a számítógépes felhasználásuk nehézségekbe ütközik*. Egy másik probléma velük kapcsolatban, hogy *egy összetett tárgykör több különböző módon is kifejezhető*, amittől csökken a reprezentációs kiszámíthatóság, ez pedig negatív hatással van a visszahívás mértékére. Az analitikus-szintetikus rendszerek legismertebb képviselői az **Egyetemes Tizedes Osztályozás** mellett Henry Evelyn Bliss 1940-as években megjelentetett osztályozási rendszere (**Bliss Classification, BC**)⁸³ és Ranganathan kettőspontos osztályozásának negyedik, ötödik és hatodik kiadása (**CC 4-6**).⁸⁴

A fentiek alapján látható, hogy az enumerikus rendszerek esetén a leírás során a rendszeralkotó fogalmak felől kell közelítenünk, azok közül kell kiválasztani a megfelelőt. Ez a kötött tárgyszavas indexeléssel rokonítja az ilyen rendszereket. Az ilyen osztályozást „**top-down**” osztályozásnak is szokták nevezni, mivel ilyenkor a fogalmak továbbosztása képezi az osztályozási tevékenység alapját. Az analitikus-szintetikus (és facettás) szemlélet ennek éppen az ellenkezője. Ilyenkor a dokumentum tartalma alapján épített új jelzetekkel bővül a rendszer fogalmi struktúrája, ezért ebben „**bottom-up**” osztályozásról szokás beszélni.⁸⁵

1.5 Irodalomjegyzék

Babiczyk Béla, és Barátné Hajdu Ágnes. 1998. *Bevezetés az információkereső nyelvek elméletébe és gyakorlatába*. Budapest: Universitas.

Barátné Hajdu Ágnes. 2011. „Az ismeretszervezés dinamizmusa, különös tekintettel az Egyetemes Tizedes Osztályozásra.” *Könyv és nevelés* 13 (1): 63-72.

Binwal, J. C. 1988. „Modes of formation of subjects and their role in information retrieval.” PhD dissz., Darwad: Karnatak University. <http://shodhganga.inflibnet.ac.in/handle/10603/94558>. (megtekintve: 2017-11-01)

Bliss Classification Association (BCA). 2015. „The Bliss Bibliographic Classification : history & description.” <http://www.blissclassification.org.uk/bchist.shtml>. (megtekintve: 2017-11-01)

Bliss, Henry Evelyn. 1929. *The organization of knowledge and the system of the sciences*. New York, NY: Henry Holt and Company.

Broughton, Vanda. 2015. *Essential Classification. Second Edition*. London: Facet Publishing.

Dahlberg, Ingetraut. 2001. „Az ismeretek egyetemes rendjének alapjai: Az egyetemes osztályozási rendszer kérdései és lehetőségei.” In *Osztályozás és információkeresés. Kommentált szöveggyűjtemény. Első kötet: Az osztályozás elmélete*, szerk. Ungváry Rudolf, és Orbán Éva, 221-260. Budapest: Országos Széchényi Könyvtár.

⁸² Gulati 2013: 150-151.

⁸³ BCA 2015, Broughton 2015: 36-37.

⁸⁴ Raghavan 2015.

⁸⁵ Broughton 2015: 33-40.

- de Grolier, Eric. 1962. *A study of general categories applicable to classification and coding in documentation*. Paris: UNESCO.
- Frické, Martin H. 2016. „Logical division.” *Knowledge Organization* 43 (7): 539-549.
- Frické, Martin H. 2017. „Faceted classification, analysis and search: some questions on their interrelations.” In *Faceted classification today: theory, technology and end users: proceedings of the International UDC Seminar 2017, London (UK), 14-15 September*, edited by Aida Slavić, and Claudio Gnoli, 25-41. Würzburg: Ergon Verlag.
- Gendler-Szabó Zoltán. 2004. „Compositionality.” In *Stanford Encyclopedia of Philosophy*. Stanford, CA: Stanford University. <https://plato.stanford.edu/entries/compositionality>. (megtekintve: 2017-11-01)
- Gnoli, Claudio. 2009. „The UDC Philosophy revision: first report.” *Extensions and Corrections to the UDC* 31: 25-31. <http://hdl.handle.net/10150/200633>. (megtekintve: 2017-11-01)
- Gnoli, Claudio. 2011. „Facets in UDC: a review of current situation.” *Extensions and Corrections to the UDC* 33: 19-36. <http://arizona.openrepository.com/arizona/handle/10150/236511>. (megtekintve: 2017-11-01)
- Gnoli, Claudio, Rodrigo De Santis és Laura Pusterla. 2015. „Commerce, see also Rhetoric: cross-discipline relationships as authority data for enhanced retrieval.” In *Classification & Authority Control: Expanding Resource Discovery : Proceedings of the International UDC Seminar 2015, 29-30 October 2015, Lisbon, Portugal*, szerk. Aida Slavić, és Maria Inês Cordeiro, 151-162. Würzburg: Ergon Verlag.
- Gnoli, Claudio. 2017. „Syntax of facets and sources of foci: a review of alternatives.” In *Faceted classification today: theory, technology and end users: proceedings of the International UDC Seminar 2017, London (UK), 14-15 September*, edited by Aida Slavić, and Claudio Gnoli, 243-256. Würzburg: Ergon Verlag.
- Gnoli, Claudio. 2018. „Notation.” *Encyclopedia of Knowledge Organization*. ISKO. <http://www.isko.org/cyclo/notation>. (megtekintve: 2018-03-10)
- Gödert, Winfried. 2014. „Facets and typed relations as tools for reasoning processes in information retrieval.” In *Metadata and Semantics Research: 8th Research Conference, MTSR 2014, Karlsruhe, Germany, November 27-29, 2014. Proceedings*, szerk. Sissi Closs, Rudi Studer, Emmanouel Garoufallou, és Miguel Angel Sicilia, 128-140. Cham: Springer International Publishing. <http://arxiv.org/abs/1409.0491>. (megtekintve: 2017-11-01)
- Gulati, Dipti. 2013. „Use of library classification schemes in the ICT environment in selected libraries in national capital region: a study.” PhD dissz., Delhi: University of Delhi. <http://shodhganga.inflibnet.ac.in:8080/jspui/handle/10603/31769>. (megtekintve: 2017-11-01)
- Hjørland, Birger. 2013. „Facet analysis: The logical approach to knowledge organization.” *Information Processing and Management: an International Journal* 49: 545–557.
- Hjørland, Birger. 2017a. „Subject (of Documents).” *Knowledge Organization* 44 (1): 55-64. <http://www.isko.org/cyclo/subject>. (megtekintve: 2017-01-01)
- Hjørland, Birger. 2017b. „Facet analysis as one among other theories of classification.” Felkért előadás. UDC-seminar London, September 14.-15., 2017. http://seminar.udcc.org/2017/files/BHjorland_UDCSeminar2017_slides.pdf. (megtekintve: 2017-01-01)

- Hutchins, Williams John. 2001. „Indexelő és osztályozó nyelvek. Nyelvészeti tanulmány struktúrákról és funkciókról.” In *Osztályozás és információkeresés. Kommentált szöveggyűjtemény. Első kötet: Az osztályozás elmélete*, szerk. Ungváry Rudolf, és Orbán Éva, 388-469. Budapest: Országos Széchényi Könyvtár.
- International Federation of Library Associations (IFLA). 2010. „Functional Requirements for Subject Authority Data (FRSAD): Final Report.” <https://www.ifla.org/node/5849>. (megtekintve: 2017-11-01)
- „International UDC Seminar 2017.” UDCC (UDC Consortium). <http://seminar.udcc.org/2017/>. (megtekintve: 2017-11-01)
- Iyer, Hemalata. 2012. *Classificatory Structures : Concepts, Relations and Representation*. Würzburg: Ergon Verlag.
- Horváth Tibor (1968). „A második cranfieldi jelentés.” *Könyvtári Figyelő* 14 (5): 351-369.
- Horváth Tibor. 1997. „Don Quijoték útjain.” *Könyvtári Figyelő* 44 (4): 665-674. <http://ki.oszk.hu/kf/kfarchiv/1997/4/horvath.html>. (megtekintve: 2017-11-01)
- Horváth Tibor, és Sütő Péter. 2001. „A tartalmi feltárás.” In *Könyvtárosok kézikönyve. 2. kötet. Feltárás és visszakeresés*, szerk. Horváth Tibor, és Papp István, 35-186. Budapest: Osiris Kiadó, 2001.
- Kleineberg, Michael. 2017. „Integrative levels.” *Knowledge Organization* 44 (5): 349-379.
- Library of Congress (LOC). 2008. *The subject headings manual*. Washington, D.C: Library of Congress, Policy and Standards Division. <https://www.loc.gov/aba/publications/FreeSHM/freeshtm.html>. (megtekintve: 2017-11-01)
- Neelameghan Arashanipalai. 1973. „Agglomerate basic subjects.” *Library Science with a Slant to Documentation* 10 (2): 202-206.
- Ogden, C. K., and I. A. Richards. 1946. *The meaning of meaning: a study of the influence of language upon thought and of the science of symbolism*. New York: Harcourt, Brace & World. Inc.
- O'Neill, Edward T., Rick Bennett, és Kerre Kammerer. 2014. „Using Authorities to Improve Subject Searches.” *Cataloguing & Classification Quarterly* 52 (1): 6-19.
- Otlet, Paul. 1990. „On the Structure of Classification Numbers”. In *International Organization and Dissemination of Knowledge: Selected Essays of Paul Otlet*, ford. és szerk. W. Boyd Rayward, 51-61. Amsterdam: Elsevier, 1990.
- Pálvölgyi Mihály. 1998. *Információfeldolgozás, információkereső nyelvek*. Szombathely: Berzsenyi Dániel Tanárképző Főiskola.
- Perrault, Jean Martin. 1974. „Az ETO alapelveiről.” [Tömörítette Horváth Magda] *Könyvtári Figyelő* 20 (2): 126-131.
- Perrault, Jean Martin. 2001. „A kategóriák és relátorok új rendszere.” In *Osztályozás és információkeresés. Kommentált szöveggyűjtemény. Első kötet: Az osztályozás elmélete*, szerk. Ungváry Rudolf, és Orbán Éva, 200-217. Budapest: Országos Széchényi Könyvtár.
- Piros Attila. 2017a. „The thought behind the symbol: about the automatic interpretation and representation of UDC numbers.” In *Faceted classification today: theory, technology and end users: proceedings of the International UDC Seminar 2017, London (UK), 14-15 September*, edited by Aida Slavić, and Claudio Gnoli, 203-218. Würzburg: Ergon Verlag.

- Piros Attila. 2017b. „A facettás osztályozás napjainkban: elmélet, technológia és a végfelhasználók.” *Könyv, könyvtár, könyvtáros* 26 (11): 22-30.
- Raghavan, K. S. 2015. „The Colon Classification: A few considerations on its future.” *Annals of Library and Information Studies* 62 (4): 231-238. <http://nopr.niscair.res.in/handle/123456789/33717>. (megtekintve: 2017-11-01)
- Ranganathan, Siyali Ramamrita. 1956. „Depth classification 21 phase and intra-facet relations.” *Annals of Library and Information Studies (ALIS)* 3 (4): 109-116. <http://nopr.niscair.res.in/handle/123456789/28577>. (megtekintve: 2017-11-01)
- Ranganathan, S. R. 1967. *Prolegomena to library classification. 3rd ed.* New York: Asia Publishing House. <http://hdl.handle.net/10150/106370>. (megtekintve: 2017-11-01)
- Satija, Mohinder P. 2017. „Colon Classification (CC).” *Knowledge Organization* 44 (4): 291-307. http://www.isko.org/cyclo/colon_classification. (megtekintve: 2017-11-01)
- Singh, K. P. 2008. „UDC and its use: a case study of libraries and information centres in Delhi.” *Extensions & Corrections to the UDC* 30: 59-66. <http://arizona.openrepository.com/arizona/handle/10150/110843>. (megtekintve: 2017-11-01)
- Schulte-Albert, Hans G. 2009. „Cyprian Kinner and the Idea of a Faceted Classification.” *Libri* 24 (4): 324-337.
- Slavić, Aida, és Maria Inês Cordeiro. 2004. "Core requirements for automation of analytico-synthetic classifications." In *Knowledge Organization and the Global Information Society: Proceedings of the Eighth International ISKO Conference 13-16 July 2004, London, UK*, szerk. I. C. McIlwaine, 187-92. <http://hdl.handle.net/10150/105814>. (megtekintve: 2017-11-01)
- Soergel, Dagobert. 1994. „Indexing and retrieval performance: The logical evidence.” *Journal of the American Society for Information Science* 45 (8): 589-599. <http://www.dsoergel.com/UBLIS571DS-13.2-1Reading3SoergelIndexingandRetrievalPerformance.pdf>. (megtekintve: 2017-11-01)
- Tartaglia, S. 2004. „Authority Control and Subject Indexing Languages.” *Cataloging & Classification Quarterly* 39 (1/2): 365-377.
- Ungváry Rudolf. 2000. „Az ETO szükségessége.” *Iskolakultúra* 10 (4): 27-31. http://real.mtak.hu/61411/1/EPA00011_iskolakultura_2000_04_027-031.pdf. (megtekintve: 2017-11-01)
- Ungváry Rudolf, és Orbán Éva. 2001. *Osztályozás és információkeresés. Kommentált szöveggyűjtemény. Első kötet: Az osztályozás elmélete*. Budapest: Országos Széchényi Könyvtár. <http://mek.oszk.hu/01600/01683/pdf/01683-1.pdf>. (megtekintve: 2017-11-01)
- Vukadin, Ana. 2015. „Development of a classification-oriented authority control: the experience of the National and University Library in Zagreb.” In *Classification & Authority Control: Expanding Resource Discovery: Proceedings of the International UDC Seminar 2015, 29-30 October 2015, Lisbon, Portugal*, szerk. Aida Slavić, és Maria Inês Cordeiro, 111-120. Würzburg: Ergon Verlag.
- Wersig, Gernot. 2001. „Információ, kommunikáció, dokumentáció. Hozzájárulás a fogalmak tisztázásához az információ és dokumentációtudományban.” In *Osztályozás és információkeresés. Kommentált szöveggyűjtemény. Első kötet: Az osztályozás elmélete*, szerk. Ungváry Rudolf, és Orbán Éva, 358-386. Budapest: Országos Széchényi Könyvtár.

2 Az Egyetemes Tizedes Osztályozás

„Az ETO az ismeretek valamennyi területére kiterjedő osztályozási séma, kifinomult indexelő és visszakereső eszköz.”⁸⁶ (UDC Consortium)

2.1 Az Egyetemes Tizedes Osztályozás története és jelentősége

Az Egyetemes Tizedes Osztályozás⁸⁷ (ETO) egyike (az LCC⁸⁸ és a DDC⁸⁹ mellett) a három „nagy” klasszikus osztályozási rendszernek, amelyek a XIX. század végén keletkeztek, és világszerte ma is a legismertebb és legelterjedtebb könyvtári tartalmi feltáró rendszereknek számítanak. A másik két rendszertől eltérően azonban az ETO már eredetileg sem könyvtári, hanem határozottan dokumentációs célokra jött létre.⁹⁰

Az ETO-t 1895-ben Brüsszelben alkotta meg a két belga internacionalista jogász, bibliográfus és politikus, **Paul Otlet**⁹¹ és **Henri La Fontaine**.⁹² Nagyszabású víziójuk egy, a világ összes nyomtatott dokumentumának számbavételére létrehozott cédulakatalógus, az *Egyetemes Bibliográfiai Repertórium*, és az azon alapuló nemzetközi tájékoztató szolgálat megalkotása volt.⁹³ Tervükhöz elnyerték a belga kormányzat támogatását is, így megszervezhették az első bibliográfiai kongresszust 1895-ben, melyen megalakították a Nemzetközi Bibliográfiai Intézetet (International Institute of Bibliography, IIB⁹⁴). A kongresszuson jelentették be egy új, a Repertórium céljaira a *Dewey Tizedes Osztályozás (DDC) alapján létrehozott osztályozási rendszer* tervét.⁹⁵ A világkatalógusban való

⁸⁶ „UDCC”.

⁸⁷ „About Universal Decimal Classification (UDC)”.

⁸⁸ „Library of Congress Classification”.

⁸⁹ „Dewey Decimal Classification summaries”.

⁹⁰ Satija 2008. Az Amerikai Könyvtárak Egyesületének megfogalmazása szerint „az ETO nem könyvtári rendszer, inkább egy könyvtárak számára fenntartott rendszer”, amely elsősorban a fogalmak leírására, és nem dokumentumok rendezésére való (San Segundo 2007).

⁹¹ Paul Otlet (1868-1944) jogász, internacionalista és békeaktivista, bibliográfus, „az információtudomány atyja”, az egyetemes nemzetközi bibliográfiai számbavétel és tájékoztatás vezéralakja. A mikrofotográfia dokumentációs célú alkalmazásának úttörője (Goldschmidt és Otlet 1990a), amelyet a Nemzetközi Dokumentációs Iroda (International Offices of Documentation, OID) égisze alatt létrejövő Egyetemes Dokumentációs Hálózat technikai alapjának tartott (Goldschmidt és Otlet 1990b), amellyel az általánosan elterjedt nézet szerint lefektette a hipertext elméleti alapjait, még ha az eltérő filozófiai alapok miatt ez elsősorban formai hasonlóságot jelent is (Rayward 1994b). Életrajza, a munkásságáról készült dokumentumfilm és tanulmányok elérhetők életrajzírójának, W. Boyd Rawyardnak a honlapján („Rayward’s Otlet Page: Paul Otlet and Documentation”).

⁹² Henri La Fontaine (1854-1943) internacionalista és pacifista politikus, a Nemzetközi Békeiroda elnökeként (1907-1943) az 1913. év Nobel-békedíjasa, a nemzetközi jog professzora, a belga szenátus képviselője (1895-1932), alpinista és bibliográfus („The Nobel Peace Prize 1913”).

⁹³ A világbibliográfiára való törekvés eszméje legkorábban a XVI-ik századig, Johann Trithem és Konrad Gesner munkásságáig nyúlik vissza. Otlet és La Fontaine, bár eszközeiket a könyvtárak kialakulófélben lévő módszertanára (cédulakatalógus, osztályozási rendszerek) alapozták, tevékenységüket tudatosan elkülönítették azokétól, amit annak elnevezésével („dokumentáció”) is igyekeztek kihangsúlyozni (Shera 1970).

⁹⁴ Szintén La Fontaine és Otlet nevéhez fűződik az IIB-hez hasonló nemzetközi szervezetek munkájának összehangolására létrejött és ma is aktív Nemzetközi Szervezetek Uniójának 1907. évi megalapítása is („UIA’s History”).

⁹⁵ Rayward 1994a.

felhasználhatóság érdekében az osztályozást francia nyelvre fordították, valamint számos módosítást és kiterjesztést eszközöltek az eredeti rendszerhez képest⁹⁶, ezzel létrehozva a világ első analitikus-szintetikus és facettás osztályozási rendszerét.⁹⁷ Melvil Dewey, akit az IIB első alelnökének is megválasztottak, támogatta a módosításokat, így a DDC és az ETO fejlesztése összhangban folyt egészen a múlt század harmincas éveig.⁹⁸ A Repertórium építését az 1930-as években leállították⁹⁹, de az időközben könyvtári osztályozásként is egyre népszerűbb ETO annak megszűnése után, önálló rendszerként is tovább tudott fejlődni.¹⁰⁰

Az ETO története szorosan összefonódik a Nemzetközi Dokumentációs Szövetségével. A Nemzetközi Bibliográfiai Intézet az 1920-as években átalakult természetes személyek szövetségéből nemzeti intézmények szövetségévé, és felvette előbb a Nemzetközi Dokumentációs Intézet (IID), majd 1937-ben a Nemzetközi Dokumentációs Szövetség (FID) nevet.¹⁰¹ A FID gondozta az ETO nemzeti kiadásainak alapját képező adatbázist csaknem száz évig. A szervezett és rendszeres revíziós munka a második világháborút követően, a harmadik, német nyelvű kiadás (1934-1953) megjelenésével párhuzamosan alakult ki.¹⁰² A FID 1949 óta évente adta közre a táblázatokban eszközölt legfrissebb bővítéseket és változásokat tartalmazó „**Extensions and Corrections to the UDC**” című kiadványát.

Ezt követően számos nyelven jelentek meg teljes, közép-, rövidített vagy szakterületi ETO kiadások a nyolcvanas évekig.¹⁰³

Az 1960-as évektől a tudományok változásai, az osztályozástudomány fejlődése a korábbi bővítések mellett szükségessé tették a táblázatok szerkezetének megváltoztatását is¹⁰⁴, ami maga után vonta a FID-en belüli ETO szakterületi és nemzeti revíziós bizottságok, illetve az azokat koordináló Központi

⁹⁶ OIB 1970 és Otlet 1970.

⁹⁷ Az ETO első – francia nyelvű – teljes kiadása 1905-1907 között jelent meg a Repertóriumhoz készült segédanyag részeként („UDC Editions”), tehát több, mint huszonöt évvel előzte meg Bliss és Ranganathan rendszereit (Iyer 2012: 108-109.)

⁹⁸ McIlwaine 1998.

⁹⁹ A Nemzetközi Bibliográfiai Repertórium építése 1934-ben szűnt meg, amikor a belga kormány bezáratta a gondozásáért felelős nemzetközi központot, a Mundaneumot. Ekkor a bibliográfia csaknem 16 millió rekordot tartalmazott. (Rayward 1997). A Repertórium 2013-ban felkerült az UNESCO Világ emlékezete listájára is („Memory of the World. Inscriptions of the documentary heritage in 2013.”). A megmaradt katalógusszekrények egy része 1998 óta a Mundaneum múzeumban van kiállítva, a belgiumi Mons-ban („Universal Bibliographic Repertory”).

¹⁰⁰ A második teljes kiadás 1927-1933, a harmadik – német nyelvű – pedig 1934-1953 között jelent meg („UDC Editions”).

¹⁰¹ Pogányné 2007.

¹⁰² McIlwaine 1998.

¹⁰³ A teljes kiadás kb. 200.000 osztályt tartalmazott. A középkiadások közé a 20-100.000 osztályt, a rövidített közé a húszezernél kevesebb osztályt tartalmazó kiadásokat szokták sorolni. Ezen kívül létezhetnek szintén teljes, rövidített, közepes vagy akár kiterjesztett szakterületi kiadások, amelyek egy meghatározott felhasználói kör vagy szakterület számára készülnek (Slavić 2004). 1993 óta a hivatalos terminológia a korábbi teljes és közép- kiadások helyett a „kiterjesztett” és „standard” szavakat használja (Slavić-Overfield 2005: 221).

¹⁰⁴ Korábban a DDC-vel való összhangra való törekvés is újtában állt az alapvető változtatásoknak. A közös alapot a FID 1962-es döntése szüntette meg végérvényesen (Dahlberg 1972).

Osztályozási Bizottság (Central Classification Committee, FID/CCC) együttműködésén alapuló struktúra kialakítását.¹⁰⁵

Ebben az időben az információs szakemberek éppen kiábrándulóban voltak a klasszikus könyvtári osztályozási rendszerekből, elsősorban azok rugalmatlanságai és az általuk használt tudományfelosztások elavulása miatt. 1967-ben az UNISIST és az International Council of Scientific Unions által megrendelt összehasonlító tanulmány ugyan az ETO-t találta a meglévő rendszerek közül a legalkalmasabbnak egy információs hálózat kiépítésére, de a végső javaslat mégis egy teljesen új rendszer kialakítása volt. A tanulmány megjelenését követően az ETO strukturális átalakítása és a hatékony menedzseléséhez és karbantartásához szükséges átalakítások kérdései egyaránt előtérbe kerültek.¹⁰⁶

A hatvanas években egy másik, átfogó kutatás is zajlott, az Amerikai Fizikai Intézet (American Institutes of Physics, AIP) által, Pauline Atherton és Robert R. Freeman vezetésével.¹⁰⁷ A projekt célja az *ETO gépesített felhasználási lehetőségeinek megvizsgálása* volt. A projekt eredményeként számos, máig érvényes megoldás kidolgozása mellett javaslatokat tettek a szintaktikai szabályok precízebb rögzítésére és a jelölések pontosítására is.¹⁰⁸

Egyre nyilvánvalóbbá vált az is, hogy az osztályozási bizottságokra alapuló „túlzottan demokratikus” struktúra nem teszi lehetővé az osztályozás hatékony karbantartását és módosítását. Elsőként Wellisch vetette fel a revíziós rendszer centralizált módon történő átalakítását és egy pontos revíziós szabályzat elkészítésének a szükségességét¹⁰⁹, és már ebben az időben felmerült a táblázatok gépesítésének lehetősége is, az eredményesebb együttműködés és a módosítások hatékonyabb végrehajtása érdekében.¹¹⁰ Érdekes és értékes járulékos eredményként említhetők *Ingetraut Dahlberg* elgondolásai¹¹¹, melyeket későbbi, egyetemes osztályozási rendszerekkel kapcsolatos eredményeiben, illetve saját osztályozási rendszerében fejlesztett tovább.¹¹²

A munkálatok eredményeként először *a jelzetszerkesztési szemlélet megváltoztatására és az ETO új szabályzatának a kiadására* került sor.¹¹³ A nyolcvanas években előbb a CCC szűnt meg, majd – egy az

¹⁰⁵ Babiczky 1978. A revízió és kiadás szabályait a „UDC revision and publication procedure” című kiadványban rögzítették (FID 1968; hiv. Slavić, Cordeiro és Riesthuis 2008).

¹⁰⁶ Slavić, Cordeiro és Riesthuis 2008.

¹⁰⁷ A projektet az Egyesült Államok Nemzeti Tudományos Alapja (National Science Foundation, NSF) támogatta, miután a szovjet Referatnij Zsurnal referátumait elkezdték ETO-val indexelni (Bourne és Hahn 2003: 69).

¹⁰⁸ Atherton és Freeman 1969; hiv. Slavić, Cordeiro és Riesthuis 2008.

¹⁰⁹ Wellisch 1971; hiv. Slavić, Cordeiro és Riesthuis 2008.

¹¹⁰ Rigby 1971; hiv. Slavić-Overfield 2005: 16.

¹¹¹ Dahlberg 1972.

¹¹² Dahlberg 1978 és Dahlberg 2017.

¹¹³ Babiczky 1978 és 1985. A régi jelzetszerkesztési elvek inkább Ranganathan „kép-fal elvének” feleltek meg (Ranganathan 1967: 425-426), az újabbak inkább a jelzetelemek posztkoordinált visszakeresésének elősegítését tartották szem előtt, bár a „kép-fal elv” továbbra is megjelenik a nem független általánosan és a korlátozottan közös alosztások kapcsán. Az ETO alapelveit, illetve a revíziójára vonatkozó szabályokat tartalmazó kiadvány német, francia és angol nyelven 1981-ben (FID 1981), a magyar nyelvű fordítással is kiegészített változata pedig 1983-ban jelent meg (FID 1983).

ETO fenntartási lehetőségeit vizsgáló munkacsoport ¹¹⁴ javaslatai alapján – a korábbi középkiadásoknak megfelelő méretű központi adatbázis, a **Master Reference File (MRF, ETO mesterfájl)** készült el.¹¹⁵ Ezzel párhuzamosan az ETO gondozását átvette egy újonnan alapított, az ETO különböző nemzetiségű kiadói álló konzorcium, melynek alapításakor a FID is tagja volt. A munkacsoport javaslatai közé tartozott az osztályozás facettás jellegének tudatos erősítése is.¹¹⁶

Az **ETO Konzorcium (UDC Consortium, UDCC)** önfinanszírozó, nonprofit szervezet, melynek célja az ETO fejlesztése és terjesztése.¹¹⁷ Az ETO-t – beleértve az MRF forgalmazásának a jogait – 1992. január 1. óta a konzorcium birtokolja. *A séma lefordítása, kiadása és minden egyéb felhasználása a megfelelő licenz megszerzéséhez van kötve.*¹¹⁸ 2001 óta az adatbázis *online szolgáltatás* formájában is elérhető.¹¹⁹

Az ETO szerkesztőbizottsága 2007-ben lett átalakítva, ettől az évtől a főszerkesztő egy kisebb létszámú szerkesztői csoporttal és egy nagyobb létszámú tanácsadói bizottsággal együttműködve gondoskodik a táblázatok karbantartásáról, aktualizálásáról és felújításáról.¹²⁰

A tulajdonosváltás az Extensions and Corrections kiadásában is változást hozott: 1993-tól 2011-ig a kiadvány évente, (nyomtatott vagy digitális) könyv formában jelent meg, mely az ETO változásai mellett a rendszerrel és a konzorciummal kapcsolatos híreket, ajánlásokat, publikációkat is tartalmazott. A 2012-13-as illetve 2014-2015-ös számok összevonva jelentek meg, a továbbiak pedig csak digitálisan (ingyenesen) lesznek elérhetőek. Szintén évente kerül publikálásra (az Extensions and Correctionsben és a konzorcium honlapján is) az ETO-val kapcsolatos publikációk bibliográfiája.¹²¹

Időközben a legfontosabb szellemi produktumáról lemondó FID az átszervezések ellenére sem tudott megbirkózni az új kihívásokkal, és 2001-ben hivatalosan is beszüntette működését.¹²² Az ETO-t

¹¹⁴ Az ETO tulajdonosi szerkezetének átalakítását, az MRF adatbázis kialakítását és a további következetes facettásítást javasoló munkacsoport tagjai A. R. Harala (Finnország), H. Jobst (Ausztria), I. C. McIlwaine (Egyesült Királyság), Gerhard Riesthuis (Hollandia) és Nancy Williamson (Egyesült Királyság) voltak, a felügyeletet pedig David Strachan (FID) és A. Gilchrist látták el (FID 1990; hiv. Slavić, Cordeiro és Riesthuis 2008).

¹¹⁵ A Master Reference File géppel olvasható formában 1993 közepén lett publikálva (McIlwaine 1998). Jelenleg kb. 70.000 osztályt tartalmaz, melyek központilag, angol nyelven vannak karbantartva („UDC Master Reference File”).

¹¹⁶ McIlwaine 1998 és FID 1990; hiv. Slavić, Cordeiro, és Riesthuis 2008.

¹¹⁷ „UDC Consortium (UDCC)”. A UDC Consortium a holland jog szerint, holland alapítványként bejegyezve működik, székhelye Hágában, a Holland Nemzeti Könyvtár épületében található (Slavić, Cordeiro és Riesthuis 2008).

¹¹⁸ „UDC Licences”.

¹¹⁹ A UDC Online előfizetés esetén a táblázatok legfrissebb változatához nyújt hozzáférést, kereshető és böngészhető formában („UDC Online”). Kutatási és oktatási célokra ingyenes licenz igényelhető („UDC Online English – Prices”). Emellett a mintegy 2600 osztályt tartalmazó UDC Summary ingyenesen is elérhető („UDC Summary”).

¹²⁰ Slavić, Cordeiro és Riesthuis 2008. Jelenleg az ETO Konzorcium elnöke John Akeroyd, alelnöke Maria Inês Cordeiro, az ETO főszerkesztője pedig Aida Slavić. A szerkesztői csoport és a tanácsadói bizottság tagjainak listája megtekinthető a konzorcium honlapján („UDC Editorial Team & Advisory Board”).

¹²¹ „Extensions and Corrections to the UDC”.

¹²² Földi 2001 és Shimmon 2001. A FID archívuma jelenleg a Holland Nemzeti Könyvtárban van elhelyezve. Az archívum feldolgozása és digitalizálása 2012. végén kezdődött el (Slavić 2012 a).

továbbra is a konzorcium tartja karban, folytatva a korszerűsítés - a tudásterületek átrendeződésének követése, a facettanalízis elvének következetes és elméletileg is megalapozott érvényesítése¹²³, a gépesíthetőség megkönnyítése, a rendszer nagyobb fokú következetességének elérése¹²⁴ - irányába mutató átszervezéseket¹²⁵, a rendszer népszerűsítésére és megismertetésére tett törekvéseket¹²⁶, az MRF és az online ETO karbantartását és az osztályozással kapcsolatos kutatások támogatását. Utóbbi tevékenységgel kapcsolatosan kiemelendő a 2007 óta két évente megszervezett **ETO Szeminárium (UDC Seminar)**, amely a világ legrangosabb osztályozástudományi konferenciái közé tartozik.¹²⁷

2.2 Az ETO elterjedtsége

Az ETO már megjelenésekor nagy nemzetközi sikert aratott. Már az első kiadás elkészülte előtt, vagy közvetlen annak megjelenését követően elkezdtek használni könyvtári célokra is, például Magyarországon¹²⁸ vagy Spanyolországban¹²⁹.

Az azóta eltelt százhusz évben az ETO *a második legelterjedtebb osztályozási rendszerré* vált, kevéssel elmaradva a DDC mögött, és megelőzve a LCC-t.¹³⁰

A legutóbbi nemzetközi felmérést az ETO-t használó könyvtárak számára vonatkozóan az ETO szerkesztőbizottsága indította 2012 elején: az előzetes adatok alapján csak *Európában legalább 28 ország 144.000 könyvtára* alkalmazza ezt az osztályozási rendszert.¹³¹ Ezt megelőzően a legutóbbi teljes körű felmérés eredménye 2006-ban lett publikálva.¹³² Ennek során a világ *208 országát* mérték fel, melyek közül *124-ben használják, 34-ben első számú osztályozási rendszerként*. Utóbbi elsősorban Európában és Afrikában jellemző, illetve Ázsia azon országaiban, melyek egykor a Szovjetunióhoz tartoztak. Ázsiában leginkább Indiában van elterjedve (Ranganathan kettőspontos osztályozása és a DDC mellett), ahol egyébként is nagyon nagy hagyományai vannak az osztályozási rendszerek használatának és kutatásának.¹³³ Észak- és Közép-Amerikában inkább a másik két nagy osztályozási

¹²³ Broughton 2010.

¹²⁴ Slavić és Davies 2017.

¹²⁵ A változások nagyságrendjét jól mutatja az olyan teljesen új osztályok részletes kidolgozása, mint az informatika (004) és a menedzsment (005), vagy a tulajdonságok (-02) és műveletek (-04) általánosan közös alosztásai, illetve, a vallás és teológia (2) és a történelem (94) már megvalósult, illetve a filozófia és az orvostudomány folyamatban lévő teljes körű átszervezése („Major Changes to the UDC 1993-2013”, McIlwaine 2006, McIlwaine és Williamson 2008 és Gnoli 2009). Az eredeti terv szerint tíz év alatt az egész osztályozás revíziójára sor került volna, ez azonban tarthatatlannak bizonyult (Slavić, Cordeiro és Riesthuis 2007), és a munkálatok azóta is tartanak.

¹²⁶ Érdekes és értékes az ETO oktatási szolgáltatás, amelynek keretében olyan kurzusokat lehet igényelni, amelyek standard tananyag vagy egyedi igények alapján, a lehető legmagasabb színvonalon segítik elő az osztályozás megismerését (minden érdeklődő számára a kezdő osztályozóktól a professzionálisokig, akik a rendszer változásaival és jelenlegi állapotával ismerkedhetnek meg) („UDC Training Service”).

¹²⁷ „International UDC Seminars”.

¹²⁸ Barátné 2007b.

¹²⁹ San Segundo 2007.

¹³⁰ Slavić 2008. Ugyan szólnak érvek az ETO-DDC (Fandino 2008), illetve ETO-LCC (San Segundo 2008) váltás mellett és ellen is, maga az átállási folyamat is nagyon költséges, ezért ezek a számok nagyságrendileg általában csak lassan változnak.

¹³¹ Slavić 2012b.

¹³² Slavić 2008.

¹³³ Satija 2008 és Singh 2008.

rendszer használata általános, Dél-Amerikában pedig, bár az országok többségében használják, de inkább csak tudományos és speciális gyűjtőkörű könyvtárakban. A jelentés szerint az óceániai térségben egyedül Ausztráliában mondható jelentősnek az ETO-t használó intézmények száma.

Az idézett felmérések szerint jelenleg az ETO *leginkább az Ibériai-félszigeten, illetve Kelet- és Közép-Európában számít dominánsnak*. Tulajdonképpen valamennyi olyan európai ország, ahol első számú osztályozási rendszernek számít itt található, és jelenleg az ETO Konzorcium tagjai is mind ebből a két körzetből kerülnek ki.¹³⁴

Spanyolországban nagy hagyományai vannak az ETO-nak¹³⁵, Portugáliában pedig különösen a Portugál Nemzeti Bibliográfiai Adatbázisban (PORBASE) betöltött szerepe miatt jelentős.¹³⁶

Az idézett 2006-os felmérés idején Kelet-Közép Európa országai közül csak három országban nem számított első számú osztályozási rendszernek, ám ezek az adatok Ukrajna¹³⁷ és különösen Oroszország¹³⁸ esetében nem tükrözik a rendszer valódi jelentőségét. *Oroszország* azért is különleges, mert ez az egyetlen ország, amelyik a Konzorcium megalakulása és az MRF megjelenése után *továbbra is karbantartja és kiadja a teljes ETO táblázatokat*.¹³⁹ A szovjet utódállamok közül az osztályozásnak a Baltikumban is sikerült megtartania a vezető szerepét.¹⁴⁰

A 2006-os felmérés 30 nemzeti fordítást említ (ez a vizsgált országok 27 százalékát jelenti). Azoknak az országoknak a többségében, ahol nem volt ilyen, az angol táblázatokat használták. Konkrétan az ETO fordításaira vonatkozó teljességre törekvő felmérés legutóbb 2004-ben készült, ez *39 különböző fordítást* említ.¹⁴¹ A bejelentett ETO kiadások 2011-ben lezárt bibliográfiája szintén az utóbbi adatot

¹³⁴ A konzorcium tagjai jelenleg: a Spanyol Szabványosítási és Érvényesítési Testület (AENOR), az Orosz Akadémia Összorosz Tudományos és Műszaki Információs Intézete (VINITI), a Cseh Köztársaság Nemzeti Könyvtára, a Szlovák Nemzeti Könyvtár, a Ljubljana Nemzeti és Egyetemi Könyvtár és a Zágrábi Nemzeti és Egyetemi Könyvtár („UDC Consortium”).

¹³⁵ Az ország az elsők közé tartozott, ahol könyvtárakban alkalmazták, később ötven évig rendelet írta elő használatát a közkönyvtárakban, és jelenleg is csaknem minden könyvtárban ez az elsődleges tartalmi feltáró eszköz (San Segundo 2007). A legújabb rövidített ETO kiadás is Spanyolországban készült el 2016 nyarán (New Spanish Abridged Edition). A UDC Online is elérhető spanyolul („UDC Online”), a UDC Summary pedig baszk, katalán és galíciai nyelvre is le van fordítva („UDC Summary Translations”).

¹³⁶ „Classificação Decimal Universal – CDU”. A Portugál Nemzeti Könyvtár az egyetlen olyan könyvtár, amelynek az OCLC The Virtual International Authority File-jába feltöltött rekordjai tartalmazzak osztályozási jelzeteket („VIAF: The Virtual International Authority File.”) és a Portugál Digitális Nemzeti Könyvtár – nyitott linked data-ként is elérhető – állománya is indexelve van az ETO-val („Digital National Library of Portugal”). A Nemzeti Könyvtár és a Konzorcium közötti szoros kapcsolatot az is mutatja, hogy az előbbi főigazgatója és az utóbbi alelnöke ugyanaz a személy, Dr. Maria Inês Cordeiro.

¹³⁷ Ukrajnában elsősorban a főiskolai és tudományos információs központok használják, 2008-ban megjelent az ukrán nyelvű kiadása is, illetve 2009-ben szabványként is elfogadták (Nabhan 2010).

¹³⁸ Oroszországban ugyan másodlagos osztályozási rendszernek számít az ETO, mert „csak” a tudományos és műszaki információs központok és könyvtárak, valamint a VINITI használják, azonban ez is több, mint 5.000 könyvtárat jelent. A rendszer jelentőségét mutatja az is, hogy 2006-ben GOST szabvánnyá vált (GOST 7.90-2007. 2006).

¹³⁹ A legutóbbi, negyedik orosz nyelvű teljes kiadás 2001 és 2009 között jelent meg (Arszkij 2008).

¹⁴⁰ Noreikiene 2008 és Makke 2016. A bejelentett kiadások listája szerint a legutóbbi, speciális lett nyelvű kiadás 1996-ban készült el („Last reported editions of the Universal Decimal Classification”).

¹⁴¹ Slavić 2004.

támasztja alá.¹⁴² Mindkét forrás jól mutatja, milyen új nemzeti kiadások születtek Kelet-Közép-Európa geopolitikai változásait követően, ugyanakkor azt is, hogy a megjelentetéshez szükséges jogdíjak hogyan csökkentették a korszerűsített, új kiadások számát.¹⁴³

A **UDC Online** felületen keresztül *hét nyelven* (angolul, hollandul, csehül, horvátul, spanyolul, szlovénul és franciául – a német fordítás egyelőre még nem aktív) érhetők el a táblázatok¹⁴⁴, a **UDC Summary** pedig *ötvenhét nyelvre lefordítva* tartalmazza a szabadon elérhető 2600 osztályt.¹⁴⁵

Az ETO-val indexelt *legnagyobb gyűjtemény* egyértelműen a VINITI RAS 28 millió rekordot tartalmazó adatbázisa, melyet a szlovák nemzeti könyvtár 4 milliós állománya követ. De használják az ETO-t a Svájci Könyvtárak és Információs Központok Hálózata (NEBIS) és Csehország, Magyarország, Portugália, Románia, Szlovénia közös katalógusai is. Több, mint húsz európai (és néhány afrikai, illetve egyetlen ázsiai) ország pedig *nemzeti bibliográfiájában* is alkalmazza, ahogyan az Amerikai Meteorológiai Társaság is a Meteorological & Geostrophysical Abstracts adatbázisában.¹⁴⁶

A feltárás alapjául szolgáló gyűjtemény nem csak a klasszikus értelemben vett dokumentumok és műalkotások, hanem például kutatási vagy adminisztratív adatok állományát is jelentheti. Az ETO elterjedtségét és a modern igényeknek való megfelelését mutatják az *új típusú adatok feltárására* irányuló olyan próbálkozások, mint a UK Data Service kísérleti projektje, melynek során az általuk tárolt adatokat a korábban használt tezaurusz tárgyszavakon kívül ETO-jelzetekkel is ellátták a hatékonyabb tárgyi hozzáférés érdekében.¹⁴⁷

Az ETO nem csak közvetlenül szolgálhat gyűjtemények feltárására. Gazdag szókészlete miatt gyakran használják *tezauruszok és más osztályozási rendszerek terminológiai forrásaként* is. Ilyen speciális osztályozási rendszerek pl. a BBC archívuma által használt LonClass¹⁴⁸, a Skót Sarkkutató Intézet (SPRI) által használt Polar UDC¹⁴⁹ vagy a VINITI által kiadott Referativni Zhurnal matematikai osztályozási rendszere¹⁵⁰. Az is előfordul, hogy *kiegészítik egymást egy tárgyi osztályozással* úgy, hogy az adott szakterületen a szakosztályozás, a többi területen pedig az ETO kerül alkalmazásra, mint például a Global Forest Decimal Classification (GFDC) esetében.¹⁵¹

¹⁴² „Last reported editions of the Universal Decimal Classification”. A bibliográfia 40 nyelvet sorol fel, de az azerbajdzsánit bibliográfiai adatok nélkül, megerősítetlenül említi, ugyanakkor hiányoznak belőle a legutóbbi magyar nyomtatott kiadásra vonatkozó adatok.

¹⁴³ Slavić et. al. 2007.

¹⁴⁴ „UDC Online”.

¹⁴⁵ Az ETO elterjedtsége szempontjából a képet árnyalja az, hogy az 56 nyelv tartalmazza az eszperantót, a már említett baszkot, katalánt és galíciait, illetve az, hogy hat nyelvet kizárólag, további hatot pedig nagyobb részt Indiában beszélnek („UDC Summary Translations”). Edgardo Civallero latin-amerikai indián nyelvekre való fordítással is kísérletezett, sikertelenül (nem elsősorban a nyelvi, inkább a kulturális-gondolkodási eltérések által okozott akadályok miatt) (Civallero 2016).

¹⁴⁶ „Collections indexed by UDC”.

¹⁴⁷ Barbalet 2015, Barbalet és Cunningham 2017.

¹⁴⁸ Alexander et. al. 2010.

¹⁴⁹ Gilbert és Lane 2008.

¹⁵⁰ „Referativni Zhurnal Classification Scheme”.

¹⁵¹ „Global Forest Decimal Classification Database”. A GFDC főosztályai a UDC Summaryben is megtalálhatók, a 630-as (Erdészeti) osztály alatt, külső forrásból származó jelzetekként (ld. <http://udcdata.info/050821>).

Az ETO és más nagy osztályozási rendszerek egymásra is kölcsönös befolyással voltak és vannak. Mint az első facettás és analitikus-szintetikus osztályozás, amely negyven évvel megelőzte Ranganathan és Bliss rendszereit, természetesen meghatározta az azok kialakítása során használt elveket és módszereket. Bár saját állítása szerint az összetett tárgyak koncepcióját egy gyermekjáték sugallta neki¹⁵², feltételezhető, hogy Ranganathan a kettőspont alkalmazásának ötletét az ETO-ból merítette, de egyéb direkt és indirekt behatások is kimutathatók az osztályozási rendszerében. Ugyanakkor a későbbi – főleg az elmúlt húsz évben történt – revíziókon keresztül Ranganathan elméleti munkássága és a Kettőspontos Osztályozás is nagy hatással volt az ETO további fejlődésére.¹⁵³ A Bliss Classification Association és az ETO Konzorcium között pedig még formális megállapodás is született arról, hogy átvehetik egymás szemantikai struktúráit.¹⁵⁴ Az ETO-ban elsősorban a teljesen facettás formában átalakított kettes („Vallás. Teológia”) főosztály tükrözi a BC2 kialakítása során szerzett tapasztalatokat¹⁵⁵, és ennek a főosztálynak a DDC-vel való összehangolása mutatja a Dewey-féle osztályozással továbbra is élő kapcsolatot.¹⁵⁶ Egy másik példa hasonló együttműködésre az ETO és az orosz Rubrikátor folyamatos összehangolása a VINITI által.¹⁵⁷ Szintén a VINITI munkatársai kísérleteznek az Oroszországban használt tárgyi leíró szótárak – köztük az ETO - automatikus összekapcsolásával a Web of Science (WoS) tárgyi kategóriáin keresztül.¹⁵⁸

2.2.1 Az ETO Magyarországon

Az ETO-nak Magyarországon különösen nagy hagyományai vannak. *Valamennyi könyvtártípusban elterjedt*, és a legtöbb könyvtár elmondhatja magáról, hogy vagy jelenleg is használja, vagy a múltban használta legalább egyszer. *A könyvtárhasználat és -tudomány oktatásának valamennyi szintjén foglalkoznak vele*, mind elméleti, mind gyakorlati szinten.¹⁵⁹ A legfrissebb, 2016 első negyedévében készült felmérésben megkérdezett 239 könyvtár 84 százaléka használja ezt az osztályozást (és a maradék majdnem fele nyilatkozott úgy, hogy korábban használta, de valamilyen okból felhagyott vele), túlnyomó többségük aktív feldolgozó munkát is végezve vele. Más tartalmi feltáró eszköz (jellemzően valamilyen tárgyszórendszert) ezeknek a könyvtáraknak csak a fele használ.¹⁶⁰

Az ETO-nak már a kezdetektől vannak magyar vonatkozásai. A magyar szakirodalomban gyakran említett tény, hogy *Mandelló Gyula* hívta fel Paul Otlet figyelmét az ETO-ra, ahogyan az is, hogy Magyarországon elsőként **Szabó Ervin** alkalmazta a rendszert 1900-ban¹⁶¹ a Kereskedelmi és Ipartestületi Kamara könyvtárában, később pedig a Fővárosi Könyvtárban is, és az első nyomtatott kiadás megjelentetése (1914-ben), illetve az ETO-használat oktatásának megkezdése is az ő nevéhez fűződik. Az *Országos Széchényi Könyvtár* 1936-ban kezdte használni az Egyetemes Tizedes

¹⁵² Garfield 1984.

¹⁵³ Chatterjee 2015.

¹⁵⁴ McIlwaine 1998.

¹⁵⁵ Broughton 2010. Az orvostudományok több, mint húsz éve zajló revíziója szintén a BC2 megfelelő osztálya alapján kezdődött el 1993-ban (McIlwaine és Williamson 2008).

¹⁵⁶ McIlwaine 2006.

¹⁵⁷ Arszkij 2008.

¹⁵⁸ Beloozerov et. al. 2017.

¹⁵⁹ Barátné 2007b.

¹⁶⁰ Bognár és Tóth 2016.

¹⁶¹ Tehát az ETO-t könyvtári célra világszinten is Szabó Ervin használta először, még az első, francia nyelvű kiadás publikálása előtt.

Osztályozást, az *első tankönyv* pedig – Káplány Géza munkája – 1943-ban jelent meg.¹⁶² Széles körű elterjedése a második világháború utánra tehető, 1948 után szinte valamennyi típusú könyvtár és dokumentációs intézmény vállalta az ETO használatát, a *Magyar Nemzeti Bibliográfia* 1946, a *Magyar Folyóiratok Repertórium*a (később Magyar Időszaki Kiadványok Repertórium) pedig 1947 óta adja közre ETO-jelzetekkel is bővítve a bibliográfiai rekordokat.¹⁶³

Számos nem hivatalos rövidített kiadás után az *első hivatalos, szabványként megjelentetett táblázat* 1966-ban jelent meg, a következő pedig 1977-ben. Szintén 1966-ban kezdődött el a *teljes magyar nyelvű ETO* kiadása. Ennek utolsó kötete 1971-ben látott napvilágot, de ezt később folyamatosan követték az egyes kötetek korszerűsített újra kiadásai is.¹⁶⁴ A következő rövidített kiadásra **1990/1991**-ig kellett várni, a teljes kiadás utolsó két kötete pedig 1992-ben jelent meg. Ezt hosszabb szünet követte, amíg **2005**-ben, az MRF 2000-ben lezárt anyagával, a korábbiaktól eltérő módon és bővebb fogalomtárral megjelent a legutóbbi nyomtatott magyar nyelvű ETO-kiadás¹⁶⁵, az OSZK Könyvtári Intézetének gondozásában.¹⁶⁶ A *legújabb, elektronikus kiadás előkészületei* 2016 első felében kezdődtek el. Első lépésben az adatbázis valószínűleg az OSZK saját tulajdonú Relex névtér kezelő alkalmazásába kerül feltöltésre, amely az ETO mellett teauruszok kezelésére is alkalmas, így lehetőség nyílik az osztályozás és a tárgyszórendszerek összekapcsolására.¹⁶⁷

Magyarország a FID-ben – különböző intézményeken keresztül és változó aktivitással – 1937 óta képviseltette magát.¹⁶⁸ A szerkesztőbizottságban az ETO Konzorcium megalakulásáig hét magyar tag szerepelt, a 33-as főosztály albizottságának pedig magyar elnöke is volt, Földi Tamás személyében.¹⁶⁹ Jelenleg a konzorcium szerkesztői csoportjában Barátné dr. habil. Hajdu Ágnes képviseli hazánkat.¹⁷⁰

A magyar könyvtárosok levelezőlistáján, a Katalisten többször is voltak elhúzódo, mély viták az ETO további használatának szükségességéről, illetve annak alternatíváiról.¹⁷¹ Mint minden vitában, itt is értékes érvek és hasznos információk merültek fel mindkét oldal részéről. Ám az ETO-val szembe helyezkedők többnyire nem konkrétan az ETO, hanem a mesterséges információkereső nyelvek hasznosságát vitatták, illetve az informatikai támogatottság hiányából fakadó problémákat vetettek fel. Maga a vita ténye viszont azt is mutatja, hogy az ETO továbbra is foglalkoztatja a magyar könyvtáros közvéleményt.

¹⁶² Barátné 2007b.

¹⁶³ Babiczky 1975.

¹⁶⁴ Babiczky 1987.

¹⁶⁵ A 2005-ös az első magyar nyelvű kiadás, amely már az MRF alapján készült. A magyar nyelvű szakirodalom és oktatás, hangsúlyozandó a korábbiakhoz viszonyított változás mértékét sokáig csak „új ETO-ként” hivatkozott erre a kötetre. Ugyanakkor a könyvtárakban, továbbra is elterjedt az 1990/1991-es kiadás használata (Bognár és Tóth 2016).

¹⁶⁶ Az új kiadás (OSZK KI 2005) elsősorban Barátné Hajdu Ágnes és Fejős László áldozatos munkájának köszönhetően jött létre (Barátné és Fejős 2005 és Barátné 2007a). Mivel ez volt az első magyar kiadás az MRF létrehozását követően, a változásokat tartalmazó konkordanciátáblázatok elkészítése is nagy kihívást jelentett (Fejős 2007, Fejős 2008a, Fejős 2008b, Fejős 2008c és Fejős 2009).

¹⁶⁷ Dancs 2016. A Relex névtérkezelő használatáról ld. Ungváry 2013.

¹⁶⁸ Balázs 1960.

¹⁶⁹ Barátné 2007b.

¹⁷⁰ „UDC Editorial Team & Advisory Board”.

¹⁷¹ Fejős 2000 és „Katalist. 2014 januári archívum”.

A könyvtár és információtudománynak központi területe az osztályozási rendszerek elmélete és gyakorlata, ahogyan a könyvtáros identitásnak is fontos eleme legalább egy osztályozási rendszer alkalmazásának képessége.¹⁷² Elterjedtsége és az oktatásban betöltött szerepe miatt Magyarországon ez általában az Egyetemes Tizedes Osztályozás ismeretét jelenti.

2.3 Irodalomjegyzék

„About Universal Decimal Classification (UDC).” UDCC (UDC Consortium).

<http://udcc.org/index.php/site/page?view=about>. (megtekintve: 2017-11-01)

Alexander, Fran, Kathryn Stickley, Vicky Buser, és Libby Miller. 2010. UDC at the BBC. *Extensions and Corrections to the UDC* 32: 21-24.

Arszkij, Jurij M. 2008. „UDC activities in Russia: A VINITI report.” *Extensions and Corrections to the UDC* 30: 49-52. <http://arizona.openrepository.com/arizona/handle/10150/105158>. (megtekintve: 2017-11-01)

Atherton, Pauline, Robert R. Freeman. 1969. „Final report of the research project for the evaluation of the UDC as the indexing language for a mechanized retrieval system.” In *Proceedings of the First Seminar on UDC in a Mechanized Retrieval System conducted by R. R. Freeman and P. Atherton, Copenhagen, 2nd-6th September, 1968. (FID/CR Report no 9)*, szerk. R. Molgaard-Hansen, és Malcolm Rigby, 29-37. Copenhagen: Danish Centre for Documentation.

Babiczyk Béla. 1975. „A 80 éves Nemzetközi Dokumentációs Szövetség és az Egyetemes Tizedes Osztályozás Magyarországon.” *Tudományos és műszaki tájékoztatás* 22 (9): 649–655. http://tmt-archive.omikk.bme.hu/show_news.html?id=2312&issue_id=86. (megtekintve: 2017-11-01)

Babiczyk Béla. 1978. „Új szakasz az Egyetemes Tizedes Osztályozás fejlődésében.” *Könyvtári figyelő* 24 (6): 659-671.

Babiczyk Béla. 1985. „Szemléletváltozás az ETO jelzetszerkesztésében.” *Könyvtári Figyelő* 31 (1): 17–27.

Babiczyk Béla. 1987. „A magyar ETO-kiadások húsz éve (1966-1986).” *Tudományos és műszaki tájékoztatás* 34 (11): 543–550. http://tmt-archive.omikk.bme.hu/show_news.html?id=3079&issue_id=375. (megtekintve: 2017-11-01)

Balázs Sándor. 1960. „Magyarország képviselte és szerepe a FID-ben.” *Magyar könyvszemle* 76 (1): 58-64.

Barátné Hajdu Ágnes. 2007a. „A magyar ETO-kiadás munkálatai. Hazai és nemzetközi tapasztalatok.” *Könyv, könyvtár, könyvtáros* 16 (2): 20-39. <http://epa.oszk.hu/01300/01367/00122/pdf/02muhelykerdesek.pdf>. (megtekintve: 2017-11-01)

Barátné Hajdu Ágnes. 2007b. „Multilevel education, training, traditions and research on UDC in Hungary.” *Extensions and Corrections to the UDC* 29: 273-284. <http://hdl.handle.net/10150/105607>. (megtekintve: 2017-11-01)

Barátné Hajdu Ágnes. 2011. „Az ismeretszervezés dinamizmusa, különös tekintettel az Egyetemes Tizedes Osztályozásra.” *Könyv és nevelés* 13 (1): 63-72.

¹⁷² Barátné 2011.

- Barátné Hajdu Ágnes, és Fejős László. 2005. „New Hungarian UDC edition.” *Extensions and corrections to the UDC 27*: 8-12.
- Barbalet, Suzanne. 2015. „Enhancing subject authority control at the UK Data Archive: a pilot study using UDC.” In *Classification & authority control: expanding resource discovery: proceedings of the International UDC Seminar 2015, 29-30 October 2015, Lisbon, Portugal*, szerk. Aida Slavić & Maria Inês Cordeiro, 203-212. Würzburg: Ergon Verlag.
- Barbalet, Suzanne, és Nathan Cunningham. 2017. „The challenge of managing access to new and novel forms of data: an application of UDC.” In *Faceted classification today: theory, technology and end users: proceedings of the International UDC Seminar 2017, London (UK), 14-15 September*, szerk. Aida Slavić, és Claudio Gnoli, 133-146. Würzburg: Ergon Verlag.
- Beloozerov, Viktor N., Marat R. Biktimirov, Aleksandr B. Antopolskij, Olga A. Antoshkova, Tatyana S. Astaxova, és Olga V. Smirnova. 2017. „Establishing correspondences between Web of Science subject areas and UDC.” In *Faceted classification today: theory, technology and end users: proceedings of the International UDC Seminar 2017, London (UK), 14-15 September*, szerk. Aida Slavić, és Claudio Gnoli, 305. Würzburg: Ergon Verlag.
- Bognár Noémi Erika, és Tóth Máté. 2016. „Tartalmi feltáró eszközök használata a magyarországi könyvtárakban.” *Könyv, könyvtár, könyvtáros* 25 (6): 18-30.
<http://ki.oszk.hu/3k/2016/07/tartalmi-feltaro-eszkozok-hasznalata-a-magyarorszag-i-konyvtarakban/>. (megtekintve: 2017-11-01)
- Bourne, Charles P., és Trudi Bellardo Hahn. 2003. *A History of Online Information Services, 1963-1976*. Cambridge, MA, and London: MIT Press.
- Broughton, Vanda. 2010. „Concepts and Terms in the Faceted Classification: the Case of UDC.” *Knowledge Organization* 37 (4): 270-279.
- Chatterjee, Amitabha. 2015. „Universal Decimal Classification and Colon Classification. Their mutual impact.” *Annals of Library and Information Studies (ALIS)* 62 (4): 226-230.
<http://op.niscair.res.in/index.php/ALIS/article/view/11208>. (megtekintve: 2017-11-01)
- Civallero, Edgardo. 2016. „UDC and Indigenous People.” *IFLA Metadata Newsletter* 2 (2): 37-39.
http://www.ifla.org/files/assets/classification-and-indexing/ifla_metadata_newsletter_dec2016.pdf. (megtekintve: 2017-11-01)
- „Classificação Decimal Universal - CDU.” Biblioteca Nacional de Portugal (BNP).
http://www.bnportugal.pt/index.php?option=com_content&view=article&id=485&Itemid=546&lang=en. (megtekintve: 2017-11-01)
- „Collections indexed by UDC.” UDCC (UDC Consortium).
<http://udcc.org/index.php/site/page?view=collections>. (megtekintve: 2017-11-01)
- Dahlberg, Ingetraut. 1972. „Az ETO újjáalakításának lehetőségeiről” [Fordította Horváth Magda]. *Könyvtári Figyelő* 18 (4): 412-430.
- Dahlberg, Ingetraut. 1978. *Ontical Structures and Universal Classification*. Bangalore. Sarada Ranganathan Endowment for Library Science.
- Dahlberg, Ingetraut. 2017. „Why a New Universal Classification System is Needed.” *Knowledge Organization* 44 (1): 67-71.
- Dancs Szabolcs. 2016. „Ismét az ETO hasznáról - az új elektronikus kiadás fényében.” *Könyv, könyvtár, könyvtáros* 25 (4): 3-8.

- http://epa.oszk.hu/01300/01367/00277/pdf/EPA01367_3K_2016_04_003-008.pdf.
(megtekintve: 2017-11-01)
- „Dewey Decimal Classification summaries.” Online Computer Library Center (OCLC).
<https://www.oclc.org/dewey/features/summaries.en.html>. (megtekintve: 2017-11-01)
- „Digital National Library of Portugal.” The European Library. National Library of Portugal.
<http://www.theeuropeanlibrary.org/tel4/collection/a0005>. (megtekintve: 2017-11-01)
- „Extensions and Corrections to the UDC.” UDCC (UDC Consortium).
<http://udcc.org/index.php/site/page?view=ec>. (megtekintve: 2017-11-01)
- Fandino, Marta. 2008. „UDC or DDC: a note about the suitable choice for the National Library of Liechtenstein.” *Extensions and Corrections to the UDC* 30: 45-48.
- Fejős László. 2000. „Kell-e még az ETO?” *Könyv, könyvtár, könyvtáros* 9 (3): 23-40.
<http://epa.oszk.hu/01300/01367/00003/pdf/04muhelykerdesek.pdf>. (megtekintve: 2017-11-01)
- Fejős László. 2007. „Az ETO adatbázis magyar kiadása és konkordanciája.” *Könyvtári levelező/lap* 19 (7): 3-6. <http://epa.oszk.hu/00300/00365/00068/pdf/00068.pdf>. (megtekintve: 2017-11-01)
- Fejős László. 2008a. „Adalékok az ETO új kiadásához és konkordanciájához.” *Könyv, könyvtár, könyvtáros* 17 (3): 19-28. <http://epa.oszk.hu/01300/01367/00187/pdf/02muhelykerdesek.pdf>.
(megtekintve: 2017-11-01)
- Fejős László. 2008b. Adalékok az ETO új kiadásához és konkordanciájához (folytatás). *Könyv, könyvtár, könyvtáros* 17 (4): 26-32.
<http://epa.oszk.hu/01300/01367/00188/pdf/04muhelykerdesek.pdf>. (megtekintve: 2017-11-01)
- Fejős László. 2008c. Nyomdakész az ETO konkordancia. *Könyvtári Levelező/lap* 20 (12): 8, 20.
<http://www.epa.oszk.hu/00300/00365/00085/pdf/00085.pdf>. (megtekintve: 2017-11-01)
- Fejős László. 2009. ETO gondozás, konkordancia. *Könyvtári levelező/lap* 21 (2): 3-4.
http://www.epa.oszk.hu/00300/00365/00092/pdf/KLL_EPA00365_2009-02_003-004.pdf.
(megtekintve: 2017-11-01)
- FID (International Federation de Documentation). 1968. *Universal decimal classification (UDC) revision and publication procedure = Procedure de revision et de publication de la classification decimale universelle (CDU) = Das Verfahren bei Bearbeitung und Veröffentlichung der universellen Dezimalklassifikation (DK)*. Hága: International Federation de Documentation.
- FID (International Federation de Documentation). 1981. *Principles of the Universal Decimal Classification (UDC) and rules for its revision and publication (FID 598)*. Hága: International Federation de Documentation.
- FID (International Federation de Documentation). 1983. *Az Egyetemes Tizedes Osztályozás (ETO) alapelvei, revíziójának és kiadásának szabályai (FID 603)*. Budapest: OMIKK.
- FID (International Federation de Documentation). 1990. *Task Force For System Development: final report*. Hága: International Federation de Documentation.
- Földi Tamás. 2001. „Búcsú a FID-től.” *Tudományos és műszaki tájékoztatás* 48 (4): 160-162.
http://tmt-archive.omikk.bme.hu/show_news.html?id=1107&issue_id=37. (megtekintve: 2017-11-01)
- Garfield, Eugene. 1984. „A Tribute to S. R. Ranganathan, the Father of Indian Library Science. Part 1. Life and Works”. *Essays of an Information Scientist* 7: 37-44

- Gilbert, Mark, és Heather Lane. 2008. „Forty-five numbers for snow: a brief introduction to the UDC for Polar libraries.” *Extensions and Corrections to the UDC* 30: 23-25.
<http://hdl.handle.net/10150/105332>. (megtekintve: 2017-11-01)
- „Global Forest Decimal Classification (GFDC) Database.” International Union of Forest Research Organizations (IUFRO). <http://iufro.forintek.ca/GFDCDefault.aspx>. (megtekintve: 2017-11-01)
- Gnoli, Claudio. 2009. „The UDC Philosophy Revision: First Report.” *Extensions and Corrections to the UDC* 31: 25-31. <http://hdl.handle.net/10150/200633>. (megtekintve: 2017-11-01)
- Gnoli, Claudio, Philippe Cousson, Tom Pullman, Gabriele Merli, és Rick Szostak. 2011. „Representing the structural elements of a freely faceted classification.” In *Classification and ontology: formal approaches and access to knowledge: Proceedings of the International UDC Seminar, 19-20 September 2011, The Hague, The Netherlands*, szerk. Aida Slavić, és Edgardo Civalero, 193-206. Würzburg: Ergon Verlag.
- Goldschmidt, Robert, és Paul Otlet. 1990a. „On a New Form of the Book: The Microphotographic Book.” In *International Organization and Dissemination of Knowledge: Selected Essays of Paul Otlet*, ford. és szerk. W. Boyd Rayward, 87-93. Amsterdam: Elsevier, 1990.
- Goldschmidt, Robert, és Paul Otlet. 1990b. „The Preservation and International Diffusion of Thought; the Microphotographic Book.” In *International Organization and Dissemination of Knowledge: Selected Essays of Paul Otlet*, ford. és szerk. W. Boyd Rayward, 204-208. Amsterdam: Elsevier, 1990.
- GOST 7.90-2007. ГОСТ 7.90-2007. Система стандартов по информации, библиотечному и издательскому делу. Универсальная десятичная классификация. Структура, правила ведения и индексирования. 2006. Interstate Council for Standardization, Metrology and Certification (ISC).
- „International UDC Seminars.” UDCC (UDC Consortium). <http://seminar.udcc.org/>. (megtekintve: 2017-11-01)
- Iyer, Hemalata. 2012. *Classificatory Structures : Concepts, Relations and Representation*. Würzburg: Ergon Verlag.
- „Katalist. 2014 januári archívum.” Katalist. <https://listserv.niif.hu/pipermail/katalist/2014-January/>. (megtekintve: 2017-11-01)
- „Last reported editions of the Universal Decimal Classification (UDC).” UDC Consortium (UDCC), legutóbb módosítva 2011. márc. 10-én.
http://www.udcc.org/files/UDCeditions_bibliography10March2011.pdf. (megtekintve: 2017-11-01)
- „Library of Congress Classification.” The Library of Congress.
<https://www.loc.gov/catdir/cpsolcc.html> (megtekintve: 2017-11-01)
- „Major Changes to the UDC 1993-2013.” UDCC (UDC Consortium).
http://www.udcc.org/index.php/site/page?view=major_revisions. (megtekintve: 2017-11-01)
- Makke, Jane. 2016. „Who Needs UDC? Notes from Estonian Libraries.” *IFLA Metadata Newsletter* 2 (2): 34-37. http://www.ifla.org/files/assets/classification-and-indexing/ifla_metadata_newsletter_dec2016.pdf. (megtekintve: 2017-11-01)
- McIlwaine, I. C. 1998. „The Universal Decimal Classification: some factors concerning its origins, development and influence.” In *Historical studies in information science*, szerk. Trudi Bellardo Hahn, és Michael Buckland, 94-106. Medford, NJ: Information Today.

- McIlwaine, I. C. 2006. „The new ecumenism: Exploration of a DDC/UDC view of religion.” *Extensions and Corrections to the UDC* 28: 9-16.
- McIlwaine, I. C., és Nancy Williamson. 2008. „Medicine and the UDC: the process of restructuring Class 61.” *Extensions and Corrections to the UDC* 30: 9-16.
- „Memory of the World. Inscriptions of the documentary heritage in 2013.” UNESCO.
<http://www.unesco.org/new/en/communication-and-information/memory-of-the-world/register/access-by-year/2013/>. (megtekintve: 2017-11-01)
- Nabhan, Yulia. 2010. UDC News in Ukraine. *Extensions and Corrections to the UDC* 32: 31-34.
- Noreikiene, Dalia. 2008. UDC In Lithuania. *Extensions and Corrections to the UDC* 30: 53-54.
<http://hdl.handle.net/10150/110844>. (megtekintve: 2017-11-01)
- Office international de Bibliographic (OIB). 1970. „Rules for Developing the Decimal Classification.” In *International Organization and Dissemination of Knowledge: Selected Essays of Paul Otlet*, ford. és szerk. W. Boyd Rayward, 63-69. Amsterdam: Elsevier, 1990.
- Országos Széchényi Könyvtár Könyvtári Intézet (OSZK KI). 2005. *Egyetemes tizedes osztályozás. 1. kötet Táblázatok 1-2. Rész. UDCC Publication P057*. Budapest: Országos Széchényi Könyvtár Könyvtári Intézet.
- Otlet, Paul. 1990. „On the Structure of Classification Numbers”. In *International Organization and Dissemination of Knowledge: Selected Essays of Paul Otlet*, ford. és szerk. W. Boyd Rayward, 51-61. Amsterdam: Elsevier, 1990.
- Pogányné Rózsa Gabriella. 2007. „Törekvések a világ tudásának egyetemes számbavételére.” *Könyvtári Figyelő* 53 (1): 55-73. <http://ki.oszk.hu/kf/2007/04/torekvesek-a-vilag-tudasanak-egyetemes-szambavetelere/>. (megtekintve: 2017-11-01)
- Ranganathan, S. R. 1967. *Prolegomena to library classification. 3rd ed.* New York: Asia Publishing House. <http://hdl.handle.net/10150/106370>. (megtekintve: 2017-11-01)
- Rayward, W. Boyd. 1994 a. „The International Federation for Information and Documentation (FID).” In *Encyclopedia of Library History*, szerk. Wayne A. Wiegand, és Donald G. Davies, Jr, 290-294. New York: Garland Press. <http://people.ischool.illinois.edu/~wrayward/otlet/FIDHIST2.htm>. (megtekintve: 2017-11-01)
- Rayward, W. Boyd. 1994 b. „Visions of Xanadu: Paul Otlet (1868-1944) and Hypertext.” *Journal of the American Society of Information Science* 45: 235-250.
<http://people.ischool.illinois.edu/~wrayward/otlet/xanadu.htm>. (megtekintve: 2017-11-01)
- Rayward, W. Boyd. 1997. „The Origins of Information Science and the Work of the International Institute of Bibliography/International Federation for Documentation and Information (FID).” *Journal of the American Society for Information Science* 48: 289-300.
<http://people.ischool.illinois.edu/~wrayward/OriginsofInfoScience.pdf>. (megtekintve: 2017-11-01)
- „Rayward’s Otlet Page: Paul Otlet and Documentation.” Rayward, W. Boyd.
<http://people.ischool.illinois.edu/~wrayward/otlet/otletpage.htm>. (megtekintve: 2017-11-01)
- „Referativni Zhurnal Classification Scheme.” legutóbb módosítva 2011. június 1-én.
<http://www.mathontheweb.org/mathweb/Classif/RZhClassification.html>. (megtekintve: 2017-11-01)

- Rigby, Malcolm. 1971. „The role of the UDC in automated information and data systems.” *Proceedings of the Second Seminar on UDC in Mechanized Information Systems conducted by R. R. Freeman, Frankfurt, 1st-5th June, 1970. (FID/CR Report no 11)*, szerk. R. Molgaard-Hansen, és M. Westring-Nielsen, 182-192. Köpenhága: Danish Centre for Documentation, 1971.
- San Segundo Manuel, Rosa. 2007. „Use of the UDC in Spain: implementation, application, teaching and research.” *Extensions and Corrections to the UDC* 29: 285-296.
- San Segundo Manuel, Rosa. 2008. „Some arguments against the suitability of Library of Congress Classification for Spanish Libraries.” *Extensions and Corrections to the UDC* 30: 35-44.
- Satija, Mohinder. 2008. „Universal Decimal Classification: past and present: a state of the art.” *DESIDOC Journal of Library and Information Technology* 28 (6): 3-10.
<http://publications.drdo.gov.in/ojs/index.php/djlit/article/view/217/120>. (megtekintve: 2017-11-01)
- Shera, Jesse Hauk. 1970. „Könyvtárosság, dokumentáció, tájékoztatástudomány.” *Könyvtári Figyelő* 16 (3): 222-229.
http://epa.oszk.hu/00100/00143/00199/pdf/EPA00143_konyvtari_figyelo_1970_03_222-229.pdf. (megtekintve: 2017-11-01)
- Shimmon, Ross. 2001. „IFLA & FID - an Open Letter.” *inCite: News magazine of the Australian Library and Information Association* 22 (8): 6.
<http://www.austlii.edu.au/au/journals/inCiteALIA/2001/217.html>. (megtekintve: 2017-11-01)
- Singh, K. P. 2008. „UDC and its use: a case study of libraries and information centres in Delhi.” *Extensions & Corrections to the UDC* 30: 59-66.
<http://arizona.openrepository.com/arizona/handle/10150/110843>. (megtekintve: 2017-11-01)
- Slavić, Aida. 2004. „UDC Translations: a 2004 Survey Report and Bibliography.” *Extensions & Corrections to the UDC* 26: 58-80. <http://hdl.handle.net/10150/106363>. (megtekintve: 2017-11-01)
- Slavić-Overfield, Aida. 2005. „Classification Management and use in a Networked Environment: The Case of the Universal Decimal Classification.” PhD diss., London: University College London.
http://discovery.ucl.ac.uk/1334914/1/AidaSlavićOverfield_thesis_UCL2005.pdf. (megtekintve: 2017-11-01)
- Slavić, Aida, Jiri Pika, Gerhard Riesthuis, és Chris Overfield. 2007. „German UDC Translation Project.” *Extensions and Corrections to the UDC* 29: 11-15.
- Slavić, Aida, Maria Inês Cordeiro, és Gerhard Riesthuis. 2007. „Enhancement of UDC data for use and sharing in a networked environment.” Cikk a „The 31st Annual Conference of the German Classification Society on Data Analysis, Machine Learning, and Applications. March 7-9, 2007, Freiburg i. Br., Germany” konferencián elhangzott előadás alapján.
<http://arizona.openrepository.com/arizona/handle/10150/106330>. (megtekintve: 2017-11-01)
- Slavić, Aida. 2008. „Use of the Universal Decimal Classification: A worldwide survey.” *Journal of Documentation* 64 (2): 211-228.
- Slavić, Aida, Maria Inês Cordeiro, és Gerhard Riesthuis. 2008. „Maintenance of the Universal Decimal Classification: Overview of the Past and Preparations for the Future.” *International Cataloging and Bibliographic Control* 37 (2): 23-29. <http://hdl.handle.net/10150/105220>. (megtekintve: 2017-11-01)

- Slavić, Aida. 2012a. „FID & UDC Archive News.” *Universal Decimal Classification* (blog), *UDC Consortium*, 2012. október 8. <http://universaldecimalclassification.blogspot.hu/2012/10/fid-udc-archive-news.html>. (megtekintve: 2017-11-01)
- Slavić, Aida. 2012b. „UDC libraries in the world - 2012 study.” *Universal Decimal Classification* (blog), *UDC Consortium*, 2012. augusztus 20. <http://universaldecimalclassification.blogspot.hu/2012/08/udc-libraries-in-world-2012-study.html>. (megtekintve: 2017-01-11)
- Slavić, Aida, és Sylvie Davies. 2017. „Facet Analysis in UDC: questions of structure, functionality and data formality.” In *Faceted classification today: theory, technology and end users: proceedings of the International UDC Seminar 2017, London (UK), 14-15 September*, szerk. Aida Slavić, és Claudio Gnoli, 181-200. Würzburg: Ergon Verlag.
- „The Nobel Peace Prize 1913.” Nobel Media AB. http://www.nobelprize.org/nobel_prizes/peace/laureates/1913/. (megtekintve: 2017-11-01)
- „UDC Consortium.” UDCC (UDC Consortium). http://udcc.org/index.php/site/page?view=about_udcc. (megtekintve: 2017-11-01)
- „UDC Editions.” UDCC (UDC Consortium). <http://www.udcc.org/index.php/site/page?view=editions>. (megtekintve: 2017-01-01)
- „UDC Editorial Team & Advisory Board.” UDCC (UDC Consortium). http://udcc.org/index.php/site/page?view=editorial_team. (megtekintve: 2017-11-01)
- „UDC Licences.” UDCC (UDC Consortium). <http://udcc.org/index.php/site/page?view=licences>. (megtekintve: 2017-01-01)
- „UDC Master Reference File (MRF).” UDCC (UDC Consortium). <http://www.udcc.org/index.php/site/page?view=mrf>. (megtekintve: 2017-11-01)
- „UDC Online.” UDCC (UDC Consortium). <http://www.udc-hub.com/>. (megtekintve: 2017-11-01)
- „UDC Online English – Prices.” UDCC (UDC Consortium). <http://www.udc-hub.com/en/prices.php>. (megtekintve: 2017-11-01)
- „UDC Summary. Multilingual Universal Decimal Classification Summary. UDCC Publication No. 088.” UDCC (UDC Consortium). <http://www.udcc.org/udcsummary/php/index.php>. (megtekintve: 2017-11-01)
- „UDC Summary Translations.” UDCC (UDC Consortium). <http://www.udcsummary.info/translation.htm>. (megtekintve: 2017-11-01)
- „UDC Training Service.” UDCC (UDC Consortium). http://www.udcc.org/index.php/site/page?view=training_service. (megtekintve: 2017-11-01)
- „UIA’s History.” Union of International Associations (UIA). <http://www.uia.org/history>. (megtekintve: 2017-01-01)
- Ungváry Rudolf. 2013. „A Relex névtérkezelő.” (előadás) Workshop 2013 konferencia, Sopron, 2013. március 26-28. <https://niif.videotorium.hu/hu/recordings/6195/relex-nevterkezo>. (megtekintve: 2017-11-01)
- „Universal Bibliographic Repertory.” *Mundaneum*. <http://archives.mundaneum.org/en/universal-bibliographic-repertory>. (megtekintve: 2017-11-01)

„VIAF: The Virtual International Authority File.” OCLC. <https://viaf.org/>. (megtekintve: 2017-11-01)

Wellisch, Hans. 1971. „Reorganization of the UDC.” *Nachrichten für Dokumentation* 22 (2): 55-63.

3 Az ETO mint analitikus-szintetikus osztályozási rendszer

„[az analitikus-szintetikus osztályozási rendszerek] rugalmasságot biztosítanak, amennyiben módot adnak a specifikus fogalmak létrehozására és nem korlátozzák az osztályozót az elérhető fogalmakra. Ugyanakkor komplexsége is teszi a munkáját, mivel a jelzeteket meg kell konstruálnia és nem elég kiválasztania egyet egy listából.”¹⁷³ (Dipti Gulati)

3.1 Az ETO jellege

Bár az Egyetemes Tizedes Osztályozást az első analitikus-szintetikus osztályozási rendszernek szokták tekinteni, valójában nem teljesen előzmények nélküli, hiszen a párhuzamos továbbosztások és a segédtablázatok már korábban is megjelentek például Dewey és Cutter munkáiban. Viszont együttes alkalmazásuk, valamint a szabad kombinációk és fázis kapcsolatok kifejezési lehetőségének rendszerbe illesztése a maga korában teljesen újszerűnek számított.¹⁷⁴ Ez elsősorban annak köszönhető, hogy – ahogyan azt már korábban említettük – az ETO volt az első, nem kifejezetten könyvtári (tehát elsősorban a dokumentumok elrendezésére szolgáló), hanem dokumentációs (a több szempontú visszakeresés biztosításának céljára szánt) osztályozási rendszer.¹⁷⁵

„Az Egyetemes Tizedes Osztályozás (ETO) alapelvei, revíziójának és kiadásának szabályai” között az alábbiakat olvashatjuk:¹⁷⁶

„Az ETO hierarchikus rendszer, vagyis minden egyes osztálya tovább bontható. A jelzetek terjedelme tárgykörük specifikusságával arányos. Szerkezeti felépítése 'decimális': azaz minden egyes csoport az ismeretek egymást folyamatosan követő felosztásával, a legáltalánosabbtól a legrészletezőbbig a tízes számrendszer alapján épül fel (...).

Az ETO lehetővé teszi egy tárgykör felosztását egymást követően több jellemző ismerv (facetta) szerint is.

Az ETO nézőpontok (aspektusok) szerinti osztályozás, minthogy benne egy adott fogalom különböző szemszögből tekintve, vagy különböző tudományterületek oldaláról nézve többször is előfordulhat.

Az ETO szintetikus osztályozás. A komplex fogalmak különféle összetett jelzetekkel jelölhetők.”

Az ETO tehát elsősorban egy mély fogalmi hierarchiát felépítő táblázati részt tartalmaz, mely az emberi ismeretek és tevékenységek (nem csak a tudományok) egészére kiterjed.

Az alapelv, hogy mindig a legegyszerűbb alakra törekedjünk, így, ha egy fogalomra létezik táblázati szám, azt kell használni a leírására (a szinonimák elkerülése érdekében).¹⁷⁷ Amennyiben nem létezik, illetve a dokumentum több szempontú leírására van szükség, akkor az általánosan és korlátozottan közös *alosztások segítségével* nyílik mód prekoordinált jelzetek létrehozására. Jelenleg, ahogy korábban már láthattuk, a rendszer fejlesztése éppen a facettás, illetve szintetizáló jelleg erősítése

¹⁷³ Gulati 2013: 151.

¹⁷⁴ de Groelier 1962: 17-20.

¹⁷⁵ vö. Slavić és Davies 2017.

¹⁷⁶ FID 1983.

¹⁷⁷ OSZK KI 2005: 7.

irányába halad, azaz a valamilyen szempontból ismétlődő hierarchikus továbbosztásokat igyekeznek általánosan és korlátozottan közös alosztások használatával kiváltani.

3.2 Az ETO mint hierarchikus rendszer

Amikor Melvil Dewey 1876-ban megjelentette könyvtári célra (tehát dokumentumok rendszerezésére) kidolgozott egyetemes osztályozási rendszerét, a rendszer újdonságait elsősorban az ismeretek *hierarchikus, a tizes számrendszeren alapuló elrendezése* és a *betűrendes tárgymutató* jelentették.¹⁷⁸ Rendszerében például az 500 jelenti a természettudományokat, az 510 a matematikát, ezen belül az 513 a számelméletet, 513.5 a számrendszereket stb. Ezzel egy egyszerű, kevés jelet használó, könnyen áttekinthető struktúrát sikerült megalkotnia, melyben a fogalmakat azonosító jelzetek egyben a köztük fennálló hierarchikus kapcsolatokat is kifejezik.

Mikor Paul Otlet továbbfejlesztette Dewey rendszerét, megtartotta annak főosztályait és osztályait, de az olvashatóság érdekében eltörölte a három számjegyű minimum elvét (így változott a fenti példa esetén a természettudomány főosztály jele 500-ról 5-re, a matematikáé pedig 510-ról 51-re).¹⁷⁹ A főosztályok a két rendszerben a mai napig megegyeznek, egyetlen kivétellel: *a négyes főosztály az ETO-ban jelenleg betöltetlen*, amióta a nyelvészet átkerült a nyolcas főosztályba.¹⁸⁰

A jelenlegi táblázatok leghosszabb jelzetei az elektrotechnika alatt találhatóak, ezek egyikén jól demonstrálható a tizedes alosztás elve.¹⁸¹

6 Alkalmazott tudományok. Orvostudomány. Műszaki tudományok
62 Mérnöki tudományok. Technika általában
621 Általános gépészet. Magtechnika. Nukleáris technika. Elektrotechnika.
Mechanikai technológia általában
621.3 Elektrotechnika
621.39 Híradástechnika. Távközlés, távírótechnika, távbeszélő-technika.
Rádiótechnika. Videotechnika. Távvezérlés
621.397 Videotechnika, televíziótechnika. Képrögzítés, továbbítás és lejátszás.
Készülékek és hálózatok
621.397.13 Televíziótechnika
621.397.132 Színes televízió
621.397.132.1 Szimultán rendszerek
621.397.132.12 Közös átviteli csatornával az alapszínjelekhez
621.397.132.121 NTSC-rendszer (US National Television Systems Committee)

Ilyen mélység csak a táblázatok egy-egy helyén jellemző, és azok *egyenetlenségéből* adódik, ami gyakran felmerülő kritika az ETO-val szemben.¹⁸² Ahogy már korábban említettük, *a jelenlegi fejlesztések éppen a hierarchikus mélység csökkentésére, a párhuzamosságok kiküszöbölésére, a több szempontú visszakereshetőség erősítésére törekednek.*

A táblázatokban előfordulnak olyan esetek, amikor egy osztálynak nem a nála egy decimális karakterrel rövidebb jelzettel jelölt osztály, hanem az azt tartalmazó *aggregált osztály a hierarchikus felső kapcsolata*. Például a 502.5 (földfelszín) felső kapcsolata nem az 502 (Természetvédelem),

¹⁷⁸ Babiczky 1975: 11.

¹⁷⁹ Horváth és Sütő 2001: 66.

¹⁸⁰ A négyes főosztály felszabadítása az 1960-as évek közepén történt. Betöltésére azóta több elképzelés is született, jelenleg legvalószínűbb, hogy az orvostudomány lesz ide áthelyezve (Benito 2007, McIlwaine és Williamson 2008, Pestana 2015).

¹⁸¹ A harmadik rész példái a 2005. évi magyar kiadásból származnak (OSZK KI 2005), de valamennyi idézett osztály szerepel az angol nyelvű UDC Online-ban is („UDC Online”).

¹⁸² Az egyenetlenséget nagyon jól mutatja a tárgykörök lefedettségének grafikonja is, amely az ETO Konzorcium honlapján tekinthető meg („UDC Subject Coverage”).

hanem az az alatt található 502.3/.7 (A környezet részei) kiterjesztés, vagy az (571) földrajzi alosztásé (Az Orosz Föderáció ázsiai része) a (470+571) összekötés (Orosz Föderáció). Ilyen esetekben a jelzet nem tükrözi a valós hierarchiát a megszokott formában. (Az utóbbi példából pedig az is látszik, hogy nem csak a fő-, hanem a segédtablázatok – jelen esetben a Földrajzi hely szerinti általánosan közös alosztások – is hierarchikus rendszerben tartalmazzák a fogalmakat.)

A jelzetben nem kifejeződő, tehát *rejtett hierarchia* tipikus esete, amikor a tízes osztás nem elegendő az alosztályok kifejezésére, ezért százas elven osztjuk tovább az aktuális osztályt. De hasonló problémát okozhat például a jelzetek rövidítése érdekében alkalmazott ún. *teleszkópos hierarchia*, amikor az osztály és a leszármazottai azonos szinten jelennek meg. *Ugró hierarchiáról* akkor beszélünk, ha a hierarchia szintjei közül egy vagy több kimarad, tehát pl. egy három számjeggyel jelölt osztálynak közvetlen leszármazottja egy öt számjeggyel jelölt. Szintén tipikus probléma a *hamis hierarchia*, azaz amikor a jelzetek nem tökéletesen tükrözik a kívánt hierarchikus kapcsolatokat. Ennek legjellemzőbb példái a földrajzi alosztásokban a történelmi és geopolitikai változások miatt kialakuló következetlenségek.¹⁸³

Gyakori bírálat tárgya az *ETO elavult tudományfelosztása*, ami természetes is, hiszen egy száznegyven évvel ezelőtt kialakított hierarchiáról van szó. Ugyanakkor egy a Wikipedia kategóriákat és az ETO osztályokat összehasonlító közelmúltbeli kutatás azt mutatta ki, hogy szinte valamennyi Wikipedia kategória első vagy második szinten van az ETO-ban reprezentálva¹⁸⁴. Ez arra utal, hogy egy modern, teljesen gyakorlati jellegű kategorizálás sem mutat gyökeres eltérést az ETO alapfelosztásától. Másrészt ugyanez a példa megerősíti a régi kritikát is, hogy az erőltetett *tízes felosztás a hierarchiaszintek eltolódásához vezet*¹⁸⁵ (hiszen a kategóriák az ETO-ban különböző hierarchiaszinteken jelennek meg), aminek a következménye a főosztályok fentebb látott egyenetlensége is.

A *monohierarchikus megközelítés* azt jelenti, hogy választani kell a lehetséges felosztási szempontok közül. Ennek következtében a felosztások egymásutánja alárendeltséget eredményez független szempontok esetén is.¹⁸⁶ Például a matematika alatt először a matematika ágai alapján történik a felosztás, így kapjuk az 517.9 (Differenciál egyenletek, integrál egyenletek stb.) jelzetet az 517 (analízis) alatt és az 519.6-ot (numerikus módszerek), majd mindkettő alatt megtaláljuk a közönséges (517.922/.929 és 519.62), illetve a parciális differenciálegyenletekre (517.95 és 519.63) vonatkozó jelzeteket. Hasonló probléma merülhet fel az új, hagyományos diszciplínákon átívelő tudományterületek esetében¹⁸⁷. De az ETO-ban az „arany” kifejezés is hasonlóan jelenik meg: a nemesfémekkel kapcsolatos teljes irodalom összegyűjtéséhez legalább nyolc helyet kell figyelembe venni a táblázatokban, attól függően, hogy a téma a pénzérmék (336.743.22), a szervesetlen kémia (546.59), az ásványtan (549.283), a bányászat (622.342) stb. szempontjából van tárgyalva.¹⁸⁸

¹⁸³ Slavić 2008.

¹⁸⁴ Salah et. al. 2011.

¹⁸⁵ Csík 1995

¹⁸⁶ Csík 1995.

¹⁸⁷ A leggyakrabban idézett példa a genderelmélet, de említhető a fraktálok elmélete is (Barátné 2011).

¹⁸⁸ Ezekre a problémákra a rendszer facettás, illetve szintetizáló formában történő átszervezése tűnik megoldásnak (ami illeszkedik a jelenlegi fejlesztési tendenciákra), bár ennek is megvannak a hátrányai (Buxton 2011).

Amennyiben a párhuzamosan kialakuló független hierarchiák szerkezete teljesen megegyezik, akkor nem szükséges minden esetben felsorolnunk a leszármaztatott osztályokat, elég csak utalni arra a helyre, ahol az enumerikus felsorolás megtalálható.¹⁸⁹ Például a Galaktikus ködök hierarchiája az alábbi:

524.5 *Intersztelláris anyag. Galaktikus ködök*
.52 *Intersztelláris gáz*
.522 *H I tartományok*
.523 *H II tartományok*
.527 *Intersztelláris molekulák*
.54 *Emissziós ködök*
.57 *Intersztelláris por. Sötét ködök*

Az intergalaktikus anyag (524.78) jelzeténél nem találjuk meg ugyanezt a felsorolást, csupán egy utalást annak helyére ($\cong$ 524.5). Ezt a megoldást – amelyet **párhuzamos továbbosztásnak** nevezünk – már Dewey is alkalmazta, a táblázatok rövidebbé tétele érdekében.¹⁹⁰

3.3 Az ETO mint facettás rendszer

Korábban már láttuk, hogy a facetták olyan jellemzők, amelyek egy adott osztályon belül tipikusan elő szoktak fordulni és pontosítják az osztályhoz tartozó fogalmakat. Említettük Ranganathan és Vickery szempontrendszereit is, azaz a kategóriákat, melyeket a facetták továbbosztásának alapjaként meghatároztak. Az ETO esetében csak egy *ajánlás* létezik arra, hogy a fogalmakat milyen kategóriák szerint célszerű csoportosítani¹⁹¹, de nincs olyan általános érvényű kategória lista, mint Ranganathan esetében a PMETS, és a korábbi fejlesztésekből is hiányzott ez a típusú szisztematikus megközelítés.¹⁹²

3.3.1 A korlátozottan közös alosztások

A fenti meghatározás alapján az ETO-ban a facetták általában **korlátozottan közös alosztások**ként jelennek meg. Ezek olyan segéd táblázatok, illetve speciális jelzetszerkesztési megoldások, amelyek csak egy bizonyos osztályon belül alkalmazhatók.

A korlátozottan közös alosztások *legtöbbször enumerikus formában jelennek meg*, tehát egyszerűen fel vannak sorolva annak az osztálynak az elején, ahol alkalmazhatók. Például a teljes matematika (51) osztályban a következő korlátozottan közös alosztások alkalmazhatók:

51-3 *Számítási eljárások*
51-32 *Számítás segédeszközök nélkül*
51-33 *Táblázatok használata*
51-35 *Mechanikai, grafikai és más eszközök használata*
51-37 *Számológép, számítógép használata a matematikában*
51-7 *Matematikai tudományok és módszerek alkalmazása más tudományok területén*
51-8 *Matematikai játékok és fejtörők*

A matematikán belül csak a csoportelmélet osztályban (512.54) alkalmazhatók az alábbi alosztások:

512.54
.01 *Csoportok axiomatizálható osztályai*
.03 *Csoportok különböző osztályainak elemi elmélete*
.05 *Algoritmikus problémák a csoportelméletben. A szó-probléma*

¹⁸⁹ FID 1983: 16.

¹⁹⁰ Shenoy 1998: 109.

¹⁹¹ Rendszerek, Teljes entitások, Részek, Anyagok, Tulajdonságok, Tevékenységek, Cselekvők (ágensek), Elvek és elméletek (UDCC 2005).

¹⁹² Gnoli 2011.

Mindkét alosztás alkalmazható az 512.541.2 jelzeten, például az „Abel-csoportok magasabb rendű részcsoportjai osztályainak Cayley-táblázatok segítségével történő vizsgálata” kifejezésére: 512.541.2.03-33.

A fenti példában láthattuk az enumerikus korlátozottan közös alosztások két alaptípusát, a **kötőjellel**, illetve **.0 karaktersorozattal bevezetett** alosztásokat. Utóbbi gyakran visszatérő kritika tárgya, mivel nem minden esetben dönthető el automatikusan, hogy egy ilyen karaktersor a táblázati szám szerves része-e vagy egy facettát vezet be, különösen párhuzamos továbbosztással képzett jelzetek esetében.¹⁹³

Az enumerikus korlátozottan közös alosztások még ennél is nehezebben automatizálható esetét jelentik a **számjelzetes** korlátozottan közös alosztások, melyeket mindenféle bevezető karakter nélkül kell a táblázati szám után írni. Ilyenek találhatók például korábbi kiadásokban a kohászat (669) alatt, ahol pl. a rézkohászatot a 669.3 (réz) és a 669...3 (kohászat) összevonásával 669.33-ként kell kifejezni.¹⁹⁴ Ezeket a típusú alosztásokat a kilencvenes évek revíziói többnyire megszüntették, illetve más megoldásokkal helyettesítették.

A korlátozottan közös alosztások negyedik típusát az **aposztróffal képzett** alosztások jelentik. Az aposztróf eredetileg arra lett bevezetve, hogy lehetővé tegye két, azonos osztály alá tartozó fogalom összekapcsolását, „szintetizálását”. Tehát egy olyan relációról van szó, amely csak egy adott osztályon belül alkalmazható, ezzel inkább az analitikus-szintetikus, mintsem a facettás jellege erősödik a rendszernek.¹⁹⁵ A táblázatok bővítése során azonban az eredeti szándék nem volt következetesen képviselve, és számos olyan alosztás lett bevezetve, amelyben az aposztróf a kötőjelhez és a .0-hoz hasonlóan egyszerű bevezető karakterként van használva, és az eredeti, szintetizáló funkció csak néhány osztályban maradt meg (pl. a szervetlen kémia – 546 – vagy a politikai mozgalmak és pártok – 329.1/.6 – alatt).¹⁹⁶

3.3.2 Nem korlátozottan közös alosztásokkal kifejezett facetták

Van példa nem korlátozottan közös alosztásokkal meghatározott facettákra is. Például az 1994. évi átalakítás óta az egyes földrajzi területek vagy népek történelme *általánosan közös alosztások facettás módon való alkalmazásával* határozható meg:

94 - Általános történelem
(439.165) - Hajdú-Bihar megye

¹⁹³ A problémára már Pauline Atherton és Robert R. Freeman felhívták a figyelmet a hatvanas években (Freeman és Atherton 1967; hiv. Buxton 1990), de a korlátozottan közös alosztások esetén azóta sem érintették a revíziók a pont-nullás alosztásokat (szemben a szempont szerinti alosztásokkal, amelyek törölve lettek). A párhuzamos továbbosztások szisztematikus kivezetése azonban csökkentette az általuk okozott lehetséges szinonimák számát (Piros 2015, Mándy 2011).

¹⁹⁴ OMIKK és OSZK-KMK 1990: 317.

¹⁹⁵ Ezt a jelleget megerősítették a hetvenes évek átszervezései, amelyek lehetővé tették az aposztróffal szintetizált jelzetek permutálását, ezzel implicite kimondva, hogy ezeket a fogalmakat valamennyi komponens szerint visszakéreshetővé kell tenni (Piros 2014, Babiczky 1985).

¹⁹⁶ McIlwaine 1998, Wimmer 2000. Ebben a munkában jelzetszintézisről akkor beszélünk, amikor a művelet a táblázatok ugyanazon pontján álló jelzeteket köti össze, melyek egyenrangúak és a sorrendnek legfeljebb jelentést árnyaló szerepe van. Azok az esetek, amikor az aposztróf a táblázatok más helyéről származó számot vezet be (pl. a 669.35 : (OSZK KI: 2/375) rézötvözetek vagy 547.29 alatt az észterek esetében (OSZK KI: 1/442) olyan enumerikus alosztásoknak tekintendők, melyek forrása a táblázat más helyén található.

Olyan is előfordul, hogy *külső forrásból származó vagy betűjeles alosztások használatával* kell kifejezni egy adott attribútumot, például a programozási nyelvek esetén így adható meg a nyelv (pl. 004.912=93C++). A földrajzi alosztások alatt is hasonló módon lehet meghatározni a koordinátákat (a (160) alatt), illetve a méreteket (a (18) alatt).

A formai, illetve a tulajdonságok szerinti általánosan közös alosztások alatt lehetőség van új fogalom képzésére úgy, hogy egy tetszőleges fő táblázati számot a (0) formai alosztáshoz, illetve a -029 tulajdonságok szerinti alosztáshoz kapcsoljuk *viszonyítással*:

(0:82-31) – *regény forma*
-029:133 – *okkult tulajdonságok*

A korábban kifejtett *párhuzamos alosztások* eredetileg a rendszer hierarchikus jellegéből következnek, és függetlenek a facettáktól. Ha azonban a párhuzamos továbbosztás a fő táblázatok szintjére utal ($\cong 0/9$ formában), akkor ez a gyakorlatban azt jelenti, hogy az osztályt egy tetszőleges másik fő táblázati szám hozzáírásával pontosítjuk. Például korábban így lehetett kifejezni a tantárgyakat a 372.8 vagy a szakfőiskolákat a 378.6 alatt.¹⁹⁷

3.4 Az ETO mint analitikus-szintetikus rendszer

Az ETO analitikus-szintetikus jellegét az **általánosan közös alosztások** biztosítják, melyeknek két nagy csoportja van: az általánosan közös **alosztás táblázatok**, melyekkel a facettákhoz hasonlóan fejezhetők ki az egyes fogalmak attribútumai tetszőleges osztály esetén, és a jelzetek összekapcsolására szolgáló **alosztási jelek**, melyekkel a komplex és az aggregált tárgykörök fejezhetők ki.

3.4.1 Általánosan közös alosztási táblázatok

Ezek a táblázatok olyan *ismétlődő jellemzőket sorolnak fel*, melyek a fő táblázati osztályok által jelölt fogalmak pontosítására felhasználhatók, a fő táblázati és a segéd táblázati jelzet egymás után írásával. Az általánosan közös alosztások *bármelyik fő táblázati számmal használhatók*, mivel olyan általános kategóriák szerinti pontosításra szolgálnak, mint a hely, a szöveg nyelve, fizikai megjelenése stb.

Mivel az *alosztási számok jelzetén belüli sorrendjére semmilyen merev megkötés nincs*¹⁹⁸ – akár a fő táblázati számok előtt, vagy azok számjegyei közé ékelve is állhatnak – fontosak a **bevezető (esetleg záró) karakterek**, amelyek biztosítják a jelzetelemek elkülöníthetőségét.

A 1. számú táblázat az ETO általánosan közös alosztási táblázatait tartalmazza.¹⁹⁹

Táblázat azonosítója	Megnevezése	Jele (bevezető és záró karakter)
1c táblázat	Nyelvi általánosan közös alosztások	=...

¹⁹⁷ OMIKK és OSZK-KMK 1990: 133-134.

¹⁹⁸ A jelzetek besorolási sorrendjével szemben a jelzetelemek sorrendjére nem létezik központi megállapított szabály. A különböző ETO kiadások előszavai és az osztályozási kézikönyvek általában egy a besorolási sorrend megfordításán alapuló sorrendet javasolnak, az azonban csak ajánlás, melytől a leírás jelentésének finomítása, illetve a gyűjtemény vagy a katalógus speciális besorolási gyakorlata érdekében el lehet térni (Robinson 2003).

¹⁹⁹ A táblázatok forrása a UDC Summary magyar nyelvű fordítása („UDC Summary 2011. Hungarian (Magyar).”).

1d táblázat	Formai általánosan közös alosztások	(0...)
1e táblázat	Hely szerinti (földrajzi) általánosan közös alosztások	(...)
1f táblázat	Az emberek származása, etnikai csoportok és nemzetiség szerinti (etnikai) általánosan közös alosztások	(=...)
1g táblázat	Idő szerinti általánosan közös alosztások	"..."
1h táblázat	Külső forrásból vett jelzetek és A/Z névalosztások	*... és A/Z
1i táblázat	Szempontra szerinti nem önálló általánosan közös alosztások ²⁰⁰	.00...
1k táblázat	Általános ismérvek szerinti nem önálló általánosan közös alosztások	-0...

1. táblázat Az ETO általánosan közös alosztásai

Az **önálló általánosan közös alosztások** a fő táblázati jelzetekkel egyenértékű elemként használhatók, akár önállóan is szerepelhetnek a jelzetekben, illetve az összetett jelzet élére is kerülhetnek, ha a téma leírása ezt kívánja meg. Ezzel szemben a **nem önálló általánosan közös alosztások** jelzetei csak fő táblázati számokhoz kapcsolódva, azok után, az értelmüket módosítva használhatók.

Az általánosan közös alosztások nem csak táblázati számokat és azok kiterjesztéseit tartalmazhatják, hanem további általános²⁰¹ és korlátozottan közös alosztásokat, illetve alosztási jeleket²⁰² is. Ezek használata a táblázatok megfelelő helyein van szabályozva.

Lentebb néhány példát láthatunk az általánosan közös alosztások használatára.

634.8-057.117"199" (450) (086.8)=131.1 Kilencvenes évekbeli olaszországi szőlőültetvényeken dolgozó munkásokról szóló videofelvétel

347.6-055.62-055.2"18" (410) A lányági öröklődésre vonatkozó angliai törvények a XIX. században

„2013"591.5 (043)=161.1 Állatok viselkedéséről szóló 2013-ban írott doktori disszertáció orosz nyelven

628.3 (410.111) "18" Szennyvízfeldolgozás Londonban a XIX. században²⁰³

²⁰⁰ A szempont szerinti általánosan közös alosztások 1999-ben és 2001-ben törölve lettek („Cancellations to the UDC”).

²⁰¹ Ilyenre példa az etnikai övezetek kifejezése a földrajzi alosztásba ágyazott etnikai alosztással, (például a (494=112.2) jelöli Német Svájcot (OSZK KI 2005: 1/34))

²⁰² A legismertebb példák az országok közötti kapcsolatok kifejezésére szolgáló viszonyítás egy földrajzi alosztáson belül (pl. 339.5(44:450) jelöli a kereskedelmi kapcsolatokat Francia- és Olaszország között (OSZK KI 2005: 1/34)), illetve a viszonyítás használata új formai alosztás képzésére létező fő táblázati számból (pl. 929(0:82-31) – Életrajz regény formában (OSZK KI 2005: 1/26)), de számos más esetben is használhatók alosztási jelek alosztásokon belül.

²⁰³ A példák forrása Batley (2014: 114-118).

338.5(410.111.11) Árak London belvárosában
 51(=14) ".../-1" (091) A hellén matematika története az Kr. e. I. századig
 622(410).33"195"(01) Bibliográfia Nagy-Britannia szénbányászatáról az ötvenes években
 741.7-026.613.24-023.3*8 Narancssárga nyolcszögek kivágása ollóval
 94(399.7=822.111) "14/15" (043.3)=134.2 Disszertáció az inkák XV/XVI. századbeli történelméről, spanyol nyelven

A külső forrásból vett jelzetek és A/Z névalosztások – főleg a jelzetek automatikus feldolgozása esetén – felvetik a kérdést, hogy az alosztást követő zárójelek, plusz- és perjelek, kettőspontok stb. mikor tartoznak az alosztáshoz, és mikor vezetnek be új jelzetelemet. Erre az lenne a megoldás, ha ezeknek a lehetőségeknek is biterminális jeleik lennének.²⁰⁴ Ennek hiányában két megközelítés lehetséges. Az egyik szerint A/Z után már minden a külső jelzet része.²⁰⁵ A másik szerint az ismert szeparátor karaktereket akkor is feldolgozzuk, ha előttük külső forrásból származó alosztás állt. Az utóbbi mellett szól, hogy a táblázatok tartalmazzanak olyan szabványos jelzetépítési megoldásokat, amelyek lehetővé, sőt szükségessé teszik, hogy a külső forrásból származó alosztásokat más alosztások kövessék.²⁰⁶

3.4.2 Jelzetek összekapcsolására szolgáló általánosan közös alosztási jelek

A táblázati számok, facetták és általánosan közös alosztások segítségével kifejezett egyszerű és összetett fogalmakból a jelzetek összekapcsolására szolgáló általánosan közös alosztási jelekkel lehet agglomerált és összetett tárgyköröket készíteni. Az alosztási jelek száma az évek során bővült, esetenként a szerepük is változott. Jelenleg az alábbi kapcsolati módok használatára van lehetőség:²⁰⁷

1a táblázat Összekötő jelek

- Koordináció (pluszjel, +)
- Folyamatos kiterjesztés (perjel, /)

1b táblázat Relációs jelek

- Mellérendelő viszonyítás (kettőspont, :)
- Sorrendrögzítés (dupla kettőspont, ::)
- Csoportképzés (szögletes zárójel, []).

3.4.2.1 Összekötés és kiterjesztés

Az összekötés és kiterjesztés az *agglomeráció* kifejezésére szolgálnak. A kiterjesztést a táblázatokban egymás mellett álló tárgykörök összefoglalására lehet alkalmazni, míg az összekötés nem egymás mellett álló tárgyköröket is összekapcsolhat.²⁰⁸

²⁰⁴ „A biterminal symbol for non-UDC notation”.

²⁰⁵ Csabay 1999.

²⁰⁶ Például az irodalom osztályban a szerzőt, a mű címét és a kiadásfajtaát, fordításokat felváltva szerepeltetett betűjeles és korlátozottan közös alosztásokkal kell jelölni (pl. 821.133.1MOL7TART.03=511.141 jelöli Molière Tartuffe-jének magyar fordítását) (ld. OSZK KI 2005: 2/529), vagy térbeli méretek jelölésekor is megelőzi a mértékegység jelölése a mennyiséget (181km427) (OSZK KI 2005: 1/39). De említhetjük a programozási nyelvek kifejezését is, amely a =93 alatt, A/Z alosztással történik (OSZK KI 2005: 1/25), vagy a földrajzi nevek névalosztással történő kifejezését.

²⁰⁷ „UDC Online.”

²⁰⁸ Chatterjee 2015.

Ez a két alosztási jel *bővíti az egyes osztályok tartalmát* más, kapcsolódó osztályok hozzáadásával, ezáltal kiterjeszti a táblázati számok jelentését. A plusz- vagy perjellel összekapcsolt számok tágabb fogalmakat jelölnek, mint az egyszerű számok, például a 73+75 (szobrászat és festészet), illetve a 73/75 (szobrászattól a festészetig, beleértve a 74-es osztályt [„Rajzművészet. Formatervezés. Alkalmazott művészet, iparművészet.”] bővebb fogalmat jelöl, mint a 73 (szobrászat). Mivel az ETO besorolási szabályai az általánostól a specifikus felé haladás elvét követik, először az összekötéssel, majd a kiterjesztéssel alkotott számokat kell besorolni, ezeket követhetik az egyszerű főtáblázati számok.²⁰⁹

Teljes dokumentumok leírásakor az **összekötés** leggyakrabban *témák egymás melletti, független előfordulását* fejezi ki, ezért – a részjelzetek visszakereshetőségét megkönnyítendő – a gyakorlatban elterjedt az a megoldás, hogy a pluszjel használata helyett a részjelzetek külön-külön kerülnek rögzítésre.²¹⁰

Ez talán a visszakeresés szempontjából kielégítő megoldás, de ellentmond annak a korábbi fejezetekben már említett alapelvnek, hogy a jelzetnek tükröznie kell a leírt szöveg fogalmi viszonyait a relevanciadöntések hatékony támogatása érdekében.²¹¹ Authority control alkalmazása esetén pedig természetes, hogy az összetett és a részjelzeteknek külön rekordokat kell létrehozni.

A **kiterjesztés** azonos hierarchiaszinten, *közvetlenül egymás mellett álló számok összekapcsolására* szolgál. Ezért úgy is szokás kezelni, mint az egymás mellett álló számok összekötésének rövidített kifejezését. Ha azonban teljesen pontosak akarunk lenni, akkor – mivel a kiterjesztéshez csak egységes egészként kapcsolódhatnak további alosztások, az egyes elemeihez nem – egy szögletes zárójelben álló összekötésként kell rá tekintenünk. Például a 26/28-532.2 (zsidó, keresztény és mohamedán csoportos szertartások) helyes felbontása nem 26+27+28-532.2, hanem [26+27+28]-532.2.

Az elsődleges probléma a kiterjesztéssel kapcsolatban az, hogy az ETO-jelzeteket egységes egészként kezelő *programok nem tudják kezelni az intervallum belsejében álló számokat*, mivel azok semmilyen formában nem jelennek meg a jelzetben.²¹² Ezért a nyolcvanas évek elején az is felmerült, hogy törlik ezt a műveletet.²¹³ A legújabb magyar nyelvű nyomtatott ETO-kiadás is a mellőzését ajánlja (helyette az összekötés használatát javasolva), amennyiben minden számot szeretnénk visszakereshetővé tenni.²¹⁴

Azonban vannak olyan intervallumok, amelyek a táblázatokban is egységként jelennek meg (tehát maga az intervallum is szerepel a táblázatokban). Ezekben az esetekben, illetve amennyiben a jelzetek dokumentumok polcon történő besorolására szolgálnak, ragaszkodni kell az intervallumok használatához.²¹⁵ Egyéb esetekben az a nézet terjedt el, hogy célszerűbb lebontani az intervallumot az

²⁰⁹ Id. Table 1a („UDC Online”).

²¹⁰ Babiczky 1985, Pálvölgyi 1998: 94.

²¹¹ Piros 2014.

²¹² Piros 2014.

²¹³ Babiczky 1985.

²¹⁴ OSZK KI 2005: 1/11.

²¹⁵ „Should the slash be abolished?”

összetevőire. Ez azonban szintén csökkentheti a jelzet mint szurrogátum megfelelőségét. Helyesebbnek tűnik megtartani az intervallumokat, és a programokat felkészíteni azok kezelésére.²¹⁶

3.4.2.2 Viszonyítás és sorrendrögzítés

Az ETO *legfontosabb alosztási jele* a komplex tárgykörök leírására használható **viszonyítás** - a rendszer megjelenésekor ez számított a legnagyobb újításnak.

Jelenlegi meghatározása szerint a kettőspont „*kapcsolatot fejez ki két vagy több tárgykör között az ETO számuk összekapcsolásával. Szemben a plusz- és a perjellel, a kettőspont nem kiterjeszti, hanem pontosítja az összekapcsolt tárgyköröket.*”²¹⁷ Mivel a kapcsolat jellegére nincs utalás, a viszonyítást gyakran feleltetik meg a Boole-algebra „és” műveletének – némileg pontatlanul, mivel ezzel figyelmen kívül hagyjuk a relációt magát, amely szintén befolyásolja a tárgykör jelentését.

Az ötvenes és hatvanas években, amikor a fogalmi kapcsolatok vizsgálata állt a kutatások középpontjában, többen is kritizálták a rendszert, amiért nem ad lehetőséget a fáziskapcsolatok típusának részletesebb kifejtésére. Több különböző javaslat is született²¹⁸, végül azonban ezek nem kerültek elfogadásra, és a kapcsolatok kettőspontos kifejezése önmagában továbbra sem utal a hatás irányára, sem a jellegére.

A hatvanas években került bevezetésre a dupla kettőspont a **sorrendrögzítés**, azaz az *irányított kapcsolat jelölésére*.²¹⁹ A hagyományos cédulakatalógusokban ez úgy értelmezhető, hogy csak az egyik fogalom alapján szükséges besorolni a jelzetet, a számítógépes programok azonban már azt is lehetővé tehetik, hogy az összekapcsolt fogalmak, illetve az összetett fogalom alapján, de a sorrend figyelembe vételével is visszakereshető legyen.

A viszonyításra és a sorrendrögzítésre a UDC Online-on az alábbi példákat találhatjuk:²²⁰

17:7 Etika a művészetrel kapcsolatban
341.63(44:430) Nemzetközi egyeztetési eljárások Franciaország és Németország között
37-042.3:32 A politika befolyása az oktatásra
37-042.4:004 Számítógépek alkalmazása az oktatásban
628.463:692.758 Szeméttledobó aknákat használó hulladékgyűjtő rendszerek
061.2(100)::[54+66] IUPAC International Union of Pure and Applied Chemistry
77.044::355.4 Háborús fotózás

²¹⁶ Buxton (1990) olyan algoritmus használatát javasolja, amely az intervallumot automatikusan felbontva teszi azt visszakereshetővé a részei alapján, Riesthuis (1997, 1999) pedig példaprogramot is publikált a kiterjesztések felbontására. Ennek a megoldásnak egy alternatívája a kiterjesztés intervallumként való ábrázolása, amely egyébként a relációk szempontjából az egyszerű számokkal azonos módon viselkedik. Így, ha szükséges, az összekötés később felbontható, de a jelzet eredeti formájában is reprodukálható marad. Ha az intervallum végpontjai alapján indexelünk, a „között” operátorral a hierarchiát is követő, tehát gyorsabb keresés is megvalósítható (Piros 2014). Végül, de nem utolsó sorban az utóbbi megoldást támogatják a MARC21 és UNIMARC osztályozási formátumai (LOC 2000 és IFLA 2000).

²¹⁷ Id. Simple relation („UDC Online”).

²¹⁸ Ranganathan 1956, de Grolier 1962: 20-21, illetve Perrault 2001.

²¹⁹ Babiczky 1985.

²²⁰ „UDC Online.”

A közelmúltban a fázisnak, tehát a kapcsolat jellegének jelölésére is született megoldás: 2003. óta - a -042 alatti általánosan közös alosztások használhatók erre a célra.²²¹ Például a 33:32 („a politika és a gazdaság kapcsolata”) jelzet a következő formában finomítható tovább: 33-042.3:32 („a politika hatása a gazdaságra”).²²²

3.4.2.3 Csoportképzés

A szögletes zárójellel jelölt **csoportképzés** egy a komplex tárgykörök kifejezésében nagyon hasznos, ugyanakkor meglehetősen *ellentmondásos* alosztási jel az ETO-n belül.

Eredetileg *egy jelzetnek egy másikba való közbeékelésére* lehetett használni, ha a visszakereshetőséget csak egy irányból volt szükséges biztosítani. Magyarországon ez a felhasználás terjedt el, és még ma is használják ezen a módon, annak ellenére, hogy az alosztásra vonatkozó szabályok még a hatvanas évek végén megváltoztak, és a szögletes zárójel régi funkcióját a sorrendrögzítés vette át.²²³

A csoportképzés *jelenlegi funkciója a csoportosítás „az ETO számok többszörös összetétele esetén”, „a tárgykörök közötti viszonyok egyértelmű meghatározására”*²²⁴: „képzésére akkor lehet szükség, amikor a pluszjellel, törtvonallal vagy kettősponttal összekapcsolt ETO számokkal jelölt összetett tárgy (fogalom) mint egész kapcsolódik kettősponttal egy másik számhoz, vagy mint egészet módosítja egy általánosan vagy korlátozottan közös alosztás.”²²⁵

A korlátozottan közös alosztás természetesen csak abban az esetben értelmezhető, ha a zárójel több, alosztási jelekkel összekapcsolt, a táblázatban azonos vagy egymás alatti hierarchiaszinten található számot tartalmaz úgy, hogy az alosztás pedig mindegyikkel használható együtt. Természetesen abban az esetben, amikor „mint egészet” módosítja egy alosztás, a csoportképzéssel képzett jelzetet összekötéssel is hozzákapcsolhatjuk egy másikhoz, a kapcsolatok jellegéből következően.²²⁶

Az 1990. évi magyar kiadásban található példát megvizsgálva azt láthatjuk, hogy ha csoportképzést használunk, *az összekötés a viszonyításra nézve disztributív*. Eszerint a 331.31: [622+629] Munkaidő a bányászatban és kohászatban viszonyítás az összekötést elemeire bontva is felírható [331.31:622]+[331.31:629].²²⁷ A klasszikus eljárás az *alosztások esetén is* hasonló lehet: a csoportképzést felbontva annak minden eleméhez hozzákapcsolhatjuk valamennyi alosztást.²²⁸

Természetesen, ha az eredeti szögletes zárójelben lévő összekötés három számot kapcsol össze, akkor a felbontás után kapott is három, összetett jelzetet fog összekapcsolni. Ha a viszonyítás első tagja is egy zárójelben lévő összekötés lenne, akkor már hat elemű összekötést kapnánk. Tehát ez a típusú felbontás hosszú, nehezen áttekinthető jelzeteket eredményez, melyeket minden egyes elem, illetve részjelzet alapján visszakereshetővé kell tenni. Arról nem is beszélve, hogy ebben az esetben is

²²¹ Ld. Simple relation és -042 („UDC Online”), illetve „Major Changes to the UDC 1993-2013”.

²²² Slavić 2009.

²²³ Babiczky 1985.

²²⁴ FID 1983: 14.

²²⁵ Ld. Subgrouping („UDC Online”) és OSZK KI: 1/12.

²²⁶ Piros 2014.

²²⁷ OMIKK és OSZK-KMK: 12.

²²⁸ Piros 2014.

sérülhetnek a szöveg reprezentációjára vonatkozó elvárások.²²⁹ Ezért a zárójel felbontásánál és az ez által kapott rengeteg kifejezés kezelésénél célszerűbbnek tűnik megtartani a csoportképzést és a jelzet szerkezetét azzal együtt rögzíteni.

A szögletes zárójelnek, bár nagyon hasznos, nem eléggé elterjedt a használata, és *gyakran hiányzik olyan jelzetekből, amelyekben szükség lenne rá a kapcsolatok világossá tételéhez*. Például a 5+6"20"(540) jelzet a leírása szerint azt jelenti, hogy XXI. századi tudomány és technológia Indiában. A Wellisch által megfogalmazott „**általános összekötési szabály**” („general bonding rule”) szerint azonban az alosztás csak a tőle balra írt elemet minősíti, és nem vonatkozik valamennyi viszonyító jelre, amely ez előtt van.²³⁰ A szabály alapján alosztás természetesen a viszonyításnál általánosabb összekötésre sem vonatkozhat. Ha ezt a szabályt figyelmen kívül hagyjuk, akkor nem egyértelmű, hogy a két alosztás az agglomerált tárgykörre vagy csak annak a második összetevőjére vonatkozik.²³¹

A fenti példával ellentétben az azt idéző cikk egy másik esetben már használja a csoportképzést: [5+6](05).²³² Helyesen mindkét esetben használni kellene a szögletes zárójelet, hogy az alosztások kapcsolódása teljesen világos legyen.

A másik nagy probléma a csoportképzéssel kapcsolatban, hogy a szabályok között nem esik szó a *zárójelek egymásba ágyazhatóságáról*. Ha csak egy szinten használhatók, az éppen a fent látható kétértelműséghez vezethet. Például az alábbi komplex példában az alosztások precedenciája felborul (a szögletes zárójelben álló 92 a teljes összekötéssel van viszonyítva, tehát az összekötésnek helyesen szögletes zárójelben kellene állnia) amiatt, mert csak a külső szinten lehet szögletes zárójelet használni.²³³

394.4:[92(100+437):329(437).15(091)+327.32(100)] A csehszlovák kommunista párthoz és a nemzetközi munkásmozgalomhoz kapcsolódó zászlók és jelképek nyilvános ünnepségeken.

3.4.3 A jelzetek besorolására vonatkozó szabályok

A dokumentumok tartalmi besorolása esetén általában *az általános témától a specifikusig haladás* elvét szokás követni, ami megfelel az információkeresés, illetve böngészés (és még általánosabban, az emberi gondolkodás) általános logikájának.²³⁴

Megállapodás szerint a csak általánosan közös alosztásokat tartalmazó jelzetek megelőzik az összes többi. Ezen belül sorrendben először a nyelvi, majd a formai, földrajzi, etnikai, végül az idő alosztásoknak kell következniük (ahogyan a táblázatokban is).

²²⁹ Például a [74+76](02.029.1) azt sugallja, hogy egy rajzokat és grafikákat egyaránt tartalmazó albumról van szó, a 74(02.029.1)+76(02.029.1) pedig azt, hogy egy rajzokat tartalmazó albumból és egy grafikákat tartalmazó albumból álló kiadványról beszélünk.

²³⁰ Wellisch 1969: 4; idézi Dahlberg 1972.

²³¹ Azt Wellisch is felismerhette, hogy a szabály önmagában kevés az ilyen jellegű ellentmondások egyértelmű és szemléletes feloldására, mert ő javasolta (Perrault mellett) a csoportképzés mai szabályok szerinti felhasználását is (Kent, Lancour és Daily 1981: 142).

²³² Az idézett példák forrása Satija 2008.

²³³ Piros 2014. A példa forrása Smiraglia et. al. (2013). Bár ugyanezt a tartalmat egyszerűbben is ki lehetne fejezni, az OCLC WorldCatban valóban létező jelzetről van szó.

²³⁴ A fejezetben szereplő szabályok forrása a BSI standard kiadásának előszava (BSI 2005: xxi-xxiii.). A táblázat szintén innen származik, de ki lett egészítve a sorrendrögzítéssel és a szempont szerinti alosztásokkal, az 1990. évi magyar kiadás alapján (OMIKK és OSZK-KMK: 7).

Az egyszerű ETO számok esetén az általánostól a specifikus felé haladás a rövidebb számok hosszabbak elé sorolását jelenti. Az egymásnak nem alárendelt fogalmak esetén a decimális számok adják a rendezés természetes sorrendjét.

Az agglomerált, komplex és összetett tárgykörök esetén a helyzet kissé bonyolultabb: ehhez először is meg kell határozni az alosztások specifikálásának sorrendjét. Az természetes, hogy az agglomerált tárgykörök általánosabbak az egyszerűeknél, hiszen az agglomeráció kibővíti a fogalmak jelentését. A fázis- és facettarelációk minden esetben pontosítják – tehát szűkítik – a jelentést. A pontos sorrendjük ugyanaz, mint amit a csak alosztásokat tartalmazó számok esetén már említettünk. A teljes besorolási sorrendet a 2. táblázat tartalmazza.

Jele	Megnevezése	Példa
=...	Nyelvi általánosan közös alosztások	=112.2
(0...)	Formai általánosan közös alosztások	(0.035.2)
(...)	Földrajzi általánosan közös alosztások	(430)
(=...)	Etnikai általánosan közös alosztások	(=1.410)
"..."	Idő szerinti általánosan közös alosztások	"18"
+	Összekötés	622+669
/	Kiterjesztés	622/623
0/9	Főtáblázati szám önmagában	622
:	Viszonyítás	622:338.3
::	<i>Sorrendrögzítés</i>	622::338.3
=...	Nyelvi általánosan közös alosztások	622=112.2
(0...)	Formai általánosan közös alosztások	622(0.035.2)
(...)	Földrajzi általánosan közös alosztások	622(430)
(=...)	Etnikai általánosan közös alosztások	622(=1.366)
"..."	Idő szerinti általánosan közös alosztások	622"18"
*...	Külső forrásból vett jelzetek	622*FE ₂ O ₃
A/Z	A/Z névalosztások	622GOE
.00...	<i>Szempont szerinti nem önálló általánosan közös alosztások</i>	
-0...	Általános ismérvek szerinti nem önálló általánosan közös alosztások	622-057.2
-1/-9	Kötőjeles korlátozottan közös alosztások	622-78
.0...	Pont nullás korlátozottan közös alosztások	622.03
'...	Aposztrófus korlátozottan közös alosztások	622'17
	Következő egyszerű főtáblázati szám	623

2. táblázat *Az alosztások besorolási sorrendje*

Besoroláskor tehát az azonos főtáblázati számmal jelölt dokumentumok csoportosítását „a jelzatláncban balról jobbra következő jelzet típusa, illetve jele határozza meg, ezen belül pedig újra a decimális számrend érvényesül. Ezt követően újra a következő jelet kell figyelembe venni, és így folytatni tovább a jelzatlánc végéig.”²³⁵

²³⁵ OSZK KI: 1/8.

A csoportképzés nem befolyásolja a besorolást, kivéve olyan jelzetek között, amelyek csak a szögletes zárójelben különböznek. Ilyenkor a szögletes zárójel nélküli jelzetet kell előbbre sorolni.

Az ETO lehetővé teszi az **interkalációt**, azaz egy – nyitó és záró határolójellel is rendelkező – alosztásnak egy főtablázati szám belsejébe való besúrását. Ilyen esetben a besoroláskor csak a főtablázati szám alosztás elé eső részét kell figyelembe venni, hiszen az interkaláció alkalmazását éppen ez indokolja.²³⁶

3.5 A kontextus jelentősége

Az ETO analitikus-szintetikus jellegéből következően erősen prekoordinált rendszer: szabályai úgy vannak kialakítva, hogy indexeléskor a komplex, agglomerált és összetett tárgykörök egyaránt leírhatók legyenek benne, utalva nem csak az őket alkotó fogalmakra, de azok szerepére és a kapcsolataik típusára is. Ez *a rugalmasság ugyanakkor növeli a leírásokban érvényesülő szubjektivitást.*²³⁷

Korábban már említettük, hogy a szintaktikai relációkat paradigmatis és szintagmatikus kapcsolatokra szoktuk bontani, attól függően, hogy gyakori, tartósan fennálló, esetleg az osztályozási rendszerben is szereplő vagy eseti, csak az éppen leírt dokumentumban előforduló viszonyról van szó.

Perrault az ETO kapcsolatait két nagy csoportba sorolta, megkülönböztetve a **logikai** és a **dokumentációs kapcsolatokat**, annak függvényében, hogy a fogalmi szerkezetre vagy a dokumentumra utalnak, tehát az *aboutness* vagy az *isness* kifejezésére szolgálnak.²³⁸ Ha a két csoportosítás kapcsolatát megvizsgáljuk, azt láthatjuk, hogy amíg *a dokumentációs kapcsolatok mindig szintagmatikusak*, hiszen egy adott dokumentum formai megjelenésére vonatkoznak, a logikai kapcsolatok lehetnek paradigmatis és szintagmatikusak, a vizsgált tárgykör gyakoriságától függően.²³⁹ A nyelv alkalmazásakor *elsősorban az indeterminisztikus szintagmatikus relációk kezelése jelent nagy kihívást.*

A szintaktikai relációk használata az információkeresés során elsősorban a precízió növelését célozza, illetve a relevancia minél pontosabb és gyorsabb eldöntését. Mély ismeretük azonban fontos a visszahívás szempontjából is: ha nem veszünk minden eshetőséget figyelembe az index összeállításánál, a keresőkérdés - manuális vagy automatikus – megfogalmazásakor, illetve böngészéskor, az

²³⁶ BSI 2005: xix.

²³⁷ Természetesen ez a hatás magasabb fokú kontrollal csökkenthető, illetve kiküszöbölhető. Ugyanakkor a dokumentumokhoz való – globális – hozzáférés szempontjából a rekordok megfelelő összekapcsolása esetén előnyös, ha a különböző indexelők különböző nézőpontjai is érvényesülnek (Tillet 2015).

²³⁸ Perrault szerint a csoportképzés és a formai alosztás minden esetben dokumentációs, az összekötés és a nyelvi alosztás pedig egyaránt lehet logikai és dokumentációs. A többi alosztást a dokumentációs csoportba sorolta. (Perrault: 2001) Ha az alosztások azóta megváltozott definícióit nézzük, a csoportképzés egyértelműen átkerült a dokumentációs kapcsolatok közé, és a formai alosztás is vonatkozhat a tárgyalásmódra, tehát lehet dokumentációs is.

²³⁹ Vannak olyan összetett jelzetek, amelyek a táblázatokban is szerepelnek, így már az ETO szintjén is paradigmatisak (például a (470+571) vagy 017/019).

releváns rekordok elsikkadását okozhatja. Különösen igaz ez egy annyira flexibilis rendszerben, amilyen az ETO, amely ugyanazon tárgykörök kifejezésére is számos lehetőséget kínál.²⁴⁰

A rendszer tudatos és szakszerű facettásításával párhuzamosan *a szintaktikai relációk jelentősége is növekszik*. Ugyanakkor éppen a fejlesztések átgondoltsága, módszeressége és következetessége, a jelzetképzésre vonatkozó szabályok szigorodása és az enumerikus táblázatok példákkal való helyettesítése várhatóan növeli majd a konzisztenciát a jelzetek képzésében is.

3.5.1 Az alosztási jelek jelentésmódosító hatása

A kétféle agglomeráció, azaz az összekötés és a kiterjesztés közötti különbség nyilvánvaló: amíg az összekötés olyan tartalmakra vonatkoztatható, amelyekben csak a feltüntetett fogalmak szerepelnek, a kiterjesztés rejtetten tartalmazza a közöttük lévő valamennyi fogalmat. Tehát *a kiterjesztéssel jelölt tárgykör terjedelme általában nagyobb, mint az összekötéssel jelöltté*.²⁴¹ Például amíg a 21+29 jelentése „történelem előtti és primitív vallások és modern spirituális mozgalmak”, a 21/29-é „Vallási rendszerek. Vallások. Hitfelekezetek.”²⁴², amely tartalmazza a távol keleti, keresztény stb. vallásokat is.

Ha a két jelzetet *viszonyítással* kapcsoljuk össze (21:29), akkor a kialakult komplex tárgykör jelentése: „a primitív vallások és a modern spirituális mozgalmak kapcsolata”. A kapcsolat iránya is kifejezhető a *sorrendrögzítés* (21::29) használatával, ilyen esetben a jelentés „a primitív vallások hatása a modern spirituális mozgalmakra”.

A *csoportképzés* szerepe valamivel bonyolultabb. A [27+28](4)(091) jelzetben az alosztások mindkét fő táblázati számra vonatkoznak, így a jelzet jelentése „a kereszténység és az iszlám története Európában”. Ha egyszerűen elhagyjuk a szögletes zárójelet, akkor az alosztások csak a második fő táblázati számra vonatkoznak, így a jelentés „a kereszténység és az európai iszlám története” lesz.

Az *aposztrófós jelzetszintézis* esetén is egy egyszerű példával világíthatjuk meg az aposztróffal és a kettősponttal képzett komplex tárgykörök közötti különbséget: a 329.12'17 jelentése „liberális nacionalizmus”, a „329.12:329.17” pedig a liberális és a nemzeti mozgalmak kapcsolata. Itt fontos látnunk azt is, hogy az aposztróffal szintetizált jelzet a második és további számokat rövidítve tartalmazza, amit értelmezésnél szintén figyelembe kell venni. A „329.12:17” jelentése ugyanis „a liberalizmus etikája”. Szintén érezhető a sorrend jelentést árnyaló szerepe a szintézis során.²⁴³

²⁴⁰ A rendszernek ez a rugalmassága az előnyei és hátrányai között egyaránt gyakran van említve: egyfelől lehetőséget biztosít a tartalmak egészen pontos leírására, másfelől viszont szinte lehetetlenné teszi azt, hogy meggyőződjünk arról, hogy mindig, minden lehetőséget számba vettünk.

²⁴¹ Kivéve, ha az összekötés szomszédos jelzeteket kapcsol össze. Ilyen esetben a két művelet elvileg ekvivalens, gyakorlatilag a kiterjesztés használata azt sugallhatja, hogy a tárgykörök nincsenek olyan mereven elválasztva, mint az összekötés esetében.

²⁴² A vallás főosztályba tartozó példák a UDC Online-ból („UDC Online”) származnak, tehát a főosztály teljes revíziója utáni állapotot tükrözik. A további példák forrása a 2005-ös magyar kiadás (OSZK KI 2005).

²⁴³ Megjegyzendő, hogy az ilyen jellegű jelzetszintézis az elmúlt húsz évben szisztematikusan kivezetésre kerül. Amíg a 2005. évi (2000. évvel lezárt MRF-re épülő) magyar kiadás a 329.1/.6 („politikai pártok és mozgalmak”), 355 („hadtudomány”) és az 546 („szervetlen kémia”) osztályok alatt van lehetőség ilyen jellegű jelzetkapcsolás alkalmazására, a 2005. évi BSI kiadásban a „hadászat” osztály már nem tartalmazza ezt a lehetőséget. A jelenlegi MRF-ben ez az eszköz kizárólag a kémiai vegyületek számára van fenntartva (ld. OSZK KI 2005: 239, 315 és 435, BSI 2005: 240, 301 és 392-393 és „UDC Online”).

A viszonyítás esetén nincs jelentése a sorrendnek: az ajánlás az, hogy a számokat növekvő sorrendben kell szerepeltetni, de ettől el lehet térni, ha az adott katalógus vagy gyűjtemény szempontjából ez tűnik célszerűnek.²⁴⁴

Alosztási jelek általánosan közös alosztási számok között is előfordulhatnak. Például a 339.5(44:450) jelentése „kereskedelmi kapcsolatok Franciaország és Olaszország között”, 339(44+450) pedig „kereskedelem Franciaországban és Olaszországban”. Itt is látható a jelentésbeli különbség, de arra is fontos rámutatni, hogy a jelzetek automatikus feldarabolásánál az ilyen esetek külön kezelendők.²⁴⁵

Legvégül meg kell említenünk a pont szerepét is. Ez a karakter eredetileg csak központosítási, olvasást megkönnyítő szereppel bír²⁴⁶, ezen kívül sem szemantikai sem szintaktikai szerepe nincs. Különösen automatizált jelzetfeldolgozás, keresés vagy böngészés során fontos ezt figyelembe venni. Például a 225 jelentése „Japán vallásai”, a 24, a buddhizmust jelenti. 225.24 a 225 továbbosztása, jelentése „Sekai Kyuseiko”. Nyilvánvaló, hogy ebben az esetben a jelzet 24 felőli visszakereshetősége zajt okozna.

3.5.2 A korlátozottan közös alosztások

A kontextus jelentést befolyásoló szerepe talán a legeggyértelműbben a korlátozottan közös alosztások által leírt, klasszikus értelemben vett facetták esetén jelentkezik. Ezek az alosztások csak egy adott bázisosztály alatt, tehát a táblázatok egy adott helyén használhatók, *ugyanaz a karaktorsor másik bázisosztály alatt, egészen más jelentéssel szintén előfordulhat*. Emiatt, ha pusztán a korlátozottan közös alosztást alkotó számokat tekintjük, és nem vesszük figyelembe azt a fogalmat, amelyet pontosítanak, az jelentős zajt okozhat. Például a -3 karaktorsor egyaránt jelenthet „vallásokhoz kapcsolódó személyeket”, „számítási technikákat”, „vezérművet” vagy akár „prózá”, annak megfelelően, hogy a vallásokhoz (2), a matematikához (51), a technikához (62) vagy az irodalomhoz (82) tartozó számmal jelölt bázisosztállyal van közösen használva.²⁴⁷

Az analitikus-szintetikus jelleg, azaz a facetták és a jelzetösszekapcsolások együttes használatának lehetősége egy jelzeten belül is okozhat jelentésbeli eltéréseket. Például az 512.54.03:531.111.5-3 és az 512.54-3:531.111.5.03 jelzetekben a korlátozottan közös alosztások felcserélése jelentésváltozáshoz vezet, hiszen mindkét fő táblázati számmal használható együtt a -3 és .03 is, csak eltérő jelentéssel.

Az ETO-ban a korlátozottan közös alosztások nem mindig egyszerű felsorolások, összetett tárgykörök számos esetben a táblázatok egy adott helyéről vagy akár a táblázatok tetszőleges helyéről származó jelzetek felhasználásával, az alkalmazás helyén leírt szabályok szerint képezhetők²⁴⁸, és nem csak egy bázisosztályt pontosíthatnak, hanem akár egy intervallumot, összetett, illetve – csoportképzés használata esetén – komplex vagy agglomerált tárgykört is.

²⁴⁴ McIlwaine, I. C. 2000: 37-39.; hiv. Robinson 2003. Természetesen gépi környezetben a legjobb az, ha a sorrendtől függetlenül egyformán elérhetők a jelzetek pl. a 329.12, 329.17, 329.12: 329.17 és 329.17:329.12 számok alapján keresve, illetve böngészve is.

²⁴⁵ Smiraglia et. al. 2013 cikkükben részletezik, hogyan vezetett hamis indexekhez az itt idézetthez hasonló jelzetek automatikus felbontása (az indexben négyessel kezdődő számok szerepeltek, ami nem lehetséges, mivel a négyes főosztály üres).

²⁴⁶ Kivéve a .0-ás korlátozottan közös alosztásokat és a már eltörölt szempont szerinti alosztásokat.

²⁴⁷ A könnyebb érthetőség és a szemléletesebb kifejezőmód érdekében a dolgozat további részében a bázisosztály, illetve fókusz kifejezések helyett csak osztályokról, alosztásokról, ETO számokról lesz szó.

²⁴⁸ Gnoli (2011) részletesen tárgyalja az ETO-ban előforduló facettákat, az itt felsorolt három alapesetre rendre kontextus-definiált, speciális- és általános extra-definiált fókuszként hivatkozva.

A fentiek miatt a korlátozottan közös alosztások kezelése az egyik legproblematisabb eleme az ETO számok számítógépes kezelésének.

3.5.3 Az általánosan közös alosztások

Az analitikus-szintetikus jelleg az általánosan közös alosztások esetében is vet fel problémákat. Akárcsak a korlátozottan közös alosztások esetében, *itt sem mindegy, melyik alosztás melyik fogalmat pontosítja*. Például a 341.232.3(47):330.34(540) jelentése Oroszország által Indiának nyújtott gazdasági segítség, míg 341.232.3(540):330.34(47) ennek pont a fordítottja.

Megkötés az **alosztások sorrendjére** nincs, csak ajánlás: általában a jelzetek besorolási sorrendjének az ellentétét célszerű használni, ami így a *specifikustól az általánosig* való haladást jelent. Ettől, ha a helyi környezet úgy kívánja, el lehet térni, de ügyelni kell rá, hogy adatcsere esetén a sorrendek megegyezzenek.²⁴⁹

A fentiek alapján az alosztások sorrendjének *elvileg nincs jelentést befolyásoló szerepe, a gyakorlatban azonban előfordulnak* ilyen esetek is. A 2005. évi magyar kiadás előszavában ez szerepel: „Amennyiben ezen a sorrenden valamilyen oknál fogva változtatunk, a fogalmi láncban nyomatékosabb jelentést kapnak az előrehozott elemek.” Robinson példákat is hoz olyan esetekre, amikor az általánosan közös alosztások sorrendjének eltérése számottevően befolyásolja a jelzetek jelentését.²⁵⁰ A 2005. évi angol nyelvű standard kiadás előszava a formai alosztások esetén tesz különbséget **belső** (tehát a tárgyalásmódra vonatkozó) és **külső** (tehát a dokumentum fizikai megjelenését leíró) **formai alosztások** között, azt javasolva, hogy amennyiben egy jelzet mindkettőt tartalmazza, mindig a belső kerüljön előbbre, függetlenül a számsorrendtől.²⁵¹

3.6 Irodalomjegyzék

Babiczy Béla. 1975. *Könyvtári feldolgozó munka: A tanítóképző intézet II. évfolyama számára (népművelő-könyvtáros szak). 2. rész. Osztályozás és szakkatalógus szerkesztés*. Budapest: Tankönyvkiadó.

Babiczy Béla. 1985. „Szemléletváltozás az ETO jelzetszerkesztésében.” *Könyvtári Figyelő* 31 (1): 17–27.

Batley, Susan. 2014. *Classification in Theory and Practice*. Amsterdam: Elsevier Science.

Barátné Hajdu Ágnes. 2011. „Az ismeretszervezés dinamizmusa, különös tekintettel az Egyetemes Tizedes Osztályozásra.” *Könyv és nevelés* 13 (1): 63-72.

Benito, Miguel. 2007. „Better consistency of the UDC system moving Medicine from section 61 to section 4.” *Extensions and Corrections to the UDC* 29: 311-316.
<http://hdl.handle.net/10150/105244>. (megtekintve: 2017-11-01)

²⁴⁹ Az ajánlott sorrend tehát: 0/9, korlátozottan közös alosztás, -0 (csökkenő számsorrendben), A/Z, „...”, (=...), (1/9), (0...), =... (McIlwaine 2000: 37-39.; hiv. Robinson 2003). Természetesen, ahogy a viszonyítás esetén, itt is az lenne a legjobb, ha a programok képesek lennének felismerni a csak sorrendben különböző jelzetek azonosságát, ebben az esetben az adatcsere esetére tett megszorítástól is eltekinthetnénk.

²⁵⁰ Például 94(410)20(051) („Nagy-Britannia XXI. századi történelméről szóló folyóiratok”), 94(410)(051)20(051) („Nagy-Britannia XXI. századi történelméről szóló XXI. századi folyóiratok”) és 9420(051)(410) („A XXI. század történelméről szóló Nagy-britanniai folyóiratok”), illetve 069.9(100)1851(410.111) („A Nagy Kiállítás”, London, 1851) és 069.91851(100)(410.111) („Londonban, 1851-ben tartott nemzetközi kiállítás”) (Robinson 2002, Robinson 2003).

²⁵¹ Például [792+82-2](091)(086.7) („A színházi előadások története – hanganyag”) (BSI 2005: 17).

- „A biterminal symbol for non-UDC notation.” Notes & Queries. UDC Consortium (UDCC).
<http://www.udcc.org/notes.htm#bitsymbol>. (megtekintve: 2017-11-01)
- British Standards Institution (BSI). 2005. *Universal Decimal Classification: standard edition: volume 1: systematic tables*. London: British Standards Institution.
- Buxton, Andrew. 1990. „Computer Searching of UDC Numbers.” *Journal of Documentation* 46 (3): 193-217.
- Buxton, Andrew. 2011. „Duplication of Concepts in UDC.” *Extensions & Corrections to the UDC* 33: 13-17. <http://hdl.handle.net/10150/232711>. (megtekintve: 2017-11-01)
- „Cancellations to the UDC.” UDCC (UDC Consortium).
<http://udcc.org/index.php/site/page?view=cancellations>. (megtekintve: 2017-11-01)
- Chatterjee, Amitabha. 2015. „Universal Decimal Classification and Colon Classification. Their mutual impact.” *Annals of Library and Information Studies (ALIS)* 62 (4): 226-230.
<http://op.niscair.res.in/index.php/ALIS/article/view/11208>. (megtekintve: 2017-11-01)
- Csabay Károly. 1999. Cédulák, ETO-szám, szócikkek. *Tudományos és műszaki tájékoztatás* 46 (1): 12-32. http://tmt-archive.omikk.bme.hu/show_news.html?id=1933&issue_id=4. (megtekintve: 2017-11-01)
- Csik Tibor. 1995. „Ismeretek és könyvtári osztályozás.” *Könyv, Könyvtár, Könyvtáros* 4 (4): 13-24.
<http://epa.oszk.hu/01300/01367/00163/pdf/02muhelykerdesek.pdf>. (megtekintve: 2017-11-01)
- Dahlberg, Ingetraut. 1972. „Az ETO újjáalakításának lehetőségeiről” [Fordította Horváth Magda]. *Könyvtári Figyelő* 18 (4): 412-430.
- de Grolier, Eric. 1962. *A study of general categories applicable to classification and coding in documentation*. Paris: UNESCO.
- Freeman, Robert R., és Pauline Atherton. 1967. *File Organization and Search Strategy Using the Universal Decimal Classification in Mechanized Retrieval Systems*. New York: American Institute of Physics (AIP).
- FID (International Federation de Documentation). 1983. *Az Egyetemes Tizedes Osztályozás (ETO) alapelvei, revíziójának és kiadásának szabályai (FID 603)*. Budapest: OMIKK.
- Gnoli, Claudio. 2011. „Facets in UDC: a review of current situation.” *Extensions and Corrections to the UDC* 33: 19-36. <http://arizona.openrepository.com/arizona/handle/10150/236511>. (megtekintve: 2017-11-01)
- Gulati, Dipti. 2013. „Use of library classification schemes in the ICT environment in selected libraries in national capital region: a study.” PhD dissz., Delhi: University of Delhi.
<http://shodhganga.inflibnet.ac.in:8080/jspui/handle/10603/31769>. (megtekintve: 2017-11-01)
- Horváth Tibor, és Sütő Péter. 2001. „A tartalmi feltárás.” In *Könyvtárosok kézikönyve. 2. kötet. Feltárás és visszakeresés*, szerk. Horváth Tibor, és Papp István, 35-186. Budapest: Osiris Kiadó, 2001.
- Kent, Allen, Harold Lancour, és Jay E. Daily. szerk. 1981. *Encyclopedia of Library and Information Science: Volume 32 - United Kingdom: National Film Archive to Wellcome Institute for the History of Medicine*. New York: Marcel Dekker, Inc.
- „Major Changes to the UDC 1993-2013.” UDC Consortium (UDCC).
http://www.udcc.org/index.php/site/page?view=major_revisions. (megtekintve: 2017-11-01)

- Mándy Gábor. 2011. „Régi ETO, új ETO.” *Könyvtári Figyelő* 57 (2): 317-332.
<http://ki.oszk.hu/kf/2011/07/regi-eto-uj-eto/>. (megtekintve: 2017-11-01)
- McIlwaine, I. C. 1998. „The Universal Decimal Classification: some factors concerning its origins, development and influence.” In *Historical studies in information science*, szerk. Trudi Bellardo Hahn, és Michael Buckland, 94-106. Medford, NJ: Information Today.
- McIlwaine, I. C. 2000. *The Universal Decimal Classification – a Guide to its Use*. The Hague: UDC Consortium.
- McIlwaine, I. C., és Nancy Williamson. 2008. „Medicine and the UDC: the process of restructuring Class 61.” *Extensions and Corrections to the UDC* 30: 9-16.
- Országos Széchényi Könyvtár Könyvtári Intézet (OSZK KI). 2005. *Egyetemes tizedes osztályozás. 1. kötet Táblázatok 1-2. Rész. UDCC Publication P057*. Budapest: Országos Széchényi Könyvtár Könyvtári Intézet.
- Országos Műszaki Információs Központ és Könyvtár (OMIKK), és Országos Széchényi Könyvtár Könyvtártudományi és Módszertani Központ (OSZK-KMK). 1990. *Egyetemes tizedes osztályozás. Rövidített kiadás. 1. kötet Táblázatok. FID Publication 691*. Budapest: OMIKK, OSZK-KMK.
- Pálvölgyi Mihály. 1998. *Információfeldolgozás, információkereső nyelvek*. Szombathely: Berzsényi Dániel Tanárképző Főiskola.
- Perrault, Jean Martin. 2001. „A kategóriák és relátorok új rendszere.” In *Osztályozás és információkeresés. Kommentált szöveggyűjtemény. Első kötet: Az osztályozás elmélete*, szerk. Ungváry Rudolf, és Orbán Éva, 200-217. Budapest: Országos Széchényi Könyvtár.
- Pestana, Olívia. 2015. „Alignment in Medical Sciences: towards improvements of UDC.” In *Classification & Authority Control: Expanding Resource Discovery : Proceedings of the International UDC Seminar 2015, 29-30 October 2015, Lisbon, Portugal*, szerk. Aida Slavić, és Maria Inês Cordeiro, 135-150. Würzburg: Ergon Verlag.
- Piros Attila. 2014. „Az ETO alapú számítógépes információkeresés egy új megközelítése.” *Könyvtári Figyelő* 60 (2): 156-166.
http://epa.oszk.hu/00100/00143/00091/pdf/EPA00143_konyvtari_figyelo_2014_2_156-166.pdf. (megtekintve: 2017-11-01)
- Piros Attila. 2015. „Az ETO-jelzetek automatikus elemzésének kérdései.” *Tudományos és műszaki tájékoztatás* 62 (4): 127-142. http://tmt-archive.omikk.bme.hu/show_news.html?id=6025&issue_id=569. (megtekintve: 2017-11-01)
- Ranganathan, Siyali Ramamrita. 1956. „Depth classification 21 phase and intra-facet relations.” *Annals of Library and Information Studies (ALIS)* 3 (4): 109-116.
<http://nopr.niscair.res.in/handle/123456789/28577>. (megtekintve: 2017-11-01)
- Riesthuis, Gerhard. 1997. „Decomposition of complex UDC notations.” *Extensions and Corrections to the UDC* 19: 13-19.
- Riesthuis, Gerhard. 1999. „Searching with words: re-use of subject indexing.” *Extensions & Corrections to the UDC* 21: 24-32.
- Robinson, Geoffrey. 2002. „More about the meaning of common auxiliaries.” *Extensions and Corrections to the UDC* 24: 8. <http://www.udcc.org/notes.htm#common>. (megtekintve: 2017-11-01)

- Robinson, Geoffrey. 2003. „Citation Order in UDC.” *Extensions and Corrections to the UDC* 25: 19-27.
- Salah, Almila Akdag, Cheng Gao, Krzysztof Suchecki, és Andrea Scharnhorst. 2011. „Need to categorize: A comparative look at the categories of the Universal Decimal Classification system (UDC) and Wikipedia.” In *Paper for High Throughput Humanities - a satellite meeting at the European Conference on Complex Systems 2010; Lisbon, Portugal*.
<https://arxiv.org/abs/1105.5912>. (megtekintve: 2017-11-01)
- Satija, Mohinder. 2008. „Universal Decimal Classification: past and present: a state of the art.” *DESIDOC Journal of Library and Information Technology* 28 (6): 3-10.
- Shenoy, Rekha. 1998. „Organisation of Knowledge in General Classification Schemes: a Comparative Study of Social Science Subjects in Colon Classification and Dewey Decimal Classification.” PhD dissz., Mangalore: Mangalore University.
<http://shodhganga.inflibnet.ac.in/handle/10603/131771>. (megtekintve: 2018-03-16)
- „Should the slash / be abolished?” Notes & Queries. UDC Consortium (UDCC).
<http://www.udcc.org/notes.htm#stroke>. (megtekintve: 2017-11-01)
- Slavić-Overfield, Aida. 2005. „Classification Management and use in a Networked Environment: The Case of the Universal Decimal Classification.” PhD dissz., London: University College London.
http://discovery.ucl.ac.uk/1334914/1/AidaSlavicOverfield_thesis_UCL2005.pdf. (megtekintve: 2017-11-01)
- Slavić, Aida. 2008. „Faceted classification: management and use.” *Axiomathes* 18 (2): 257-271.
- Slavić, Aida. 2009. „UDC SKOSification.” *Universal Decimal Classification* (blog), UDC Consortium, 2009. március 22. <http://universaldecimalclassification.blogspot.hu/2009/03/udc-skosification.html>. (megtekintve: 2017-11-01)
- Slavić, Aida, és Sylvie Davies. 2017. Facet Analysis in UDC: questions of structure, functionality and data formality. In *Faceted classification today: theory, technology and end users: proceedings of the International UDC Seminar 2017, London (UK), 14-15 September*, szerk. Aida Slavić, és Claudio Gnoli, 181-197. Würzburg: Ergon Verlag.
- Tillet, Barbara B. 2015. „Complementary of perspectives for resource descriptions.” In *Classification & authority control: expanding resource discovery: proceedings of the International UDC Seminar 2015, 29-30 October 2015, Lisbon, Portugal*, szerk. Aida Slavić & Maria Inês Cordeiro, 19-30. Würzburg: Ergon Verlag.
- UDCC (UDC Consortium). 2005. „Guidelines for the revision and maintenance of the UDC.” *Extensions and Corrections to the UDC* 27: 19-38.
- „UDC Online.” UDCC (UDC Consortium). <http://www.udc-hub.com/> (megtekintve: 2017-11-01)
- „UDC Subject Coverage.” UDCC (UDC Consortium).
http://udcc.org/index.php/site/page?view=subject_coverage. (megtekintve: 2017-11-01)
- „UDC Summary. Universal Decimal Classification Summary. UDCC Publication No. 088. Hungarian (Magyar).” UDCC (UDC Consortium).
<http://www.udcsummary.info/php/index.php?lang=hu&pr=Y>. (megtekintve: 2017-11-01)
- Wellisch, Hans H. 1969. *Bonds and Bonding of UDC Notational Symbols*. Collage Park, Md.: University of Maryland.

Wimmer, F. 2000. „The Apostrophe in UDC: A critical examination of an unwelcome development.”
Extensions and Corrections to the UDC 22: 19-27.

4 Az ETO gépesítése

„Egy jelzet alkotóelemei vagy az összetett osztályozási számok alapján történő keresés nehézségei abból adódnak, hogy ezek a jelenlegi könyvtári katalógusok számára, melyek nem arra vannak tervezve, hogy ilyen komplex formában tárolják az adatokat, nehezen feldolgozhatók.”²⁵² (Rosa San Segundo)

4.1 Történeti áttekintés

Ugyan az Egyetemes Tizedes Osztályozás már jóval az első mai értelemben vett számítógépek előtt megjelent, a már korábban említett dokumentációs jellege, analitikus-szintetikus struktúrája, az *indexelési lehetőségek rugalmassága, a jelzetelemek jó elkülöníthetősége a megfelelő szoftveres támogatás mellett különösen alkalmassá teszi gépi indexek építésére*. Másrésről, ahogyan korábban említettük, éppen ez az analitikus-szintetikus jelleg, a szintaktikai relációk – nem ritkán következtelen – sokrétűsége az, ami *bonyolult feladattá is teszi* a rendszer minden lehetőségére kiterjedő szoftverek írását.

Az osztályozás gyors elterjedése és gépesítésre való kivételes alkalmassága²⁵³ mellett nem meglepő, hogy már igen korán megtörténtek az első kísérletek a gépi támogatású dokumentációs és könyvtári munkában való felhasználására is.²⁵⁴ Osztályozási (a DDC-ből származó) kódok lyukkártyás rendszerben történő alkalmazására először a negyvenes évek elején került sor, az Egyesült Államokban, az *első ismert, ETO-t használó alkalmazást* pedig 1948-ban Nagy-Britanniában mutatta be **E. G. Birsch**.²⁵⁵

A korai számítógépes rendszerek korlátozott lehetőségei az osztályozási folyamat újra gondolását is szükségessé tették. Az egyik szélsőséges újítás Mortimer Taube *UNITERM* rendszere volt, amely *a teljes posztkoordináción*, a fogalmaknak a keresési fázisban, kizárólag logikai műveletekkel való összekapcsolásán alapult.²⁵⁶ Számos előnye és hátránya mögött a jelentősége leginkább abban rejlett, hogy felkeltette az érdeklődést a posztkoordinált szemlélet iránt, ami *meghatározta a többi osztályozási rendszer – így az ETO – gépesítésével kapcsolatos kutatások irányát is*. Végző soron a „relációs paradigma” ötvenes és hatvanas évekbeli fénykorát követően a relátorrendszerekkel, az ETO kapcsolatrendszerének kibővítésével kapcsolatos vizsgálatok hanyatlását is a korabeli számítógépek (és szoftverek) korlátai és a posztkoordináció olcsósága okozta.²⁵⁷

Az **ETO-jelzetek KWOC-indexének** első alkalmazása, azaz prekoordinált jelzetek elemeik alapján, posztkoordinált módon való kereshetővé tétele K. Schneider és K. H. Koch nevéhez fűződik, akik a

²⁵² San Segundo 2009.

²⁵³ Egy közelmúltbeli kutatás szerint a vizsgált osztályozási rendszerek közül a klasszikus és digitális gyűjteményekben való alkalmazhatóságot egyaránt vizsgálva a DDC és az ETO a leghatékonyabb, míg digitális gyűjteményekben az Association for Computing Machinery (ACM) digitális könyvtárának (egyelőre csak a számítástechnikát felölelő) rendszere bizonyult a legjobbnak (Ullah, Khusro és Irfan 2017).

²⁵⁴ Az ETO gépesítésének első szakaszát nagy részletességgel tárgyalja Malcolm Rigby (1974), illetve Bourne és Hahn (2003). A hatvanas évek végéig tartó időszakot magyarul Orosz Gábor dolgozta fel (1967).

²⁵⁵ Orosz 1967.

²⁵⁶ Chatterjee 2016: 197-201.

²⁵⁷ Ungváry és Orbán 2001a: 194-198.

hatvanas évek első felében építettek ilyen jellegű indexet.²⁵⁸ A megoldás UNITERM-el szembeni előnye, hogy *a kereső találkozhat az ETO-jelzettel, azaz a szurrogátummal is*, nem csak közvetlenül a dokumentummal. Hátránya, hogy a keresés és böngészés során *nem érvényesíthetők a szintaktikai relációk*, illetve, hogy a KWOC-indexet is el kell készíteni vagy manuálisan, ami drága és nyitva hagyja a hibázás lehetőségét (ami ebben az esetben adatvesztést eredményezhet), vagy automatikusan, ami egy az ETO összetettségével rendelkező rendszer esetében bonyolult algoritmusokat igényel. Hátrányai ellenére is elmondható, hogy *az authority control mellett a mai napig ez a megoldás jelenti a legtöbb ETO alapú keresőszoftver alapját*.²⁵⁹

A hatvanas években a világ számos pontján folytak kutatások az ETO gépesítésének különböző aspektusaival kapcsolatban. A FID ETO gépesítéssel foglalkozó albizottsága 1962 végén kezdte meg működését, a munkák összehangolása, illetve az információcsere biztosítása érdekében.²⁶⁰ Magyarországon ebből az időszakból elsősorban az ETO tárgymutatójának számítógépes összeállítására vonatkozó, az Országos Műszaki Könyvtár és Dokumentációs Központban végzett kísérletet lehet kiemelni.²⁶¹

A különböző egyéni és intézetek által finanszírozott projektek közül mind méretét, mind hatását tekintve kiemelkedik az Egyesült Államok Nemzeti Tudományos Alapja (National Science Foundation, NSF) által finanszírozott *Amerikai Fizikai Intézet (American Institute of Physics, AIP)* 1965 és 1967 között lefolytatott átfogó vizsgálata.²⁶²

A projekt vezetője Pauline Atherton, segítője pedig Robert R. Freeman volt. Céljuk az ETO gépesített információ visszakereső rendszerekben történő felhasználhatóságának megvizsgálása volt, amely magában foglalta az ETO hatékonyságának a felmérését és egy azt használó rendszer gyakorlatban történő kidolgozását is. Az elkészült rendszer az **AUDACIOUS** (Automatic Direct Access to Information with the On-Line UDC System, Automatikus direkt hozzáférés az információkhoz az On-line ETO rendszerrel), mely több szempontból is mérföldkőnek számított. Az egyik legfontosabb ilyen eredmény, hogy az AUDACIOUS volt *az első olyan szoftver, amely authority control használatával oldotta meg az ETO számok visszakereshetőségét* részeit, szemantikai kapcsolataik, illetve természetes nyelvű keresőszavak alapján.

Az osztályozási séma leírásánál láthattuk, hogy a *besorolás szabályai* meglehetősen komplexek, a felhasználók számára nehezen érthetők és a programok számára is nehezen érvényesíthetők. Utóbbira a megoldást szintén Atherton és Freeman találták meg azzal, hogy minden ETO számhoz hozzárendeltek egy olyan karaktersort, mely a sima karakterláncok közötti (tehát karakterkódok szerinti) rendezéssel éppen az ETO besorolási sorrendjét adja.²⁶³ Ehhez először is a jelentés nélküli

²⁵⁸ Orosz 1967. Schneider és Koch mellett Freeman és Rigby is dolgoztak KWOC-indexszel a hatvanas években (Bourne és Hahn 2003: 70).

²⁵⁹ Pontosabban a legelterjedtebb az ETO-jelzetek sima sztringként való tárolása és visszakereshetősége, amely minden szempontból nagyon hatékonytalan. Egy másik, kifinomultabb alternatíva a KWIC index csak ritkán van alkalmazva.

²⁶⁰ Az albizottságban 15 országot képviselő 18 szakember dolgozott, Dr. Malcolm Rigby majd Robert Freeman, még később J-P. Sydler vezetésével (Rigby 1974: 17-20).

²⁶¹ Orosz 1977.

²⁶² Bourne és Hahn 2003: 69-75.

²⁶³ Hasonló módszert használt Schneider és Koch is (Rigby 1974: i).

pontokat kell eltávolítani a jelzetből, majd az alosztásokat bevezető jeleket kell betűkre cserélni úgy, hogy az ábécé sorrend kövesse az alosztások besorolási sorrendjét.

Magyarországon a hetvenes években a *BIBDOSZ* rendszerben kísérleteztek az ETO számok hasonló elvek szerinti rendezésével.²⁶⁴ A kilencvenes évek elején Andrew Buxton hívta fel a figyelmet újra erre a módszerre az ETO gépesítésének nyolcvanas évek végi állapotát összefoglaló és az akkori jelentősebb ETO-t használó rendszereket bemutató cikkében²⁶⁵. Az ETO konzorcium jelenleg is az MRF-ben szereplő betűkódok használatát javasolja a jelzetrészek elkülönítése és a komplex jelzetek rendezése érdekében.

Az ugyanebben az időben az *Amerikai Geológiai Intézet*en belül Thomas W. Caless vezetésével folytatott projekt célja az Észak-Amerikán kívüli geológiai szakirodalom bulletinjének és absztraktjainak az összeállítása volt. A kifejlesztett rendszer érdekességét az adta, hogy *mátrixos megközelítést* használt az ETO-számok facettáinak kategóriák szerinti keresésére és hogy a Boolean-műveletek mellett a „*kisebb*”, „*nagyobb*” és „*egyenlő*” műveletek használatát, így a kiterjesztés és összekötés kezelését is lehetővé tette.²⁶⁶

Az AIP projekt lezárásaként 1968-ban a FID megszervezte az **ETO gépesítésével foglalkozó szemináriumot**, amely az első nagyszabású nemzetközi konferencia volt ebben a tárgykörben, és melyet később még két másik követett, 1970-ben és 1971-ben. 2007-ben az ETO Konzorcium a korai kutatók iránti tiszteletből nevezte el az osztályozástudománnyal, illetve az osztályozások hálózati környezetben történő felhasználásával foglalkozó konferenciasorozatot ETO Szemináriumnak.²⁶⁷

A *nyolcvanas években* az osztályozási rendszerek számítógépes alkalmazása iránti *érdeklődés visszaesett*. Az egyre jobban terjedő integrált könyvtári rendszerek és online katalógusok fejlesztői nem fordítottak figyelmet az osztályozási rendszerek kezelésére, azok nyilvánvaló előnyei ellenére sem.²⁶⁸ Az OPAC-ok parancs vezérelt menüje nem kedvezett a tárgykörök szemantikai kapcsolataira épülő osztályozási rendszereknek. Úgy tűnt, hogy a címek kulcsszavai, illetve a szerzők alapján történő keresés kielégíti a felhasználói igényeket, az osztályozáselméleti kutatásokban pedig *a szakértői rendszerek és az automatikus osztályozás került előtérbe* a meglévő rendszerek használata helyett.²⁶⁹ Mindezek ellenére sikeres implementációkra is voltak példák, elsősorban néhány intézmény saját fejlesztésű rendszerében.²⁷⁰ Ezek közül leggyakrabban a zürichi ETH-Könyvtár **ETHICS** (ETH Library

²⁶⁴ Pethes et. al. 1976.

²⁶⁵ Buxton 1990.

²⁶⁶ Kent, Lancour és Daily 1981: 141-142.

²⁶⁷ „International UDC Seminars”.

²⁶⁸ Pollitt 1997.

²⁶⁹ Slavić-Overfield 2005: 23-25.

²⁷⁰ A nyolcvanas évek végének legérdekesebb ETO implementációit Buxton vette sorra a már idézett cikkében (1990). Az ETHICS mellett megemlíti a British Library BLAISE-LINE, az ESA-IRS (Európai Űrügynökség Információkereső Szolgáltatása) MOLARS (a Meteorológiai Hivatal Könyvtárának keresőrendszere) és a Volkswagenwerk Data-Star adatbázisait, az OCLC egyesített online, illetve a korábban már említett Skót Sarkkutató Intézet katalógusát, illetve a Harwell Atomenergia Kutató Intézete által fejlesztett STATUSt, az Inmagic Inc., a Nemzetközi Fejlesztéskutató Intézet MINISIS adatbáziskezelőjét, valamint az UNESCO CDS/ISIS információkereső csomagját.

Information Control System) rendszerét szokták kiemelni, mint a kor legkifinomultabb ETO-t használó szoftverét.²⁷¹

A *kilencvenes évek* második felétől a szakirodalomban az a nézet terjedt el, hogy az osztályozás gépesítésére a legmegfelelőbb megoldást az **authority control** nyújtja, mely lehetővé teszi, hogy a felhasználók a mesterséges jelzetek használata (vagy akár a velük való találkozás) nélkül tudjanak az osztályozási jelzetek, esetleg a hozzájuk kapcsolódó tezaurusz deskriptorok alapján keresni, a fogalmak közötti szemantikai és asszociatív kapcsolatok alapján böngészni. Ahogy az elnevezés is mutatja, az authority fájlok másik feladata a helyi rendszeren és könyvtári hálózaton belüli kontroll, azaz a konzisztencia és az egységes, magas minőségű osztályozási gyakorlat biztosítása egy központilag karbantartott jelzetállomány nyújtásával. A módszer támogatása érdekében a Library of Congress²⁷² és az IFLA²⁷³ is létrehozta saját *osztályozási MARC formátumát*. Utóbbi az UNIMARC formátumcsalád részeként és kimondva is az ETO támogatására jött létre, mivel ebben az időben ez a formátum, illetve osztályozási rendszer volt a domináns Európában.²⁷⁴ A közelmúltban számos beszámoló is napvilágot látott *központi authority fájlok* készítéséről, illetve ilyen megoldással élő rendszerek használatáról Svájcban²⁷⁵, Romániában²⁷⁶, Csehországban²⁷⁷, Horvátországban²⁷⁸, illetve Magyarországon.²⁷⁹ A módszer jelentőségét mutatja, hogy a 2015. évi ETO Szeminárium teljes egészében ennek a témának volt szentelve.²⁸⁰

A hatvanas évektől kezdődően előbb az ETO jelzetszerkesztése²⁸¹ és szabályzata²⁸², később revíziós struktúrája, végül a kiadói és tulajdonosi köre, illetve a táblázatok mérete és karbantartásának módja is megváltozott.²⁸³ Utóbbi eredményeként, a kilencvenes évek elején létrejött a **Master Reference File (MRF)**, amely új perspektívákat nyitott az osztályozás gépesítésében. Egyfelől *a táblázatok gépi nyilvántartása megkönnyítette a karbantartásukat*, másrészt az adatbázisból készült exportokat a *szoftverek közvetlenül is fel tudják használni* pl. az ETO verziók közötti konkordancia követésére vagy a jelzetek természetes nyelven alapuló visszakeresésére. Utóbbit Gerhard Riesthuis vizsgálta behatóan: először olyan algoritmusokat dolgozott ki, amelyekkel fel tudta bontani a komplex

²⁷¹ Az ETHICS belső tezaurusza körülbelül 60 000 ETO-jelzetet tartalmazott, a hozzájuk készült összesen mintegy 350 000 angol, német és francia nyelvű szöveges leírással. Az ETO-jelzetek, illetve szöveges leírásaik alapján egyaránt lehetett keresni, és a találatok megjelenítésére is lehetőség volt mindhárom nyelven. (Hug és Nöthiger 1992).

²⁷² LOC 2000.

²⁷³ IFLA 2000.

²⁷⁴ Slavić-Overfield 2005: 194.

²⁷⁵ Pika 2007, Pika és Pika-Biolzi 2015.

²⁷⁶ Frâncu 2003, Frâncu és Sabo 2010, illetve Frâncu Dediu 2015.

²⁷⁷ Baliková 2005 és Baliková 2009.

²⁷⁸ Vukadin 2015.

²⁷⁹ Barátné 2009.

²⁸⁰ „International UDC Seminar 2015”.

²⁸¹ Babiczky 1985.

²⁸² FID 1983.

²⁸³ McIlwaine 1998.

jelzeteket alkotórészeikre, majd az MRF alapján az így kiemelt részekhez hozzárendelte a természetes nyelvű indexkifejezéseket.²⁸⁴

Az utóbbi harminc év átszervezései az *osztályozás teljesen facettás formában történő kialakítását* célozzák a facettaanalízis elvének konzisztens alkalmazása alapján, olyan teljesen facettás osztályozásokhoz hasonló módon, mint a BC2.²⁸⁵ Az átszervezések másik mozgató elve az a feltételezés, hogy a táblázatok szisztematikusabb struktúrája következetesebb jelzetelést, ennek következtében jobb jelzetkezelést eredményez online környezetben is, ezért számos olyan változtatás is történt a rendszerben, amely tudatosan a *jelzetek automatikus kezelésének megkönnyítésére* irányult.²⁸⁶

Az ETO hierarchikus szerkezete, a bővebb és szűkebb fogalmak közötti navigálás lehetősége miatt **tematikus keresők** összeállításához és indexeléséhez is hasznos eszköz lehet. A kilencvenes évek második felében a könyvtári katalógusok mellett az ilyen irányú felhasználásra is történtek kísérletek, sokáig ez volt az egyik preferált irány az ETO kutatásokon belül. Az ETO-t használó, *manuálisan indexelt tematikus keresők* (a legismertebbek a NISS - National Information Services and System és a SOSIG - Social Science Information Gateway) mellett az *automatikus indexelésre* is történtek próbálkozások, például a német GERHARD (German Harvest Automated Retrieval and Directory) és a dán Nordic WAIS/WWW rendszereiben.²⁸⁷

Az információkeresés megközelítésében újabb fordulat következett be 2001-ben, amikor Tim Berners-Lee közzétette **szemantikus web**bel kapcsolatos elképzeléseit.²⁸⁸ Ezt követően 2007-ig kellett várni arra, hogy Corey Harpey és Barbara Tillett kezdeményezésére a könyvtártudomány és információkeresés szakemberei is elkezdjék kidolgozni azokat az eszközöket, amelyek segítségével sok évtizedes tapasztalatuk, valamint a munkájuk eredményeként létrejött kontrollált szótárak és authority fájlok felhasználhatóvá válhatnak a szemantikus webben keresztül is.²⁸⁹ Az IFLA kidolgozta és a szemantikus világhálóra tervezett formátumban is közzétette a Funkcionális követelmények (Functional Requirements) bibliográfiai metaadatmodell-családot.²⁹⁰ A modellcsalád tárgyi authority rekordok számára kidolgozott ajánlása, az FR5AD (Functional Requirements for Subject Authority Data) elsősorban tezauruszok és subject heading listák leírására alkalmas, osztályozási sémákra csak a rendszerek részletes elemzése után alkalmazható.²⁹¹ A szemantikus webben alkalmazott formátumok közül *elsősorban a webes ontológiák leírására létrehozott OWL 2 (Web Ontology Language) használható az osztályozási rendszerek feldolgozására.* Az ajánlás és a formátum felhasználásával

²⁸⁴ Riesthuis 1997, Riesthuis 1999.

²⁸⁵ McIlwaine 1998, McIlwaine 2006, McIlwaine és Williamson 2008, Gnoli 2009 és Broughton 2010.

²⁸⁶ Slavić és Davies 2017. Ide lehet sorolni például a számjelzetes korlátozottan közös alosztások és a szempont szerinti alosztások megszüntetését, a $\approx 0/9$ típusú párhuzamos továbbosztások számának csökkentését vagy külső jelzetre cserélését, a szintézisek következetes megszüntetését stb. (ld. „Cancellations to the UDC.” és „Major Changes to the UDC 1993-2013”.)

²⁸⁷ Slavić-Overfield 2005: 79-85.

²⁸⁸ Berners-Lee, Hendler és Lassila 2001.

²⁸⁹ Harpey és Tillett 2006.

²⁹⁰ Dudás 2013.

²⁹¹ Mitchell, Zeng és Žumer 2011.

kapcsolatos eredmények eddig elsősorban a DDC kapcsán lettek publikálva.²⁹² Meg kell jegyezni azt is, hogy az ETO szemantikus weben való használatát egyelőre a táblázatok korlátozott elérhetősége is akadályozza.

Magyarországon a kilencvenes évek óta az ETO gépi alkalmazását többféle szempontból vizsgálták. A táblázatok gépesítése kapcsán meg kell említeni a 2005-ös kiadás²⁹³, illetve a hozzá tartozó konkordanciatáblázatok elkészítését.²⁹⁴ A legújabb táblázatok fordításának és online közreadásának előkészületei 2016-ban kezdődtek el.²⁹⁵

A megvalósítás módja miatt is említésre érdemes az Országos Széchényi Könyvtár NEKTÁR adatbázisa, amelyben az ETO-jelzetek KWIC rendszerű indexe biztosította a visszakereshetőséget.²⁹⁶ Emellett folytak és folynak kutatások a *jelzetek szemantikailag önálló részjelzetei szerint történő kereshetőségével*²⁹⁷, a *komplex jelzetek automatikus részekre bontásával*²⁹⁸, *reprezentációjával és analízisével*²⁹⁹, az *osztályozás vizualizációjával*³⁰⁰, illetve a *tezausz deszkriptorok és az osztályozási kódok konvertálásával* kapcsolatban is.³⁰¹ A közkönyvtárak számára készült *központi, általános tezausz (Köztausz)* deszkriptorai tartalmazzák az őket leíró ETO-jelzeteket is³⁰², és az ETO-jelzetek szöveges feloldásai alapján történő kereshetőség tételére is történtek kísérletek.³⁰³

A három klasszikus osztályozási rendszer közül az ETO érhető el a legrégebben a weben keresztül: az online táblázatok első (akkor még a Brit Szabványügyi Hivatal által üzemeltetett) teljes verzióját 2001-ben nyitották meg.³⁰⁴ Ennek a helyét 2013-ban vette át az ETO Konzorcium által karbantartott **UDC Online**, amely a mindenkori legfrissebb MRF alapján van frissítve, így tartalma mindig tükrözi a legújabb hivatalosan is kiadott módosításokat.³⁰⁵ Ez a teljes verzió a megfelelő licenz megvásárlása vagy igénylése esetén válik elérhetővé, de hasonlóan naprakész adatokat tartalmaz a séma mintegy négy százalékát tartalmazó, ötvenhét nyelven ingyenesen elérhető **UDC Summary** is.³⁰⁶

Említettük, hogy az ETO kapcsán a szemantikus weben történő használatot korlátozza a táblázatok elérhetősége.³⁰⁷ Egyelőre a UDC Summary valamennyi osztálya rendelkezik saját URI-jal, illetve 2011.

²⁹² Mitchell, Zeng és Žumer 2011 és Mitchell, Zeng és Žumer 2012, illetve Zeng, Panzer és Salaba 2010 és Gnoli, De Santis és Pusterla 2015.

²⁹³ Barátné 2007a.

²⁹⁴ Fejős 2007, Fejős 2008a, Fejős 2008b, Fejős 2008c és Fejős 2009.

²⁹⁵ Dancs 2016.

²⁹⁶ Zöldi 1993.

²⁹⁷ Marton és Prokné 2009 és Prokné 2013.

²⁹⁸ Mándy 2013.

²⁹⁹ Piros 2014 és Piros 2015.

³⁰⁰ Barátné 2007b.

³⁰¹ Bánki 2005.

³⁰² Barátné 2009.

³⁰³ Thék 2014.

³⁰⁴ Gedam és Paradkar 2013.

³⁰⁵ „UDC Online”.

³⁰⁶ „UDC Summary”.

³⁰⁷ Piros 2014.

novembere óta elérhető SKOS formátumban is.³⁰⁸ A teljes ETO SKOS-ra konvertálása folyamatban van.³⁰⁹

4.2 A programokkal szemben támasztott elvárások

Az ETO gépesítésének három különálló, funkcionálisan részben elkülönülő fázisa tárgyalható, ezek: **a táblázatok gépesítése, az osztályozási munka és az információ visszakeresés támogatása.**³¹⁰ Az osztályozási munka segítésének előfeltétele a hatékony visszakeresés, illetve böngészés is, hiszen az osztályozónak tudnia kell választania a már létező jelzetek között, a szinonimákat kiküszöbölő minél nagyobb kontrol, illetve a költségek csökkentése érdekében. Ezt pedig mindenképpen befolyásolják a táblázat gépesítésének a kérdései is.³¹¹

Slavić felsorolása szerint „hierarchikus rendszerezésre és részletes indexelésre egyaránt alkalmas rendszerként **az ETO-t elsősorban az alábbi területeken alkalmazzák:**

- i) objektumok / források szisztematikus rendezése:
 - a. dokumentumok polcon történő elrendezése
 - b. online információs források rendszerezése / megjelenítése
 - c. objektumok (reáliák) rendszerezése
- ii) szurrogátumok szisztematikus elrendezése információ visszakeresés és felfedezés céljából:
 - a. bibliográfiai adatbázisokban, bibliográfiákban, könyvtári katalógusokban
 - b. digitális gyűjteményekben, (beágyazott metaadattal ellátott) teljes szöveges adatbázisokban
- iii) értesítő szolgáltatások / szelektív információterjesztés (SDI) / kurrens értesítő közlőnyők
- iv) egyéb kontrollált szótárak létrehozása vagy menedzselése (összekapcsolása).³¹²

Rigby a fentiek mellett még megemlíti a nemzeti és nemzetközi dokumentációs és könyvtári hálózatokon belüli kompatibilitás biztosítását, illetve az adat nyilvántartási rendszereket is.³¹³

A felsorolt szolgáltatások bármelyikének a támogatása legyen is a cél, az osztályozási rendszernek minden esetben három feladatot kell ellátnia: **feltárás, tárolás és visszakeresés**.³¹⁴ A gépesítés során *a három feladatot egyforma prioritással kell kezelni*, egyik sem szorulhat háttérbe a többi mellett.

Ha a funkcionális részek oldaláról vizsgáljuk, azt láthatjuk, hogy az osztályozási munka támogatása során elsősorban a feltárás, illetve az elkészült jelzet tárolása áll a középpontban, de ahogyan már említettük, itt is szükség van a visszakeresés hatékony használatára.

³⁰⁸ „UDC Summary Linked Data”.

³⁰⁹ Slavić 2012.

³¹⁰ Slavić-Overfield 2005: 13-15.

³¹¹ Közvetve mindenképpen befolyásolják a táblázatok kiadásán vagy online elérésén keresztül, de a természetes nyelvű keresésen, a szemantikai kapcsolatok megfelelő feltárásán stb. keresztül közvetlenül is, ha az alkalmazott szoftver használja a Master Reference File-t.

³¹² Slavić-Overfield 2005: 13-15.

³¹³ Rigby 1974: 14-15.

³¹⁴ Babiczky és Barátné 1998: 165.

Az információ visszakeresés az eltárolt jelzetekből kiindulva használja az osztályozási rendszer visszakeresést biztosító lehetőségeit. Ide tartoznak a rendezés kérdései is, melyek a böngészést, a dokumentumok elrendezését, illetve a kereséskor kapott rekordok sorrendjét határozzák meg.

Ahogy már említettük, a táblázatok gépesítése mindegyik feladatkörre hatással van, hiszen közvetve vagy közvetlenül, de ez biztosítja magának a feltáró eszköznek az elérhetőségét.³¹⁵

4.3 A táblázatok gépesítése

A dolgozat korábbi részeiben volt szó arról, hogy a táblázatok gépesítésének az ötlete már az 1960-as években felmerült.³¹⁶ Az ETO egyes részeinek gépesített karbantartására a hatvanas évek közepén történtek az első kísérletek, ahogyan azok indexének gépi úton történő előállítására is.³¹⁷ Az utóbbival kapcsolatos kutatások a hetvenes években is folytatódtak, többek között Magyarországon is.³¹⁸

A nyolcvanas években először a karbantartás struktúrája változott meg olyan irányban, amely a korábbinál jobban támogatta a gépesítést, majd a már szintén említett ETO munkacsoport javaslatai alapján rögzítették a standard verzió létrehozásának alapelveit.³¹⁹ Ezek alapján 1992-ben Gerhard Riesthuis és David Strachan elkészítették a Master Reference File-t, illetve az azt tartalmazó CDS/ISIS³²⁰ adatbázist, a Brit Szabványügyi Hivatal (British Standard Institute, BSI) 1985-ös középkiadása alapján, a köztes időszakban megjelent Extensions and Corrections köteteiben közölt módosítások, illetve több nemzeti kiadás figyelembe vételével.³²¹

4.3.1 Az ETO mesterfájl („UDC Master Reference File”)

Az adatbázis angol nyelvű, központi verziójának létrehozása óta *folyamatosan frissítve van* az Extensions and Correctionsben közreadott módosítások alapján. 2007-ben a CDS/ISIS-t relációs adatbázisra cserélték (a jelenlegi verzió MySQL szerveren fut), de a teljes tartalom továbbra is *elérhető az alábbi formátumokban*:

- nyomtatott formátumban, publikálási célra (egyszerű, jelölőkkel ellátott szöveggént)
- CDS/ISIS adatbázisként (ISO2709³²² alapú formátumban)

³¹⁵ Jelen dolgozat csak érintőlegesen foglalkozik a felhasználói felületek, illetve a szűken vett megjelenítés kérdéseivel, elsősorban az azok háttérül szolgáló feladatokat és technikai megoldásokat vizsgálja. A felhasználói felületeket részletesen tárgyalja Slavić (Slavić-Overfield 2005: 137-164 és Slavić 2006). Barátné Hajdu Ágnes (2006), az elméleti háttér mellett magyarországi és külföldi példákon, Schallier pedig a Leuveni Egyetem (Belgium) könyvtárában implementált keresőfelület bemutatásán keresztül vizsgálja az ETO vizuális megjelenítésének kérdéseit. A keresési interfészek és az információkeresési viselkedés kapcsolatáról Bates (1989) ír részletesen.

³¹⁶ Slavić-Overfield 2005: 16.

³¹⁷ „Az ETO kiadásának gépesítése” 1965, Rigby 1974: 8-9.

³¹⁸ Orosz 1977.

³¹⁹ Slavić, Cordeiro és Riesthuis 2008.

³²⁰ A CDS/ISIS szöveges információ tárolására és visszakeresésére készült szoftver, melyet az UNESCO 1985 óta tart karban, elsősorban a fejlődő országok technológiai fejlődésének az elősegítése érdekében („CDS/ISIS Database Software”).

³²¹ A munka során a legfrissebb japán, magyar, szerb-horvát és francia középkiadásokat használták fel (Strachan és Oomes 2001).

³²² A bibliográfiai leírások és egyéb rekordok adatszeréhez (és nem a rendszereken belüli feldolgozásához) használható formátumok kialakítását szabályozó ISO2709 szabvány (illetve amerikai megfelelője, az ANSI

- platformfüggetlen, szoftveres feldolgozásra alkalmas formátumokban:
 - kiterjesztett, jelölőkkel ellátott szöveggént
 - szöveges XML-ként

A szöveges formátumok CP850 és UTF8 kódolással érhetőek el.³²³

A CDS/ISIS adatbázisban *minden fogalmat, azaz osztályt egy rekord reprezentál*, amely tartalmaz minden olyan adatot, amely szükséges lehet a rekord (és így a táblázatok) kinyomtatásához vagy megjelenítéséhez. Ezen felül olyan információk is megtalálhatók benne, melyek elsősorban a revízióhoz és az adminisztrációhoz szükségesek.³²⁴ A mezők listáját a 3. táblázat tartalmazza.³²⁵

Mezőkód	Leírás	
001	ETO szám	Az osztályozás adatmezői
002	Táblázat (kódok)	
003	Korlátozottan közös alosztások típusa (amennyiben elérhető) (kódok)	
004	Kombinációs lehetőségek (kódok)	
005	Közvetlen hierarchikus felső kapcsolat	
006	Korlátozottan közös alosztások	
013	Az alkalmazással kapcsolatos megjegyzés (111. mező) által bevezetett korlátozottan közös alosztás típusa	
100	Leírás: definíció	
105	Leírás: szöveges példák	
110	Scope note	
111	Alkalmazással kapcsolatos megjegyzés	
115	Példák a kombinációra	
125	Hivatkozások	
126	Természetes nyelvű indexkifejezések	
	Speciális karakterek használata	Kiadáshoz szükséges adatmezők
	Szerkesztői kommentárok	
901	Bevezetés dátuma	Adminisztrációs adatmezők ³²⁶
903	Bevezetés forrása	
904	Megjegyzések a bevezetéssel kapcsolatosan	
911	Megszüntetés időpontja	
912	A jelzet helyett bevezetett jelzet(ek)	
913	Megszüntetés forrása	

Z39.2) a könyvtári katalógusok közötti import és export gyakorlatilag kötelezőnek elfogadott alapja, a MARC formátumok is ennek megfelelően lettek kialakítva (ld. UNESCO 2015: 55 és „ISO 2709”).

³²³ „UDC Master Reference File”.

³²⁴ Slavić-Overfield 2005: 223-224.

³²⁵ a táblázat több forrás (Strachan és Oomes 2001, Riesthuis 2003 és Slavić-Overfield 2005: 224-226) adatait összegzi.

³²⁶ Az adminisztrációs és a rekord kezeléséhez szükséges adatmezők nem jelennek meg az exportokban.

914	Megjegyzések a megszüntetéssel kapcsolatosan	
921	Legutóbbi revízió időpontja	
922	Revízió specifikációja	
923	Revízió forrása	
924	Megjegyzések a revízióval kapcsolatosan	
925	Revíziótörténet	
951	Csak indexelt jelzetek (megjegyzésekben vagy példákban említett, máshogyan nem kereshető számok)	
952	Speciális karakterekre vonatkozó megjegyzések	
955	Szerkesztői megjegyzések és kommentárok	
957	Adminisztrátori megjegyzések	
958	A következő Extensions and Correctionsre vonatkozó információ	
999	Ideiglenes információ	
	Adatbázis neve	A rekord kezeléséhez szükséges adatok
	Rekord száma	

3. táblázat A Master Reference File adatmezői

A táblázatból hiányoznak a párhuzamos továbbosztást jelölő mezők, amelyek már korábban is többször változtak és jelenleg is átdolgozás alatt vannak, a könnyebb automatizálás érdekében. A tervek szerint új mezők fognak bekerülni a jelzetkapcsolási példák, a jelzet és a fogalom története, illetve - a jelzetösszekapcsolások automatizálását megkönnyítő - táblázat és korlátozottan közös alosztás kódok számára.³²⁷ A jelzetösszekapcsolások szempontjából különösen fontos a 004-es mező, mely a részjelzetek és a műveletek beazonosításához és a komplex jelzetek szabályoknak megfelelő besorolásához szükséges betűjeleket tartalmazza. 2003 óta valamennyi szöveges mezőben lehetőség van tetszőleges számú fordítás megadására almezők formájában, az adatbázis többnyelvű karbantarthatósága és használata érdekében.³²⁸

Az MRF ára a felhasználás céljától és formájától függ: különböző **licenszek** érhetők el a kiadások előkészítéséhez, oktatási, illetve egyéni és intézeti kutatási célokra, továbbá szoftverekben történő felhasználásra.³²⁹

4.3.2 Az ETO Online elérése

Az ETO Konzorcium feladatának tekinti a táblázatok használatának elősegítését éppúgy, ahogy a rendszer népszerűsítését és használatának oktatását is. Ennek érdekében alakították ki az ETO teljes struktúráját felölelő, online felületen keresztül elérhető többnyelvű adatbázist, a **UDC Online**-t és a rendszer kisebb részét tartalmazó, ingyen hozzáférhető **UDC Summary**-t.

³²⁷ A tervezett fejlesztésekre vonatkozó információk Dr. Aida Slavićtól származnak (magán e-mail formájában).

³²⁸ Riesthuis 2003.

³²⁹ „UDC MRF Use Licenses”. Az MRF licenszek nem jogosítanak fel a táblázatok kiadására, a kiadói licenszekhez ld. „UDC Publishing Licenses”.

A Konzorcium által közreadott és karbantartott eszközök, illetve azok fordításai mellett az ETO-nak *léteznek egyéb online elérhető verziói is*, például a svéd elektronikus kiadás³³⁰, illetve a tervek szerint ezt a kört fogja bővíteni az új magyar ETO fordítás is.³³¹

4.3.2.1 UDC Online

A három „nagy” osztályozási rendszerből az ETO volt az első, amely online elérhetővé vált, 2001-ben, majd valamivel később követte az LCC majd a DDC is.³³² Eredetileg a British Standard Institute tartotta karban, és állománya a BSI Standard kiadásának anyagát tükrözte. Miután a BSI konzorciumi tagsága 2012-ben megszűnt, az általuk karbantartott online ETO 2013. októberében elérhetetlenné vált, és helyét az ETO Konzorcium által kifejlesztett és karbantartott új UDC Online vette át.³³³

A UDC Online az ETO teljes, standard, aktualizált változatához enged hozzáférést: az állománya mindig egyeztetve van az MRF legújabb hivatalosan kiadott változatával. A szolgáltatás *az angol mellett holland, cseh, horvát, spanyol, szlovén és francia, illetve (tesztüzemben) német nyelven érhető el*. A fordítások alapját a legfrissebb nemzeti nyomtatott, digitális vagy online kiadások képezték.

A szolgáltatás a táblázatok hierarchikus, böngészhető formában teszi elérhetővé. A böngészés mellett lehetőség van a számok, illetve természetes nyelvű leírásaikban való keresésre. A láncindex is az osztályok természetes nyelvű hozzáférést könnyíti meg. Az MRF-ből a számok mellett a leírások, példák, és az alkalmazással kapcsolatos megjegyzések jeleníthetők meg, és a keresésekbe bevonhatók a csak példákban előforduló, illetve már megszűnt számok is.

A táblázatok mellett elérhető egy ETO szám összeállító („number builder”) alkalmazás is, melynek segítségével összetett, esetleg egyszerűbb komplex és agglomerált tárgyköröket képezhetünk. Az ETO-jelzet elemző alkalmazás, amely az MRF-ben szereplő táblázatazonosítók alapján lesz képes felbontani a jelzeteket, egyelőre még nem elérhető.³³⁴

A UDC Online éves **licensszel** használható, egyéni és csoportos előfizetéssel (az ár függ a felhasználók számától). Oktatási és kutatási célokra ingyenes licenzek is igényelhetők.³³⁵

4.3.2.2 UDC Summary

A többnyelvű UDC Summary³³⁶ egy, *az ETO teljes állományának csaknem négy százalékát tartalmazó ingyenesen*³³⁷ *elérhető adatbázis*. A szelekció a táblázatok valamennyi részét felöleli, és a rekordok tükrözik az MRF adatbázis szerkezetét, így tartalmazznak minden adatot, amely ott is elérhető. A két

³³⁰ „Universella Decimalklassifikationen”.

³³¹ A Konzorcium kezdeményezésére a UDC Online magyar fordításának elkészítése is felmerült, az Országos Széchényi Könyvtár azonban úgy döntött, inkább a táblázatoknak a Relex névtérkezelő alkalmazásba való feltöltését támogatja, a Köztarusszal (és egyéb teauruszokkal) való együttes használat elősegítése érdekében (Dancs 2016). A Relex névtérkezelőről és használatáról ld. Ungváry (2013).

³³² Gedam és Paradkar 2013.

³³³ Slavić 2013.

³³⁴ „UDC Online”.

³³⁵ „UDC Online English – Prices”.

³³⁶ „UDC Summary”.

³³⁷ Pontosabban a Creative Commons Nevezd meg! - Így add tovább! 3.0 Nemzetközi Licenz („CC BY-SA 3.0.”) szerint.

adatbázis évente *szinkronizálva van*, annak érdekében, hogy a kivonatban is mindig a legfrissebb adatok szerepeljenek.³³⁸

A UDC Summary ötlete 2008 végén merült fel³³⁹, és első verziója 2009 végén vált nyilvánosan elérhetővé.³⁴⁰ Az eredetileg kétezer osztály mostanra kétezer hatszázra bővült, a tizenhárom nyelven, amelyen az induláskor elérhetők voltak pedig ötvenhétre.³⁴¹ *A fordítások között olyan nyelvek is szerepelnek, amelyeken az ETO semmilyen formában nem volt korábban elérhető.* Ezt megelőzően soha nem volt példa arra, hogy a táblázatok akármilyen formában ilyen sok nyelven teljesen ingyenesen elérhetők legyenek – még ha a Summary eredetileg nem is arra lett szánva, hogy indexeléshez használják, hanem az osztályozás lehetőségeinek a demonstrálására, illetve kutatási, oktatási célokra.³⁴²

A felhasználóbarát, valamennyi nyelven való böngészést lehetővé tévő felület mellett a teljes adatbázis – szintén az összes nyelven – *elérhető linked data-ként is*: minden egyes osztály saját URI-val rendelkezik, amely tetszőlegesen hivatkozható más tudásszervező rendszerekből, programokból és szolgáltatásokból. Az azonosítók által meghatározott címeken megtalálhatók az adott osztály jele és leírása mellett annak hierarchikus szemantikai kapcsolatai is.³⁴³ A linked data verzió *elérhető SKOS formátumban* (egyetlen letölthető RDF/XML file-ban, mely az összes nyelvet tartalmazza). Ezen kívül e-mailben igényelhetők egy- és többnyelvű export file-ok szöveges file, MS Word és Excel, valamint XML formátumokban is.³⁴⁴

4.4 Az információ visszakeresés támogatása

A bevezető fejezetben láthattuk, hogy az információkeresésben az osztályozás mint tevékenység a tartalmak visszakereshető módon történő kifejezését jelenti egy információkereső nyelv, jelen esetben az Egyetemes Tizedes Osztályozás használatával. Az **információkeresés klasszikus modellje** szerint a kifejezést, azaz az információfeltárás eredményét (tehát az előállított, illetve a szöveghez hozzárendelt ETO-jelzetek összességét) **dokumentumképnek** nevezzük. A keresés során a kereső szintén az információkereső nyelven fejezi ki a számára szükséges információt, az így kialakult meghatározást, tehát a kérdéssel kapcsolatos ismérvek összességét nevezzük **keresőképnek**. A keresőkép és a dokumentumkép összehasonlítása a **profilszerkesztés**.³⁴⁵

Ha az ETO kapcsán a – klasszikus - *információkeresés számítógépes támogatásáról* beszélünk, az a *keresőkép összeállításának és a profilszerkesztésnek a támogatását*, a megfelelő felhasználói felület és kereső algoritmus kidolgozását és megvalósítását jelenti. *A technológiai hátteret mindkét esetben*

³³⁸ „What is UDC Summary?”

³³⁹ Slavić 2008.

³⁴⁰ Slavić 2009a és Slavić 2009b.

³⁴¹ A magyar nyelvű fordítás Barátné Dr. Hajdu Ágnes munkája. A fordítások, illetve fordítók listájához ld. „UDC Summary Translations”.

³⁴² Slavić et. al. 2009. A gyakorlatban már a könyvtárak is elkezdtek használni bizonyos munkafolyamatokra, vagy ha semmilyen formában nincs hozzáférésük a teljes táblázatokhoz. Ezt például a legutóbbi észlt felmérés is mutatja (Makke 2015), de a teljes szlovén ETO fordítást megelőzően a szlovén könyvtárak is egységesen a UDC Summaryt használták (Rozman 2015).

³⁴³ „UDC Summary Linked Data”. Az URI-eket a nyelvi azonosítókkal kiegészítve az egyes osztályok lefordított változatai érhetők el. Pl. a <http://udcdata.info/025636> URI a Számelmélet osztály angol verzióját jelenti (vagy böngészőben nézve azt a nyelvet, amelyik ki lett választva a menüből), a <http://udcdata.info/025636/hu> pedig a magyar nyelvűt.

³⁴⁴ „UDC Summary Exports”.

³⁴⁵ Babiczky és Barátné 1998: 37, Hutchins 2001: 466.

az osztályozási jelzetek és hozzáférési pontjaik leírására, illetve eltárolására használt formátumok jelentik.

Az ETO, ugyan dokumentációs céllal, de egy gigantikus cédulakatalógus számára készült, első kiadása is a katalógus használati útmutatójának volt a része, melyhez elsősorban a tájékoztatást végző szakemberek értek hozzá. Később a könyvtári katalógusokban és a bibliográfiákban – akárcsak más osztályozási rendszerek esetében – *sem az információ végső felhasználóinak a feladata volt a használata*, hanem a köztük és a katalógus között közvetítő *könyvtárosoké és információszakembereké*, akik rendelkeztek a nyelv hatékony használatához szükséges ismeretekkel.

A kereséshez hasonlóan az osztályozás sem a dokumentumok szellemi előállítói által történt, hanem osztályozási szakemberek által. Sokáig a dokumentumközpontú megközelítés miatt a relatíve körülményes és lassú jelzetezés sem okozott problémát.

Ez a szemlélet *a dokumentáció elterjedésével* kezdett el megváltozni, majd a folyamat az ötvenes években gyorsult fel és a hatvanas évek végére érte el a tetőpontját. A szakirodalmi tájékoztatás és a gépesítés szükségessé tette olcsóbb és gyorsabb megoldások kidolgozását. Ekkor *vált külön az osztályozás és információkeresés fogalma*, terjedt el a *posztkoordinált szemlélet*, illetve kerültek előtérbe a *természetes nyelven alapuló keresési eljárások*.³⁴⁶

A természetes nyelvű hozzáférés a klasszikus osztályozási rendszereket megtartó katalógusokban is előtérbe került. Az 1980-as években *felhasználók számára is hozzáférhető OPAC-ok* értelemszerűen magukkal hozták az igényt az általuk is könnyen használható tárgyi visszakeresésre.³⁴⁷ Az ETO táblázatok a felhasználó számára általában nem állnak rendelkezésre még nyomtatott formában sem. Az átlagos felhasználó esetében az osztályozási rendszer szabályainak az ismerete egyébként sem várható el, így *a jelzetek természetes nyelvű kereshetősége és böngészhetősége alapvető elvárás*. Ennek hiányában a felhasználók által ETO alapján végzett keresések száma a nullához közelít.

Ezt a hatást az OPAC-ok weben keresztüli – tehát a könyvtártól fizikailag is távol lévő helyen is elérhető módon – közzététele csak növelte, hiszen a katalógus távoli használata során a felhasználó a könyvtáros szakember segítségét és a helyben esetleg elérhető táblázatokat is mellőzni kénytelen. Az utóbbi két évtizedben ehhez hozzájárult *a keresési szokások megváltozása*, az, hogy a felhasználók elsődlegesen a webes keresőket és forrásokat részesítik előnyben, azokon keresztül keresnek akár könyvtári forrásokat is. A jelenlegi trend pedig *az adatközpontú szemléleté*, amelynek lényege, hogy a gyűjtemények a szemantikus weben, RDF állításokkal leírva adják közre a gyűjteményüket leíró metaadatokat.³⁴⁸

A fentiek vezettek ahhoz, hogy az osztályozási rendszerek használatával kapcsolatos paradigma megváltozzon. **A jelenleg elterjedt szemlélet** szerint:

- *az osztályozási rendszer elsődleges szerepe a hierarchikus tárgyi böngészés támogatása*, mivel a hierarchikus kapcsolatok inkluzív követése ezekben a rendszerekben technikailag könnyebben megoldható, mint a tezauruszok esetében

³⁴⁶ Ungváry és Orbán 2001a: 328, 339 és Ungváry és Orbán 2001b: 27-28.

³⁴⁷ Bates 1989.

³⁴⁸ Horváth 2015.

- a végfelhasználó nem találkozhat az osztályozási jelzetekkel a keresés során, *lehetővé kell tenni a számára a természetes nyelvű kifejezések szerinti keresést.*³⁴⁹

A két pont bizonyos szinten ellentmond egymásnak: a mesterséges nyelvű jelzetek teljes kizárása ellen szól az, hogy a hierarchikus böngészés során éppen a mesterséges jelzetek teszik áttekinthetőbbé a hierarchiát, szemben a teauruszokkal, ahol a deszkriptorok maguk nem tükrözik az alá-fölé rendeltségi viszonyokat.

A mesterséges nyelvű kifejezések teljes elrejtése azért sem helyes, mert az iteratív facettás keresés³⁵⁰, illetve az olyan keresési stratégiák, mint a hólabdakeresés vagy a relevancia-visszacsatolás³⁵¹ esetén is *hasznos lehet, ha a felhasználó pontosan látja az indexkifejezéseket.* De ahogy a korábbi fejezetekben láthattuk, mivel a jelzet a dokumentum szurrogátuma, amely annak fogalmi viszonyait is tükrözi, hozzáértő felhasználó számára *a relevancia eldöntésében is nagy szerepe lehet,* illetve a precizitás is növelhető, ha a leírásokban esetlegesen ismétlődő kifejezések által okozott zaj elkerülése végett inkább az osztályozási jelzeteket használjuk.³⁵² A helyes megközelítés ezért az, ha *nem várjuk el a felhasználótól az osztályozás használatát, de, amennyiben igényli, lehetővé tesszük a számára.*

Az első pont indoklása szerint a hierarchikus kapcsolatok inkluzív követése ezekben a rendszerekben technikailag könnyebben megoldható. Ez enumerikus osztályozási rendszerek esetében nyilvánvalónak látszik, ugyanakkor nem szabad megfeledkeznünk a hierarchikus felépítésnek a már korábban említett hátrányairól és az ETO következtetlenségeiről sem.

A **hierarchikus böngészés** facettás és analitikus-szintetikus rendszerek esetében több kérdést is felvet. Az egyértelműnek tűnik, hogy a szintetikus jelzetek egyszerű elemei – tehát az egyszerű táblázati számok és intervallumok – annak hierarchikus felső kapcsolatai, ezért, amennyiben sikerül őket beazonosítani, a böngészés és az inkluzív keresés során a közöttük lévő hierarchikus kapcsolatok felhasználhatók. A kombinált jelzetek összetett részjelzeteinek kérdése már bonyolultabb.³⁵³ Először is, mint láttuk, azok jelentését a részek mellett a szintaktikai kapcsolatok is meghatározzák, így ezeket is figyelembe kell venni a hierarchikus kapcsolatok meghatározása, illetve a rajtuk keresztül történő keresés során, ehhez ugyanis a jelzetek szerkezetét is ki kell elemezni, és össze kell hasonlítani őket.³⁵⁴

A hierarchikus böngészés mellett meg kell említeni az osztályozási jelzetek háttérben való használatának nagy előnyét is, ami a **több nyelvű visszakereshetőség** lehetősége: amennyiben a táblázatoknak, illetve a használt jelzetek leírásainak különböző fordításai is rendelkezésre állnak, ezeken keresztül a rendszer automatikusan biztosíthatja a tárgyi hozzáférést minden nyelven anélkül, hogy a feltárást végző szakembereknek ismerniük kellene ezeket a nyelveket. Hasonlóan lehetséges

³⁴⁹ Slavić-Overfield 2005: 32-34.

³⁵⁰ Frické 2017.

³⁵¹ Ld. Ungváry és Vajda 2002: 74-75.

³⁵² Ugyanakkor figyelembe kell venni, hogy ha a rendszer nem tudja megfelelően kezelni a monohierarchiából adódó fogalmi szóródást, illetve a prekoordinált kifejezések többalakúságát, az a visszahívás csökkenéséhez vezethet.

³⁵³ Ilyen lehet például csoportképzés esetén a szögletes zárójelen belüli szám, egy olyan jelzet, amely az eredeti által tartalmazott alosztásoknak csak egy valódi részhalmazát tartalmazza, vagy esetleg az abban szereplő intervallum részintervallumát.

³⁵⁴ A prekoordinált ETO-jelzetek hasonlóságának mérésére Lula és Cieraszewska (2017) dolgozott ki módszereket.

akár a meglévő tárgyi indexekhez új nyelvek automatikus hozzáadása is, érdemi fordítói munka nélkül.³⁵⁵

4.4.1 A jelzetek természetes nyelvű visszakereshetősége

Már a hetvenes években is többen vizsgálták, hogy a táblázatokban szereplő leírások alapján hogyan állíthatók elő automatikusan az egyszerű ETO számok betűjeles indexei. A konklúzió minden esetben az volt, hogy *a tárgymutató önmagában a számok definícióból teljesen mechanikus módon nem készíthető el*. A megfelelő előkészítést követően azonban Magyarországon és a világ más részein is születtek sikeres megoldások a problémára.³⁵⁶ Mándy Gábor ugyanakkor arra mutatott rá, hogy a magyar nyelvű nyomtatott kiadások betűrendes mutatói a mindennapi nyelvtől való távol állásuk miatt nem alkalmasak ilyen célokra.³⁵⁷

Láthattuk, hogy a *Master Reference File*-ban a leírások mellett már külön mezők vannak a leírásokban nem szereplő keresőkifejezések számára, melyek automatikusan vagy akár manuálisan is hozzárendelhetők a jelzetekhez, amennyiben az adatbázis hozzáférhető. Ugyanakkor a gyűjtemény egyedi szempontjai alapján egyéb keresőkifejezések manuális hozzáadására is szükség lehet.

A komplex, összetett és agglomerált jelzetek esetén *külön kell választanunk a jelzet által jelölt újonnan létrejött fogalom önálló egységként való leírását és a részjelzetek leírásait*.³⁵⁸ Amennyiben az első típusú leírások rendelkezésre állnak – tehát a prekoordinált jelzetek önálló leírásai is el vannak tárolva az adatbázisban –, ezek felhasználásával kapcsolatban ugyanazok a kérdések merülnek fel, mint az egyszerű jelzetek betűjeles indexei esetén.

A természetes nyelvű visszakeresés legegyszerűbb módja az egyszerű és összetett ETO-jelzetek természetes nyelvű leírásának és a hozzáadott kereső kifejezéseknek az eltárolása az osztályozás során, majd ezeknek a jelzeteknek az elérhetővé tétele.

Ennél valamivel kifinomultabb módszer *a jelzetek részekre bontása, majd a részek leírásai és indexkifejezései szerinti visszakeresése*. A részekhez a leírások hozzárendelése történhet az osztályozó által manuálisan, illetve a szoftver által automatikusan, amennyiben tartalmazza az ETO táblázatait.³⁵⁹ Amennyiben maga a jelzetépítés is a Master Reference File alapján történik, mind a jelzetek részekre bontása, mind a természetes nyelvű visszakereshetőség nagy hatékonysággal elvégezhető az abban szereplő kifejezések alapján, ahogyan azt Gerhard Riesthuis kutatásaiban megmutatta.³⁶⁰

Az utóbbi időkben szakmai konszenzus, hogy a tárgyi **authority** fájlok használata az osztályozási jelzetek használatának kielégítő módja, különös tekintettel a természetes nyelvű visszakeresésre. A

³⁵⁵ Pontosabban a fordítói munkát ekkor a táblázatok kiadója, illetve az authority fájl karbantartója végzi el.

³⁵⁶ Orosz 1977.

³⁵⁷ Mándy 2011.

³⁵⁸ Ahogy a kompozicionalitás kapcsán már említve volt, egy összetett kifejezés jelentését nemcsak az elemei, hanem a struktúrája is meghatározza (Gendler-Szabó 2004). Így a prekoordináció során létrejövő új fogalom nem csak örökli a részek jelentését, hanem önálló jelentéssel rendelkezik (Slavić és Isaac 2009). Szintén láttunk példákat arra, hogyan befolyásolják a jelzet jelentését a szintaktikai relációk, illetve a jelzetelemek kapcsolódási módja. Ezeket a különbségeket, amelyeket Slavić és Cordeiro „az adatelemek függetlenségének és integritásának” nevez (Slavić és Cordeiro 2005), figyelembe kell venni a visszakeresés során.

³⁵⁹ Ahogyan már volt szó róla, az ETO táblázatok nem ingyenesek, szoftvereken keresztül történő közreadásuk csak a megfelelő licenz birtokában lehetséges.

³⁶⁰ Riesthuis 1997 és Riesthuis 1999.

tárgyi authority fájlok, illetve ETO-t használó tezauruszok azonban abban az esetben is megkönnyíthetik a minőségi és következetes osztályozási munkát, ha maga a katalógus nem alkalmaz authority controlt. Magyarországon például a Köztaurusz ingyenesen elérhető, és jól használható ilyen célra.³⁶¹

Az authority fájlok alkalmazásának egyik előnye, hogy komplex jelzeteket is tartalmazhatnak, hátrányuk, hogy *általában nem a teljes ETO-hoz biztosítanak hozzáférést*.³⁶² Ez főleg a Köztauruszhoz hasonló tezauruszok esetén igaz, amelyek csak kiegészítésként tartalmazzák az osztályozási kódokat. Az ilyen típusú tezauruszoknál egyszerre előny és hátrány az is, hogy a *tudásrendszerezésük nem feltétlen az osztályozás rendszerező elveit követi*.

A megoldás másik nagy haszna, hogy a rekordok nemcsak az osztályozási jelzeteket és leírásaikat tartalmazhatják, de egyéb meta- és szemantikai adatokat is, például a jelzetek kapcsolataira, konkordanciájára, felhasználási módjaira vonatkozóan. Ugyanakkor arról is konszenzus van, hogy az authority control, legalábbis a file karbantartása nagyon *költséges eljárás*, talán a legköltségesebb a katalógizálás összes munkafolyamata közül.³⁶³ Ezért is *fontos kérdés, hogy mennyire automatizálható*, illetve támogatható számítógéppel a belépési pontok, tehát a visszakereshetővé teendő részjelzetek meghatározása.³⁶⁴

A **belépési pontok meghatározása** akkor is fontos kérdés, ha a jelzettelés intézményi szinten történik, és az osztályozói munka eredményeként újonnan létrejövő jelzeteket szeretnénk természetes nyelvű kifejezéseken keresztül visszakereshetővé tenni. Ilyen esetben sarkalatos kérdés, hogy a használt OPAC milyen formában teszi lehetővé a jelzeteknek, azok részeinek és a természetes nyelvű leírásoknak az eltárolását és összekapcsolását.

Egy lehetséges megoldás a jelzetek **KWOC-index**ének építése. Ebben az esetben a rendszer a részjelzetek manuális, esetleg valamilyen szintű automatikus felismerése után a jelzetből, tehát a kontextusukból kiemelve helyezi azokat, illetve leírásaikat kereső pozícióba.

A természetes nyelvű visszakeresés biztosítható a jelzetrészek leírásai által alkotott láncok permutálásával vagy **KWIC-index** építésével. Ennek a megoldásnak előnye, hogy *egy bizonyos szintig megtartja a kontextusra vonatkozó információt*. Hátránya, hogy a felhasználható leírások száma korlátozott, illetve, hogy a permutáció miatt nagyon nagy számú rekord eltárolására is szükség lehet.

Látható, hogy *valamennyi megoldás feltételezi azt, hogy az ETO-jelzetek külön állományokban vannak eltárolva*, amelyek tartalmazzák a természetes nyelvű leírásaikat is, és a dokumentumrekordok csak hivatkoznak ezekre az állományokra. Az is látszik, hogy a fő problémát valamennyi megoldás esetén a jelzetrészek felismerése, illetve a természetes nyelvű kifejezések hozzárendelése jelenti.

Végül megemlítendő, hogy bármelyik megoldást válasszuk, *tekintettel kell lenni a rendszer változásából adódó eltérésekre* is. Ahogy már volt róla szó, az elmúlt két évtizedben az átszervezések üteme felgyorsult. Ez egyfelől öröndetes, mivel revízió során a jobb elméleti megalapozottság mellett a számítógépesítés kérdéseit is figyelembe veszik, másfelől szükségessé teszi a felhasznált verzió

³⁶¹ „OSZK Tezaurusz / Köztaurusz”.

³⁶² Ez más megközelítés szerint előny is, mivel a file-ba felvett jelzetek használata igazolt az irodalomban való előfordulásaik („literary warrant”) által (Slavić-Overfield 2005: 180).

³⁶³ Tillett 2001, illetve O’Neill, Bennett és Kammerer 2014.

³⁶⁴ Piros 2015.

figyelembe vételét és az eltérő verziók szerint készült, ugyanazokat a fogalmakat leíró jelzetek egymásnak való megfeleltetését.³⁶⁵

4.4.2 Böngészés

Osztályozási jelzetek szerinti böngészésről akkor beszélünk, amikor a böngészés *az osztályozási rendszerben explicit és implicit módon kifejezett szemantikai kapcsolatok, elsősorban az integrációs szintek szerint*³⁶⁶ *történik akár a jelzetek, akár azok természetes nyelvű megfelelői között.*³⁶⁷ Ez a típusú böngészés jobban megfelel az emberi gondolkodásmódnak, mint a leírások betűrendje szerinti, ezért, ahogy korábban már említettük, *ez az egyik legfontosabb funkció*, amiért osztályozási rendszereket használunk az információ visszakeresése során.

A böngészést gyakran szembeállítják az információkereséssel, mint alkalmi, konkrét cél nélküli keresési tevékenységet. Más megközelítés szerint a böngészés, más néven **„félíg-irányított vagy félíg-strukturált” keresés** a standard információkeresés fontos része.³⁶⁸ A már említett két keresési stratégiának is alapvető része a keresések eredményeinek böngészése a kérdés pontosítása érdekében. Amennyiben nem az általános információkeresés, hanem a könyvtári alkalmazás szempontjából vizsgáljuk, az osztályozási kódok érdekes felhasználási módja a *virtuális könyvtári polcok* összeállítása, ami szintén a böngészés egy formája.³⁶⁹

Bár a hierarchikus osztályozási rendszerek táblázati számai tükrözik a rendszer szerkezete szerinti integrációs szinteket, ez nem mondható el azok természetes leírásairól. A természetes nyelvű leírások hierarchikus böngészésére dolgozta ki Ranganathan a **láncindexelés** módszerét. Ennek lényege, hogy a jelzetek helyett azok teljes, természetes nyelvű leírásait rendezzük mechanikusan, az osztályok alá-fölé rendeltégi viszonyainak (jelzetláncoknak) megfelelően, ahogy az a 2. ábrán látható.

Bár a láncindexelés módszere általánosan elterjedt módszer az osztályozási rendszerek természetes nyelvű indexelésére, az ETO-val együtt történő alkalmazása számos problémát felvet, amelyek elsősorban a hierarchia és a jelzetalkotási szabályok korábban már említett következetlenségeiből adódnak.³⁷⁰

Az authority fájl alapján történő böngészés megvalósítására Jiri Pika hoz példát a Zürichi ETH-könyvtár Nebis rendszeréből.³⁷¹ Az ilyen megoldás előnye, hogy a jelzetek és leírásaik között egyaránt lépegethetünk, továbbá az, hogy a mellérendelő és asszociatív kapcsolatok is feltárára kerülhetnek. Ugyanakkor azt is látnunk kell, hogy *ez nem az osztályozási kódok szerinti böngészés*: ilyen esetekben

³⁶⁵ A változások figyelembe vételének szükségessége és nehézsége egyaránt jól demonstrálható a – teljes egészében átszervezett – kettes főosztályon: míg 2000. előtt a 22/28 osztályok a kereszténység különböző kérdéseit írták le és csak a maradék foglalkozott az egyéb vallásokkal, az átszervezés óta a kereszténység teljes egészében a 27 alatt van és a további finomítások valamennyi vallás esetében egységesen, facettás módon történnek. Ennek megfelelően a vallásokat leíró számok a 2000. évi és a korábbi verziók között többnyire csak az első számjegyben egyeznek meg.

³⁶⁶ Az implicit kapcsolatok a kompozit jelzetek között a szintaktikai relációk alkalmazásával létrejövő alkalmi szemantikai kapcsolatokat jelentik, amelyek tehát nincsenek lerögzítve a táblázatokban, de következnek az ott lerögzített szemantikai kapcsolatokból és a rendszer szintaktikai szabályaiból.

³⁶⁷ A természetes nyelvű böngészéskor a felhasználói körből kell következnie, hogy a leírások ábécésorrendje vagy az osztályozási rendszer számkódjai szerint történjen-e a rendezés.

³⁶⁸ Ellis 1989; hiv. Bates 1989.

³⁶⁹ Schallier 2004.

³⁷⁰ Kumar és Parameswaran 1999. A példa forrása is az idézett cikk.

³⁷¹ Pika és Pika-Biolzi 2015.

az authority fájl tulajdonképpen egy olyan tezauruszként működik, melyben a szemantikai és asszociatív kapcsolatok statikusan, a szótár karbantartója által vannak feltárva, és az osztályozás csak a megjelenítést szolgálja.³⁷²

Lánc

02	=	Könyvtárak. Könyvtárta
025	=	Könyvtári ügyvitel
025.5	=	Tájékoztatás
025.5:027	=	Tájékoztatási szolgáltatások általános könyvtárakban
025.5:027.7	=	Tájékoztatási szolgáltatások egyetemi és főiskolai könyvtárakban

Index a lánc alapján

Főiskolai könyvtárak : Tájékoztatási szolgáltatások	025.5:027.7
Általános könyvtárak : Tájékoztatási szolgáltatások	025.5:027
Tájékoztatás	025.5
Könyvtári adminisztráció	025
Könyvtárak	02

2. ábra ETO számok láncindexe

Amikor a jelzetek között navigálunk, fontos szerepe van a jelzetek helyes **sorrendjének**. Korábban már láthattuk, hogy az ETO-nak erre vonatkozóan szigorú szabályai vannak, ahogyan azt is, hogy a jelzetelemek sorrendjére vonatkozó szabályok ezzel szemben teljesen kötetlenek, éppen azért, hogy lehetővé tegyék a besorolási szabályoktól való eltérést, ha erre a gyűjtemény vagy a felhasználói kör speciális jellemzői miatt szükség lenne. A jelzetek sorba rendezése technikai megoldását már a hatvanas években megalkották. Azóta is ez a módszer van elterjedve, csak néhány apró pontatlanságát kellett korrigálni, illetve az ETO változásai miatt volt szükség néhány módosításra. Ezt a módszert részletesen tárgyaljuk egy későbbi fejezetben.

A sorba rendezés mellett a jelzetelemek sorrendjének elsősorban megjelenítésbeli jelentősége van. A sorrendrögzítéstől eltekintve *az összetett és agglomerált jelzetek elemeinek sorrendje tetszőlegesen variálható*, így igazítható a felhasználó vagy a rendszer igényeihez. Összetett jelzet esetén láthattuk, hogy az alosztások hivatkozási sorrendjére létezik javaslat, amelynek alapelve, hogy az adott fogalmat egyre inkább specializáló jelzetek kerüljenek egyre hátrébb a láncban (ezzel próbálva igazodni az emberi gondolkodáshoz). Azt is láttuk, hogy az ettől való eltérésnek szemantikai jelentősége is lehet. Akár a számok, akár a leírások között böngészünk, érdekes kérdés, hogy *amennyiben a jelzetelemek sorrendje nem hordoz jelentést – vagy ha igen, akkor annak figyelembe vételével -, instruálható-e úgy egy számítógépes rendszer, hogy a jelzeteket különböző módokon jelenítse meg*, annak érdekében, hogy a felhasználó a lehető legbiztosabban el tudja dönteni a relevanciát. Variálható-e az egyes jelzetelemek megjelenítése, illetve sorrendje a felhasználó feltételezett preferenciái és kognitív státusza szerint?

³⁷² Egy ilyen file elsősorban a belépési pontok és a kapcsolatok osztályozási kódok alapján történő automatikus meghatározásában, illetve a szótár idegen nyelvekre – az osztályozási táblázatok fordítása alapján - történő könnyebb fordíthatóságában tud többet nyújtani egy klasszikus tezaurusnál.

Az ETO számok böngészésének speciális esete a **jelzetek közötti kapcsolatok vizualizációja**, melyhez ideális alapot nyújt az osztályozás hierarchikus szerkezete. A képi vizualizációnak fontos mentális és intellektuális szerepe van az absztrakt gondolkodás, az információk feldolgozása és rögzítése során, ami az osztályozott fogalmak közötti navigációt is megkönnyítheti és kényelmesebbé teheti.³⁷³ Ilyen célra készült a MagnaView szoftver, amely az ETO Master Reference File vizualizálására volt tervezve³⁷⁴, illetve felmerült a TotalZoom technológia³⁷⁵ használata is. Lengyelországban a Varsói Egyetem Központi Könyvtára állományának tárgyköreit jelenítették meg a katalógus ETO számai alapján, a D3 JavaScript könyvtár segítségével.³⁷⁶

4.4.2.1 A jelzetek automatikus sorba rendezése

A jelzetek besorolására vonatkozó szabályok tárgyalásánál volt arról szó, hogy azt a bennük szereplő szintaktikai relációk típusa határozza meg. Az ETO története kapcsán azt is láthattuk, hogy az alkalmazott alosztási jelek és szeparátorok meghatározására a rendszer létrejötté, illetve az első revíziók során, tehát a legelső mai értelemben vett számítógépek megjelenése előtt került sor. Ennek következtében az alosztások jelei és a besorolásokra vonatkozó szabályok nincsenek összhangban egyetlen elterjedt karakterkódolási móddal sem.³⁷⁷ Így az ETO-jelzetek eredeti formájukban, egyszerű karaktersorozatként való rendezése nem a szabványos besorolási sorrendet fogja eredményezni, ahogy azt az alábbi példában láthatjuk.

Rendezés ASCII ³⁷⁸ karaktersorozatként	Rendezés az ETO besorolási szabályai szerint
73	73+75
73"17"	73/75
73(091)	73
73(439)	73:75
73(=511)	73::75
73+75	73=511.141
73.033	73(091)

³⁷³ Barátné 2006.

³⁷⁴ Van der Linden és Van Wijk 2007. A szoftver fejlesztése sajnos azóta a kevés megrendelés miatt véget ért.

³⁷⁵ Barátné 2007b.

³⁷⁶ Kowalczyk, Skonieczny és Wornbard 2015.

³⁷⁷ A számítógépek alapvetően számokkal dolgoznak, ennek megfelelően a betűket és egyéb karaktereket is számokként ábrázolják, a karakterkódolás ennek a hozzárendelésnek a meghatározását jelenti. A karaktereknek egy adott kódolás szerinti alapértelmezett sorrendjét általában a karakterekhez rendelt számok sorrendje határozza meg, ezért az arab számjegyekhez rendelt számok általában tükrözik azok sorrendjét. Az egyéb, speciális karakterek számkódjai ugyanakkor természetesen nem veszik figyelembe az ETO besorolási szabályait, sem az alapértelmezett, sem más szabványos rendezések szerint (mint például az egyes nyelvek szerinti rendezések vagy a Microsoft Excel saját rendezése).

³⁷⁸ A táblázat az ASCII karakterkódolás szerinti sorrendet mutatja, mivel ez volt az első, közös használatra tervezett karakterkódolás („ASCII History”) és az arab számok, illetve az ETO szeparátor karakterei szinte valamennyi ma elterjedt nemzeti és nemzetközi karakterkódolás szerint az ASCII kódjainak vannak megfeleltetve. Eltérést az olyan karakterek okozhatnak, amelyeknek az ASCII megfelelőjén kívül más karakterek is megfeleltethetők, mint például a szimpla és dupla aposztróf (a ' illetve " helyett használható például a ´, ` , „ , ” stb.).

73/75	73(439)
737	73(=511)
73:75	73"17"
73::75	73.033
73=511.141	737

3. ábra A szöveggént történő és az ETO besorolási szabályai szerinti rendezés összehasonlítása

A **helyes gépi rendezés** úgy érhető el, ha - a tagolásra szolgáló pontok törlése után – *a relációk jeleit az angol ábécé betűire cseréljük úgy, hogy azok betűrendje kövesse az ETO besorolási szabályait.* Az így kapott karaktersorozatok a legtöbb karakterkódolás szabványos rendezése mellett pontosan az ETO szabályainak megfelelő sorrend szerint rendezhetők.³⁷⁹

A szeparátor karakterek betűkre való lecserélésének módszerét *Robert R. Freeman* nevéhez szokás kötni, aki a hatvanas években az AIP projektben alkalmazta ezt a módszert, bár vele egy időben *Schneider* és *Koch* is alkalmazott hasonló megoldást.³⁸⁰ Magyarországon már a hetvenes években sor került az ETO-rendezés hasonló elvek szerinti implementációjára, a *BIBDOSZ* rendszeren belül.³⁸¹ A kilencvenes évek elején *Andrew Buxton* valamennyi online rendszer számára ezt a megoldást ajánlotta, a jelzetelemek szóközzel való elválasztása mellett.³⁸² *Slavić* olyan jelölők (tagek) alkalmazását javasolja, amelyek a jelzetelemek helyének és típusának azonosítása mellett azok sorba rendezésére is alkalmasak.³⁸³

A modernebb rendszerekben az adatbázis, adatfeldolgozás, illetve megjelenítés szintjeinek elkülönülése a megjelenített és a rendezésre szolgáló jelzetek szétválasztását is lehetővé teszi: így elvben az is lehetséges, hogy *amíg a megjelenítés és keresés során a rendszer az eredeti jelzeteket használja, azok rendezése a háttérben hozzájuk rendelt, a fentiek alapján kialakított segédstringek alapján történjen.*³⁸⁴ A rendezés mellett az elválasztó karakterek betűkre cserélése *a jelzetelemek automatikus felismerését is megkönnyítheti*, ha a jelzetek nem túlságosan összetettek.

A **Master Reference File** megalkotói szem előtt tartották a segédstringek gyors és szabványos megalkotását is, ezért az MRF Kombinációs lehetőségek (kódok) mezője (a CDS/ISIS exportban a 004-es mező) tartalmazza az elválasztó karakterek leváltására javasolt kódokat.³⁸⁵ Az alábbi táblázat az MRF útmutatójában szereplő kódokat mutatja be.³⁸⁶

Jele	Megnevezése	Betűjele	Példa	
+	Összekötés	a	XaY	73a75

³⁷⁹ Ennek magyarázata, hogy a legtöbb kódolás szerint az egyszerű latin (és így az angol) ábécé betűi is azon a területen vannak, amelyik megegyezik az ASCII kódjaival. Az ASCII kódolás szerint a számok megelőzik a betűket, így a karakterkódolások szabványos rendezései is ezt követik.

³⁸⁰ Orosz 1967 és Rigby 1974: i.

³⁸¹ Pethes et. al. 1976.

³⁸² Buxton 1990. A szerző annak a lehetőségét is felveti, hogy magában az ETO-ban kellene a meglévő központosítási jeleket betűkre cserélni. Ezzel szemben Dahlberg szerint elegendő lenne csak egyes központosítási jeleket lecserélni (Dahlberg 1972).

³⁸³ Slavić-Overfield 2005: 168-170.

³⁸⁴ Ezt a lehetőséget szintén Buxton (1990) vetette fel.

³⁸⁵ Tehát az MRF a Slavić által javasolt megoldást követi.

³⁸⁶ Slavić-Overfield 2005: 120.

/	Kiterjesztés	b	XbY	73 b 75
0/9	Főtáblázati szám	c	Xc	73
	Közbeékelés	d	Xd	73 d i439I17
:	Viszonyítás	e	XeY	73 e 75
::	Sorrendrögzítés	f	XfY	73 f 75
[...]	Csoportképzés			
=...	Nyelvi alosztások	g	XgY	73 g 511141
(0...)	Formai alosztások	h	XhY	73 h 91
(...)	Földrajzi alosztások	i	XiY	73 i 439
(=...)	Etnikai alosztások	k	XkY	73 k 511
"..."	Idő szerinti alosztások	l	XlY	73 l 17
*...	Külső forrásból vett jelzetek	m	XmY	630 m 30
A/Z	A/Z névalosztások	n	XnY	929 n NAP1
.00...	Szempontra szerinti alosztások	o	XoY	51 o 185
-0...	Általános ismérvek szerinti alosztások	p	XpY	324 p 51
-1/-9	Kötőjeles korlátozottan közös alosztások	q	XqY	51 q 37
.0...	Pont nullás korlátozottan közös alosztások	r	XrY	73 r 33
'...	Aposztróf korlátozottan közös alosztások	s	XsY	6281 s 113

4. ábra Javaslat az ETO-jelzetekben használandó kódokra az MRF alapján

A Freeman által javasolt kódok között szerepelt a csoportképzés (szögletes zárójel) helyettesítője is. A fenti táblázatból ez már hiányzik, mivel a csoportképzésre vonatkozó szabályok azóta megváltoztak³⁸⁷, és jelenleg erre a műveletre speciális szabályok vonatkoznak.

A fenti megközelítés előnye a sorba rendezés mellett, hogy amennyiben a szintetizált jelzetek az MRF alapján lettek összeállítva, az alosztások azonosítóit tartalmazó segédstringek könnyebben bonthatók részekre, mint az eredeti jelzetek.³⁸⁸ Hátránya, hogy *nem alkalmas többszintű jelzetek, a csoportképzés, az egymásba ágyazott alosztások stb. kezelésére.*

A jelzetek karaktersorozatként történő sorba rendezése természetesen lehetséges *a használt kódolás, illetve rendezés felüldefiniálásával* is. Ennek a megoldásnak előnye, hogy az ETO-jelzetek eredeti formájukban, segédstringek nélkül is rendezhetők. Hátránya, hogy mivel a jelzetrészek beazonosítása nem történik meg az osztályozási fázisban, a rendezéshez szükséges a jelzetelemek utólagos felismerése és azonosítása. Ha ez nem történik meg, és a rendezés csak a karakterek alapján, mechanikusan történik, az ETO következetlenségei hibás sorrendet eredményezhetnek.

³⁸⁷ vö. például Babiczky 1975: 45-46., illetve OMIKK és OSZK-KMK: 12.

³⁸⁸ Riesthuis 1999.

4.4.3 Az ETO-jelzetek szerinti keresés

A természetes nyelvű keresés mellett tehát azért célszerű lehetővé tenni az ETO számok szerinti keresést is, mert ez hatékonyabb lehet a természetes nyelvű leírások szerintinél, illetve hasznos lehet bizonyos keresési stratégiák esetén. Ez a kérdés ugyanakkor azért is fontos, mert *a természetes nyelvű keresés esetén is ugyanazok az elvek határozzák meg, hogy egy adott kifejezésre keresve az annak megfelelő ETO-jelzet mellett, azon keresztül milyen más jelzetek lesznek még visszakeresve.* Ilyen szempontból nincs különbség aközött, hogy egy ETO számra közvetlenül, vagy természetes nyelvű leírása alapján kerestünk rá.

Az ETO-jelzetek visszakeresése első ránézésre egyszerűnek tűnhet. Az egyszerű ETO számok – akár inkluzív - visszakereshetővé tétele tényleg triviális feladat, azonban, ha jobban megvizsgáljuk a kérdést, láthatjuk, hogy a több elemű számok kérdése jóval bonyolultabb.³⁸⁹

A problémát elsődlegesen a jelzetek egymásnak való megfeleltetése okozza. A jelzeteket egyszerű karakterláncokként kezelve az azonos vagy balról csonkolva azonos jelzetek esetén könnyű megállapítani a kapcsolatot. Ez kibővíthető azzal, hogy az ilyen hasonlóságot a műveleti jelekkel kezdődően³⁹⁰ vagy a csoportképzésen belül is megállapítjuk, de ehhez már szükséges a szeparátorok pontos és zajmentes felismerése, illetve a megoldás továbbra is feltételezi, hogy a facetták sorrendje azonos.

Korábban láthattuk, hogy az ETO-ban a jelzetelemek sorrendje elvben kötetlen, valójában azonban lehet jelentésmódosító hatása. Emiatt a keresőnek – vagy a jelzet rögzítőjének – *olyan sok lehetőséget kellene számba vennie, amely a teljesség elérését gyakorlatilag lehetetlenné teszi.*³⁹¹ Ezért szükséges a jelzeteket részekre bontani, hogy akár automatikusan permutálva, akár a posztkoordinációt lehetővé téve tegyük elérhetővé őket.

Egy korábbi fejezetben láthattuk a szintaktikai relációk jelentőségét az ETO-jelzeten belül. A posztkoordinált keresés, azaz *a jelzetrészek Boolean-operátorokkal való összekapcsolása a keresés során lehetetlenné teszi ennek a pluszinformációnak az érvényesítését.*³⁹²

A **KWIC-**, illetve **permutált index** előnye elsősorban böngészéskor nyilvánulnak meg. Az ilyen index alapján történő kezelés a keresési algoritmustól függ: amennyiben csak a keresőoszlop alapján keresünk, *hatékonysága azonos a KWOC-indexével*, ha viszont a többi jelzetelemet is figyelembe vesszük, a karakterláncokra történő kereséshez hasonló *teljességi problémák* merülnek fel.

Ahogy már korábban említettük, **authority control** alkalmazása esetén elméletileg a jelzetek minden egyes belépési pontját feltárhatjuk, beleértve a komplex jelzetek közötti szemantikai kapcsolatokat, akár a különböző ETO-verziókat is figyelembe véve. Azonban, *ha a hasonlóság megállapítása itt is a karakterláncok azonosságán, illetve a Boolean-operátorok alkalmazásán alapul, akkor a file*

³⁸⁹ Itt csak az ETO analitikus-szintetikus természetéből adódó nehézségekkel foglalkozunk. A szabályok következtetlenségeiből adódó problémákkal, melyeket részletesen tárgyal például Buxton (1990), Mátyás (2011 és 2013), és Piros (2015) a jelzetelemzés kapcsán, egy későbbi fejezetben foglalkozunk.

³⁹⁰ Itt figyelembe kell venni azt is, hogy alosztási jelek általános és korlátozottan közös alosztások között is előfordulhatnak!

³⁹¹ Piros 2014.

³⁹² Egy tágabb kontextusban, a facettás keresés kapcsán Martin Frické a Boolean-alapú megközelítés funkcionálisra cserélésére tesz javaslatot, ami az ETO kapcsán is érdekes vizsgálódási terület lehet (Frické 2017).

*karbantartójának figyelembe kell vennie a karakterláncok szerinti keresés kapcsán említett valamennyi sorrendvariációt, sőt, a jelzetelemek hierarchikus felső kapcsolatait is.*³⁹³ Ez a lépés a jelzetek elemzésével automatizálható, illetve a megfeleltetés akár a keresés során is megtörténhet. Ehhez azonban szükség van a megfelelő jelzetrepresentációra és -elemzésre, mely feltárja a jelzetek belső szintaktikai szerkezetét is.

Eddig csak a jelzetrészek alapján történő – azaz az integrációs szintek szerinti - keresésről volt szó, pedig *érdekes lehet annak a vizsgálata is, hogy az egymást nem tartalmazó kombinált jelzetek mennyiben feleltethetők meg egymásnak.* Az ilyen összehasonlításhoz szintén szükség van a jelzetek elemzésére: a részek és szintaktikai kapcsolataik figyelembe vételével kidolgozhatók olyan mérési eljárások, amelyekkel vizsgálható a fogalmak szemantikai hasonlósága.³⁹⁴ A jelzetek szerkezetének elemzése és a hasonlósági számítások képezhetnék az alapját az automatikus jelzetkiegészítésnek és javaslattételnek, ami nagy mértékben növelhetné a keresési (és böngészési) interfészek hatékonyságát. Az eredménytelen keresések elkerülését célzó „hozzávetőleges egyezőségnek”³⁹⁵ ideális esetben szintén a jelzetek teljes struktúráján kell alapulnia.

Mivel korábban volt szó a felhasználók keresési szokásainak átalakulásáról, érdekességképpen megemlíthető, hogy *a weben publikussá tett könyvtári katalógusokban lehetőség van az ETO szerinti keresésre a Google-ön keresztül is.* A keresőmotor aránylag hatékony az alosztások szerinti keresésben (mivel minden nem numerikus karaktert szeparátorként kezel), viszont nem kezeli a hierarchikus kapcsolatokat. Ráadásul a keresés szűkíthető egy adott könyvtár katalógusára, az oldal megadásával.³⁹⁶

4.5 Az osztályozási munka támogatása

Az osztályozási munka támogatása *elsősorban a gépesített táblázatok és a jelzetek visszakeresésének, illetve az ezeken keresztül kiválasztott vagy létrehozott jelzetek dokumentumrekordokhoz való rendelésének lehetővé tételét jelenti az osztályozó szakember számára.* Ennek megfelelően a technikai kérdések egy része megegyezik a táblázatok gépesítése és a keresés kapcsán tárgyaltakkal.

Az információkeresés kapcsán már volt arról szó, hogy az uralkodó szemlélet szerint egy információ visszakereső rendszernek elsősorban a természetes nyelvű keresést kell a felhasználók számára biztosítania, de az osztályozási rendszer mesterséges nyelvű használatának a lehetővé tétele, illetve a jelzet megjelenítése is hasznos lehet bizonyos keresési stratégiák alkalmazása során. Természetesen az osztályozási munka során a természetes nyelvű kereshetőség mellett *mindenképpen szükséges az osztályozási kódok kezelésének minél hatékonyabb támogatása* is, hiszen ez elkerülhetetlen a jelzetek precíz összeállításához. Ebben a munkafázisban egyébként is feltételezhető, hogy olyan szakemberek kezelik a rendszert, akik járatosak az abban alkalmazott információkereső nyelvekben, és valószínűleg hatékonyabban is tudnak dolgozni, ha közvetlenül is használhatják azokat.

³⁹³Például a 004.384:[621.771.016.3:669.14] (számítógépek vas hideghengerművekben) a 004:621, 004:621.7, 004:621.7.01, 004:669.1, 004.3:669.14 stb. keresőkérdésekre is vissza kell keresni.

³⁹⁴ Lula és Cieraszevska 2017.

³⁹⁵ Slavić-Overfield 2005: 157-158.

³⁹⁶ Balogh 2014.

A táblázatok kereshető és böngészhető módon történő elérhetővé tétele, a jelzetek közötti kapcsolatok (hierarchikus kapcsolatok, lásd még kapcsolatok, a betűrendes mutató stb.) feltárása, esetleg a konkordancia, azaz a jelzetváltozások megjelenítése nagy mértékben megkönnyítheti az indexelő dolgát, különösen, ha ehhez a jelzetépítéshez nyújtott segítség is társul.³⁹⁷ *A táblázatokat az integrált könyvtári rendszer tartalmazhatja beépítve is³⁹⁸, vagy elérhetővé teheti azokat egy online kiadáson keresztül.*

Az ETO táblázatok helyett - vagy mellett – lehetőség van **authority fájl** alkalmazására, amely az összetett jelzetek esetén is biztosíthatja a több nyelvű visszakereshetőséget és a szemantikai kapcsolatok alapján történő böngészést. Ha a kontroll teljes, tehát csak a file-ban már szereplő jelzetek használhatók, akkor alapvető a hatékony visszakereshetőség és böngészhetőség, hiszen csak a rögzített fogalmak használhatók a dokumentumok leírására és fontos, hogy a lehető legmegfelelőbb fogalmat használjuk.

Fontos látni, hogy a túlzott kontroll *korlátozhatja a dokumentumok megfelelő mélységű feltárását*, egyfelől mivel az authority control nem alkalmas a szintagmatikus, eseti szintaktikai kapcsolatok, illetve a dokumentum tárgya mellett annak egyedi jellemzőinek (például nyelv, formai adatok) a leírására, ³⁹⁹ másfelől, mivel egy központi file használata esetén kevésbé érvényesíthetők a gyűjteményi szempontok. A kontroll teljes *hiánya* ugyanakkor egy az ETO-hoz hasonló komplexitású analitikus-szintetikus nyelv esetén *a szinonimák kezelhetetlen túlburjánzásához vezethet*. Ezért, amennyiben a rendszer lehetőséget biztosít új prekoordinált jelzetek létrehozására és használatára, nagyon fontos, hogy ne készüljön újabb jelzet olyan fogalom leírására, amely egyszer már szerepel az adatbázisban (legalábbis a másakra való hivatkozás vagy annak szerepeltetése nélkül). Más szavakkal, a megfelelő visszahívás biztosítása érdekében fontos a korábban már használt jelzetek minél teljesebb visszakereshetősége, illetve javaslatok adása az újonnan összeállított jelzet esetleges leváltására a már meglévők egyikével.

Itt a legnehezebb kérdés *az ekvivalens, de formailag különböző jelzetek felismerése*. Egyszerűbb esetekben a jelzetek csak sorrendben különböznek, azonban ennél bonyolultabb esetek is előfordulhatnak, ha pl. az egyik jelzet a másikban megjelenő kiterjesztés felbontását tartalmazza, vagy például az egyik a másiból előállítható a csoportképzés alkalmazásával.⁴⁰⁰

Az osztályozási munka támogatása során ügyelni kell arra is, hogy a kialakított dokumentumképet egyáltalán ne befolyásolja a rendszer és annak működése, csak maga a dokumentum, illetve a gyűjteményi szempontok.⁴⁰¹ Hasonlóan fontos lenne, hogy a jelzetek belépési pontjainak és a jelzetek kapcsolatainak a meghatározása is a lehető legnagyobb mértékben automatizált legyen, csökkentve a

³⁹⁷ Ez jelentheti azt, hogy a szoftver nemcsak a táblázati számok leírásait, hanem az alkalmazásukra vonatkozó megjegyzéseket, az alosztási jelek és táblázatok részletes leírásait is megjeleníti. További segítség lehet leírások láncindex formájában történő megjelenítése. Az itt felsorolt lehetőségek mind elérhetők a UDC Online portálon („UDC Online”). Bár az oldalon elméletben szintén elérhető egy a számok prekoordinációját segítő alkalmazás, ennek a funkcionalitásnak a teljes körű megvalósításához elengedhetetlen a jelzetek szerkezetének teljes feltárása.

³⁹⁸ Ehhez azonban szükség van a megfelelő licenszre („UDC Licences”).

³⁹⁹ Tartaglia 2004, Vukadin 2015.

⁴⁰⁰ Itt csak a szintaktikai kérdésekkel foglalkozunk, és nem térünk ki a táblázatokban fellelhető szinonimák és kváziszinonimák kérdésére.

⁴⁰¹ A gyűjteményi szempontok mellett természetesen soha nem lehet kizárni az indexelő vagy katalogizáló egyéni perspektívájából adódó különbségeket sem. Ez azonban nem feltétlen hátrány, globális kontextusban a web elméletben lehetőséget adhat arra, hogy kiaknázzuk az azonos források különböző perspektívákból történő leírásának előnyeit is (Tillett 2015).

hibák és a hiányos feltárás által okozott zajt és információvesztést. Az osztályozót abban kell támogatni, hogy a felhasznált rendszertől függetlenül, csak a dokumentum és a gyűjtemény szempontjait szem előtt tartva, a dokumentumokat legjobb tudása szerint indexelje. Más szavakkal, elég legyen csak a jelzetet beírni, a rendszer dolgozza azt fel automatikusan – de lehetőleg az utólagos korrekciót is lehetővé téve – olyan módon, hogy az alapja legyen a fenti feladatok leghatékonyabb megvalósításának.

A fentiek mindegyike igaz akkor is, ha egy könyvtár egy másik – általában nagyobb, illetve központi – könyvtárból átvett jelzeteket használ, kiegészítve a helyben létrehozottakkal.

4.5.1 A jelzetek felbontása

A jelzetek felbontása *a keresés során is releváns lehet*, amennyiben lehetővé tesszük a komplex jelzetekre való keresést, melyeket értelmezni kell, hogy megfelelő hatékonysággal össze tudjuk őket hasonlítani az eltároltakkal. *Nagyobb részt mégis az osztályozási munka, illetve az authority fájl építése során van rá szükség*, legalábbis amennyiben a jelzeteket nem önmagukban, hanem valamilyen kifinomultabb megoldást választva tároljuk.

A jelzetelemek és -részek természetesen *manuálisan is beazonosíthatók*, ahogyan a jelzetek mesterséges és természetes nyelvű belépési pontjai is, ez azonban *túl nagy terhet ró az indexelőre*, miközben növeli a valószínűségét a szinonimák rögzítésének, illetve annak, hogy egyes fontos belépési pontok kimaradnak. Hasonló a probléma a jelzetelemek szóközökkel való elválasztása esetén, amely lehetőséget ad az egyes jelzetelemek kiemelésére, de a típusuk meghatározása nélkül, a hierarchikus (pl. a „12” automatikus csonkolása nem vezet el a „1241”-hez), illetve rejtett jelzetek szerinti keresést (pl. 57/58-re keresve az 57 nem lesz a találatok között) erősen korlátozva.

Egy lehetséges megoldás az, ha *a rendszer a felbontás után lehetőséget ad az eredmény korrigálására*, a fel nem ismert jelzetelemek manuális meghatározására, illetve természetes nyelvű leírások hozzáadására is.⁴⁰²

A jelzetek felbontására *számos könyvtári rendszer* nyújt többé-kevésbé teljes megoldást. Ugyanakkor a hatvanas évek óta *tudományos igényű kutatások is folynak*, amelyek a nyelv jellegzetességeit figyelembe véve nagyobb mélységgel vizsgálják az ETO számok értelmezésének összetett problémáját. Ezekkel a kutatásokkal, illetve a jelzetelemzés kérdéseivel részletesebben a hatodik fejezetben foglalkozunk.

4.6 A jelzetek eltárolása

A jelzetek visszakeresésének és böngészésének, megjelenítésének, az adatcserének, illetve a jelzetépítés automatikus támogatásának egyaránt a jelzetek megfelelő formátumban történő elmentése az alapja. A jelzetek számítógépes katalóguson belüli tárolása meghatározza az osztályozást, és ez az alapja a keresésnek is. Ezek a funkcionalitások és a keresés eredményének a megfelelő megjelenítése *szükségessé teszi, hogy elválasszuk a jelzetek megjelenített és eltárolt formáját*: előbbinek a séma szabályai szerint kell megjelenítenie, utóbbinak pedig a jelzet elemeit és struktúráját kell tartalmazni egy további automatikus feldolgozásra alkalmas formátumban.⁴⁰³ Ideális esetben azonban az utóbbi alkalmas arra is, hogy *a jelzet eredeti, illetve attól különböző, de azzal ekvivalens formátumokban reprodukálható legyen.*

⁴⁰² Ilyen megoldást találhatunk a magyar fejlesztésű TextLib rendszerben, amely képes az egyszerűbb jelzetek elemzésére, de lehetőséget ad az eredmény korrigálására is.

⁴⁰³ Slavić-Overfield 2005: 168.

A tárolás két szinten valósul meg, ezért annak kérdéseit két nagy csoportra oszthatjuk. A tárolás módjának **technikai kérdései** közé tartozik a jelzeteket tároló file-ok formátuma, a használt adatbázis típusa, a használt adatmodell, és az ehhez kapcsolódó, szűkebben vett technikai kérdések. Ezek a kérdések egy adott rendszer tervezői, programozói és fenntartói hatáskörébe tartoznak, feltéve, hogy az adatbázisból a jelzetek, azaz a dokumentumkép a megfelelő formátumokban, a szükséges részletességgel reprezentálhatók a korábban említett funkciók támogatása érdekében.

A tárolás második, magasabb szintje a dokumentumkép, illetve az **ETO-jelzet reprezentációja**. Ez az alapja a korábban már említett valamennyi funkciónak: a keresésnek, böngészésnek, kapcsolatvizualizációnak, a járulékos információk megjelenítésének, a jelzetépítés támogatásának. A megfelelően megalkotott jelzetreprezentációnak *alkalmasnak kell lennie az előbb felsorolt funkcionalitások számára szükséges minden információ eltárolására és visszanyerésére*, legyen szó a jelzetek elemeiről, azok leírásairól, a köztük meghatározott szintaktikai relációkról vagy a jelzetek közötti fennálló szemantikai kapcsolatokról.

Az ETO-jelzetek reprezentációja szerint beszélhetünk olyan rendszerekről, amelyek a jelzeteket mindenféle járulékos információ nélkül, „ahogy vannak” kezelik, és olyanokról, amelyek törekednek azok belső szerkezetének bizonyos szintű feltárására, a szintaktikai kapcsolatokkal kapcsolatos információ megtartása érdekében.

4.6.1 A jelzetek, „ahogy vannak”

Bár az analitikus-szintetikus osztályozási rendszerek teljes körű gépesítése szükségessé tenné a jelzetek részletes elemzését⁴⁰⁴, az *online könyvtári katalógusok jelentős része az ETO-jelzeteket egyetlen mezőben, egységes karaktorsorozatként tárolja* anélkül, hogy azok belső szerkezetét és kapcsolatait – akár a bibliográfiai rekordban, akár egy kapcsolódó állományban – feltárná.

Ehhez az is hozzájárul, hogy az adattárolás alapjául szolgáló bibliográfiai *adatcsere formátumok maguk sem adnak lehetőséget az ETO számok szerkezetének részletes feltárására*. Ilyen például a magyar HUNMARC szabvány, mely csak az ETO számok egyszerű, minden járulékos információ nélküli eltárolására ad lehetőséget az ismételhető 080-as mezőben⁴⁰⁵, vagy az IFLA által karbantartott UNIMARC formátum, amelyben a 675-ös mező használható a jelzetek eltárolására (melynek megfelelő almezőiben a felhasznált ETO kiadás adatai is rögzíthetők).⁴⁰⁶ Az ISSN MARC szintén a 080-as mezőt használja az ETO-jelzet és a használt kiadás eltárolására.⁴⁰⁷

A bibliográfiai rekordokban ilyen formában *eltárolt jelzetekhez kapcsolódhatnak olyan indexek, amelyek elősegíthetik az egyes jelzetelemek szerinti visszakeresést*, ezek a megoldások azonban csak *korlátozottan alkalmasak az analitikus-szintetikus osztályozások lehetőségeinek kibontására*.

⁴⁰⁴ Slavić és Cordeiro 2005.

⁴⁰⁵ Barki és Prejczer 2003: 4.4.2. Egyetemes Tizedes Osztályozás. Érdekes, hogy a HUNMARC alapjául szolgáló MARC 21 bibliográfiai formátum az ETO-jelzetek hatékonyabb leírása miatt tartalmazza a \$x almező kódot az általánosan közös alosztások bevezetésére (LOC 1999b: 080 – Universal Decimal Classification Number). Az alosztások típusának megkülönböztetésére, a korlátozottan közös alosztások és műveletek leírására azonban ez a formátum sem ad lehetőséget.

⁴⁰⁶ IFLA 2008: 70.

⁴⁰⁷ „ISSN Manual test”: 13.1.

4.6.1.1 A jelzetek karaktersorozatként való kezelése

Az online könyvtári katalógusok, OPAC-ok jelentős részében az ETO-jelzetek tárolására egy, esetleg ismételtelhető adatmező áll rendelkezésre, amelybe a jelzetek egyszerű karaktersorozatként vannak eltárolva, a jelzetelemek automatikus kiemelése, megjelölése nélkül. Az ilyen rendszerek gyakran az ETO számok természetes nyelvű leírásainak rögzítésére és visszakeresésére sem adnak lehetőséget.

Amennyiben a jelzetek egyszerű karaktersorozatként vannak eltárolva, *a jelzetelemek szerinti keresés explicit vagy automatikus csonkolással, esetleg a Boolean-műveletekkel lehetséges.*⁴⁰⁸ Ez azonban az ETO analitikus-szintetikus jellege, jelzetépítési szabályai és a hivatkozási sorrend kötetlensége miatt egyáltalán *nem hatékony sem a precízió, sem a visszahívás szempontjából.*

A módszer nehézségeit könnyen szemléltethetjük egy összetett fogalom ETO számán keresztül, ha tekintjük például a Francia Igazságügyi Minisztérium 354.51(44) jelzetét.⁴⁰⁹ A földrajzi alosztás ebben az esetben nemcsak a fő táblázati szám után, de előtte, sőt, annak tetszőleges pozíciójára közbeékelve is előfordulhat a jelentés megváltozása nélkül. Ha a szám valamennyi formai variációját szeretnénk visszakeresni, az alábbi keresőkérdést kellene összeállítanunk:

(44)354.51 VAGY 354.51(44) VAGY 3(44)54.51 VAGY 35(44)4.51 VAGY 354(44).51 VAGY *354(44)51* VAGY 354.5(44)31.

Az inkluzív keresés érdekében ezt csonkolással egészíthetjük ki, az alábbi módon:

(44*)354.51* VAGY 354.51*(44*) VAGY 3(44*)54.51* VAGY 35(44*)4.51* VAGY 354(44*).51* VAGY 354(44*)51* VAGY 354.5(44*)31*.

A 354.51*(44*) jelzet már zajt okoz, mivel a fő táblázati szám után álló csillag helyére kerülhet alosztási jel is, és előfordulhat, hogy a földrajzi alosztás nem erre a fő táblázati számra fog vonatkozni.

Ha figyelembe szeretnénk venni azokat az eseteket, amikor a szám valamely művelet második paramétereként, esetleg egy csoportképzésen belül szerepel, akkor balról is kell csonkolnunk:

(44)354.51* VAGY *354.51*(44*) VAGY *3(44*)54.51* VAGY *35(44*)4.51* VAGY *354(44*).51* VAGY *354(44*)51* VAGY *354.5(44*)31*.

Ez tovább növeli a zajt, mivel a 354.51-re végződő egyéb fő táblázati számok is visszakeresésre kerülnek.

Amennyiben feltételezzük, hogy egyéb alosztások is kapcsolódhatnak a keresett fő táblázati számhoz, újabb csonkolásokat kell hozzáadnunk, amelyek tovább növelik a zajt, ugyanakkor nagyon figyelmesen kell eljárunk, ha egyetlen esetet sem szeretnénk kifejezteni. Figyelembe kell venni például a földrajzi alosztáson belüli esetleges műveleteket és egyéb alosztásokat is.

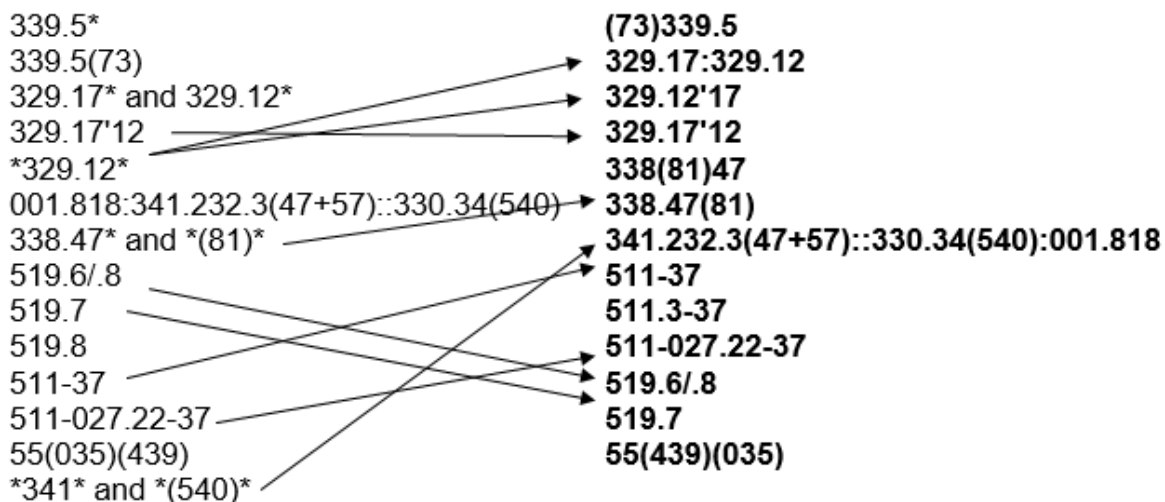
Természetesen lehetőség lenne a *354.51* ÉS *(44*) keresőkérdés használatára is, ez azonban nagyon nagy zajjal is járhat, hiszen ez a feltétel minden 354.51-re végződő fő táblázati szám és 44-re végződő földrajzi (etnikai, esetleg formai) alosztás esetén is igaz lesz.

A csonkolással való keresés következtelenségeire és hiányosságaira láthatunk néhány szemléletesebb példát a 5. ábrán.⁴¹⁰

⁴⁰⁸ A csonkolással való keresésnek a BME Aleph rendszeréből származó példákon keresztüli ismertetéseit találhatók Marton József és Prokné Palik Mária (2009), illetve Prokné Palik Mária (2013) cikkeiben.

⁴⁰⁹ Piros 2014.

⁴¹⁰ Piros 2015.



5. ábra ETO-jelzetek elérése csonkolással

Amennyiben a jelzeteket tároló mező ismétlődhet, a jelzetelemek szerinti kereshetőség növelhető a különböző sorrendvariációk, esetleg a jelzetelemek önálló jelzetként való rögzítésével. Előbbi az osztályozótól kívánja meg az összes lehetőség számbavételét, utóbbi pedig nemcsak a precíziót csökkenti (mivel az elemek közötti kapcsolat elvész), de ellentmond annak a korábban már említett elvnek is, hogy a leírásoknak minél pontosabban tükrözniük kell a leírt dokumentum fogalmi viszonyait.

Természetes nyelvű keresés az – esetleg komplex – tárgyi kifejezések leírásának szavai alapján lehetséges, ha a rendszer lehetővé teszi ezek rögzítését. A jelzetrészek nem kerülnek felismerésre, ezért azok leírásai alapján közvetlenül nem lehet keresni.

Mivel a szemantikai kapcsolatok nincsenek explicit módon feltárva, és a szintaktikai kapcsolatok kifejezése hiányában automatikusan sem ismerhetők fel, az így tárolt jelzetek között böngészni csak a jelzetek, mint karaktersorozatok formai hasonlósága alapján lehet. Ez éppen a fent látott sokalakúság miatt nagyon pontatlan jelzetsorrendet eredményezhet. Többé-kevésbé helyes sorrend még rendezési segédstringek alkalmazása esetén is csak akkor érhető el, ha ragaszkodunk a jelzetelemek hivatkozási sorrendjére tett ajánlások szigorú betartásához, illetve a műveletek valamennyi lehetséges permutációjának rögzítéséhez.

A fent látott nehézségek eredménye, hogy azokban a rendszerekben, amelyekben az ETO számok további feldolgozás nélkül, egyszerű karaktersorozatként vannak eltárolva, a jelzetek többnyire csak a dokumentumok szabványos leírásához járulnak hozzá⁴¹¹, de a keresett információ fellelését érdemben nem segítik. Buxton egyenesen úgy fogalmaz, hogy a jelzetek sima sztringként való kezelése az osztályozó idejének és az osztályozási rendszerben felhalmozott tudásnak az elherdálása.⁴¹²

⁴¹¹ Az angol nyelvű szakirodalom erre a „mark and park” kifejezést használja.

⁴¹² Buxton 1990.

4.6.1.2 A jelzetelemek KWOC-indexe

A **KWOC** mozaikszó a „**Keywords out of context**” („szövegkörnyezeten kívüli vezérszó”) rövidítése. Az eredetileg címindexelésre kitalált módszer lényege, hogy a cím visszakereshetővé tenni kívánt szavait, azaz *a vezérszavakat a címből kiemelten de azok mellett, külön keresőoszlopban szerepeltetjük*. ETO-jelzetek esetén *a cím helyén az ETO szám áll, a vezérszavak pedig a jelzetnek azon alkotói, amelyek önálló indexfogalmakként alkalmazhatók, azaz amelyekre keresés történhet.*⁴¹³ Ez elméletben lehet komplex jelzet is, de a gyakorlatban általában az egyszerű táblázati számok, esetleg intervallumok kiemelése szokott megtörténni. A korlátozottan közös alosztások szerinti visszakeresésre az ilyen indexet alkalmazó rendszerekben általában nincs lehetőség, bár elméletben ez is megoldható.⁴¹⁴

Ahogy azt már korábban említettük, az ETO-jelzetek KWOC-jellegű indexének első alkalmazása Schneider és Koch nevéhez fűződik⁴¹⁵. Olcsósága és könnyű alkalmazhatósága miatt a módszer széles körben elterjedt, és *általában ezt tekintik a jelzetelemek szerinti keresés kielégítő módjának*, amennyiben az authority control alkalmazására nincs lehetőség.

A KWOC-index módszere általában véve és az ETO-ra alkalmazva is elsősorban az információkeresés **standard Boolean-modelljét** támogatja.⁴¹⁶ Ebben az esetben az ETO-jelzetre tekintünk úgy, mint összetevői egyszerű halmazára, melyet azok Boolean-műveletekkel való összekapcsolásával kereshetünk vissza. A kiemelt jelzetelemek közötti *szintaktikai kapcsolatok és az általuk hordozott többletjelentés tehát nem érvényesíthetők* a keresés során. Ennek a visszahívásra és precízióra tett hatásairól korábban, a szintaktikai kapcsolatok, illetve a kontextus ETO-jelzeten belüli jelentősége kapcsán már részletesebben volt szó.

A teljesség érdekében külön figyelmet kell fordítani arra is, hogy a kiterjesztés által alkotott intervallumok belsejében lévő számok, illetve a jelölési konvenciókból adódóan rövidített számok (például az intervallumok végpontjai⁴¹⁷ vagy az aposztrófos jelzetszintézis elemei⁴¹⁸) is megfelelően szerepeljenek az indexben.

A precízió növelhető, amennyiben komplex jelzeteket is szerepeltetünk a keresőoszlopban, ebben az esetben azonban ugyanabba a problémába ütközünk, amiről a jelzetek karaktersorozatként való kezelése kapcsán már beszéltünk, nevezetesen, hogy az ETO rugalmas jelzetszerkesztési szabályai miatt a meglehetősen nagy számú lehetőség mindegyikét figyelembe kell venni a vezérszavak összeállítása során⁴¹⁹, ez pedig még az egyszerű jelzetrészek felismerésénél is nehezebben automatizálható.

⁴¹³ Orosz 1967.

⁴¹⁴ A korlátozottan közös alosztások KWOC-szerű kezelésére Mándy Gábor javasolt egy megoldást (2013). Eszerint a fókusz, tehát a korlátozottan közös alosztási számot az alkalmazhatósági helyével együtt kell szerepeltetni az indexben (tehát azzal a számmal, amely alatt a táblázatokban szerepel).

⁴¹⁵ Orosz 1967.

⁴¹⁶ Nyamisa, Mwangi és Cheruiyot 2017.

⁴¹⁷ Ilyen megoldást alkalmaznak a MARC osztályozási formátumok (LOC 2000 és IFLA 2000).

⁴¹⁸ Ezt a megoldást említi Buxton (1990) és Riesthuis (1999) is.

⁴¹⁹ Ez visszatérő probléma minden olyan módszer esetén, amely nem tárja fel a jelzetek belső szerkezetét részletesen, a sorrendtől függetlenül, reprodukálható módon.

A jelzetelemek automatikus felismerése ugyanis az ETO következetlen szabályai miatt egyébként sem könnyű feladat, ahogyan arra már számos korábbi kutatás rámutatott.⁴²⁰ Ahogy már korábban említettük, az ETO revíziós munkálatainak egyik fő célja az utóbbi időben éppen ezeknek a következetlenségeknek a csökkentése, a jelzetkezelés javítása érdekében.⁴²¹

Természetesen az index manuálisan is összeállítható. Ez azonban részint megsokszorozza a jelzetelés munkaigényét, így növeli annak idő, erőforrás és így az anyagi költségeit is. Másrészt növeli az indexelőre nehezedő felelősséget is, mivel nem csak a dokumentum megfelelő leírására kell figyelnie, hanem a vezérszavak megfelelő kiválasztására is. A lehetséges keresőszavak esetleges kifejejtése ugyanis csökkenti a visszahívást, szélsőséges esetekben visszakereshetetlenné is teheti a jelzetet.

Ezért *a leghatékonyabb az automatikus és manuális jelzetelemzés kombinálása*. Egyes rendszerek a jelzetelemek felismerését úgy oldják meg, hogy minden szóközt szeparátornak tekintenek, így valamennyi olyan jelzetrész visszakereshető lesz, amelyet az osztályozó szóközökkel választ el a többitől. Ez a módszer csak a felbontási algoritmust egyszerűsíti le, azt nagyjából az indexelőre hárítva, a több felvetett problémára nem kínál megoldást. További hátránya, hogy csak manuális osztályozás során alkalmazható, retrospektív konverzió vagy valamely külső adatbázis automatikus feldolgozása során nem.

4.6.1.3 A jelzetelemek KWIC-indexe

A **KWIC-index** („**Keywords-in-context**”, „Vezérszavak a kontextusukban”) eredetileg szintén címindexelési módszer. Lényege, hogy *a cím visszakeresendő vezérszavait a címen belül hozzuk kereső pozícióba*, tehát a keresőoszlop címen belüli pozícióját változtatjuk annak megfelelően, mely szavakat kívánjuk visszakereshetővé tenni. ETO-jelzetek esetén ez *a jelzetrészek kereső pozícióba hozását* jelenti, ahogy azt a 6. ábrán szereplő példában láthatjuk.

6. ábra Jelzetelemek KWIC-indexe

Erre a módszerre a legismertebb magyarországi példa az Országos Széchényi Könyvtár NEKTÁR

73:72.04.035.5	(439-2 Bp.)
73:72.04.035.5	(439 -2 Bp.)
73:72.04.035.5	(439-2 Bp.)
73:72.04	.035.5 (439-2 Bp.)
73:72	.04.035.5 (439-2 Bp.)
73:	72.04.035.5 (439-2 Bp.)
	73:72.04.035.5 (439-2 Bp.)

adatbázisa.⁴²²

A KWIC-indexelés módszere egyesíti a jelzetek karaktersorozatként való ábrázolásának és a KWOC-indexnek mind az előnyeit, mind a hátrányait. Akárcsak a KWOC-index esetében, itt sincs szükség baloldali csonkolásra a keresés során, ez *csökkenti a zaj mértékét, ugyanakkor megőrződik a kontextus által hordozott információ* is. Ez különösen hasznos a böngészés során, amikor a felsorolt rekordok

⁴²⁰ A teljesség igénye nélkül például Freeman és Atherton (1967), Buxton (1990), Riesthuis (1997), Mándy (2011), Mándy (2013) és Piros (2015) is részletesen vizsgálta ezeket a kérdéseket. A felmerülő problémákról részletesebben egy későbbi fejezetben lesz szó.

⁴²¹ Slavić és Davies 2017. A jelzetelemek felismerésének problémáiról egy későbbi fejezetben lesz részletesen szó.

⁴²² Zöldi 1993. A példa is ebből a cikkből származik.

relevanciáját kell eldönteni. Ezzel szemben, akárcsak a karaktersorozatként való kezelés esetén, az utóbbi információ *igazán csak akkor érvényesíthető a keresés és böngészés során, ha egy jelzet összes lehetséges sorrendvariációja rögzítve van* az adatbázisban, illetve, a fenti módszerekhez hasonlóan *megoldatlan a rejtett jelzetek, a nem teljes jelzetelemek, illetve a korlátozottan közös alosztások keresésének problémája*. Utóbbiak a KWOC-indexekben megoldhatók a megfelelő, jelzetben nem szereplő vezérszavak hozzáadásával, itt azonban erre nincs lehetőség.

A módszer gépesítésének legnehezebb kérdése a különböző jelzetvariánsok előállítása. Ez automatikusan nagyon nehezen megoldható, manuálisan viszont az indexelőre ró nagy felelősséget, ahogyan azt a korábbiakban már láthattuk. Egyébként a jelzetelemek automatikus kereső pozícióba hozása során ugyanolyan módszerekre támaszkodhatunk, és ugyanazokkal a nehézségekkel is szembesülünk, mint a KWOC-index esetében.

4.6.2 A kombinált jelzetek strukturált adatként való eltárolása

A prekoordinált jelzetek megfelelő kezelése szükségessé teszi a jelzetek részeinek és azok szintaktikai kapcsolatainak, a szemantikai és asszociatív kapcsolatoknak, illetve a kódok mellett azok természetes nyelvű leírásainak és indexkifejezéseinek az eltárolását és megfelelő kezelését is. Ehhez mindenképpen szükséges egy olyan állomány kialakítása, amely a jelzetek megjelenített formája mellett ezeket az információkat is tartalmazza. A tárgyi visszakeresés automatizálása kapcsán ezt a megoldást nevezzük **authority control**nak.⁴²³

A fentiekből is látszik, hogy az authority control céljaira minden olyan jelzettárolási módszer, illetve formátum megfelel, amely alkalmas a jelzetek mesterséges és természetes nyelvű kifejezésének, a keresési kulcsszavainak és a szemantikai kapcsolatoknak az eltárolására és feltárására. *Analitikus-szintetikus esetben azonban speciális koncepcionális modellre, illetve formátumra van szükség.*⁴²⁴

A történeti áttekintés kapcsán volt arról szó, hogy az IFLA bibliográfiai metaadatmodell-családjának tárgyi authority rekordok számára szánt ajánlása, az **FRSAD (Functional Requirements for Subject Authority Data)** elsősorban tezauruszok és subject heading listák leírására alkalmas.⁴²⁵ Ennek oka, hogy a modell a tartalom („aboutness”) leírására készült, a „mű-téma-nomen” hármas alapján.⁴²⁶ Emiatt osztályozási sémákra csak a rendszerek alapos analízisét követően alkalmazható.⁴²⁷ Ilyen kutatással kapcsolatos eredmények eddig a DDC és a CLC (Kínai Könyvtári Osztályozás) kapcsán lettek publikálva.⁴²⁸

A **szemantikus webre** kidolgozott formátumokat megvizsgálva hasonló helyzettel találkozhatunk. Bár a kontrollált szótárak leírására szánt **SKOS**⁴²⁹ *egyelőre nem alkalmas facettás osztályozások*

⁴²³ A kifejezés jelentése tehát kissé eltér attól, ahogy azt a bibliográfiai adatok kapcsán alkalmazzuk, mivel itt az entitások kapcsolatainak feltárása legalább olyan fontos cél, mint az egységes névalak meghatározása.

⁴²⁴ Slavić és Cordeiro 2005.

⁴²⁵ IFLA 2010. A modell korábbi neve FRSAR (Functional Requirements for Subject Authority Records).

⁴²⁶ A megközelítés elemei tehát hasonlóak az Ogden-Richards háromszögéhez, de ez a modell a kapcsolatokat lineárisan ábrázolja, így kevesebbet árul el a grammatikai kifejezési módok és az objektum kapcsolatának jellegéről.

⁴²⁷ Mitchell, Zeng és Žumer 2011.

⁴²⁸ Mitchell, Zeng és Žumer 2011 és Mitchell, Zeng és Žumer 2012.

⁴²⁹ „SKOS Simple Knowledge Organization System - Home Page”.

leírására, az OWL 2 ontológialeíró nyelv⁴³⁰ viszont alkalmazható erre a célra, de ehhez szükséges az osztályozási rendszer részletekbe menő vizsgálata.⁴³¹ Ennek a nyelvnek a felhasználására a nagy osztályozási rendszerek közül szintén a DDC kapcsán történtek kísérletek.⁴³²

Az authority control egyik *legfontosabb célja a kontroll biztosítása* a hálózaton belül, hogy a tárgykörök és szemantikai kapcsolataik leírásait az arra legfelkészültebb szakemberek készítsék el, majd a leírásokat a különböző intézmények egységes formában használják. Ezért az alkalmazása során *különösen fontos az adatcsere biztosítása*. Mindezek alapján a **MARC formátumok** természetes választásnak tűnnek a jelzetek authority fájlokban való eltárolására és továbbítására vagy akár az osztályozások karbantartására.⁴³³ Ehhez azonban szükség lenne olyan MARC szabványokra, amelyek alkalmazkodnak ezeknek a rendszereknek a sajátosságaihoz.

4.6.2.1 Osztályozási MARC formátumok

Ákárcsak a bibliográfiai, az általános és tárgyi authority control céljaira készült MARC formátumok is csak az osztályozási jelzetek egyszerű eltárolását teszik lehetővé, esetleg az osztályozás felhasznált kiadására, vagy a jelzet szintaktikai szerkezetére vonatkozó minimális információ társaságában. A részletesebb leírásokra külső adatbázisokban, azok rekordjaira kapcsoló mezőkben való hivatkozással van lehetőség.⁴³⁴

A kilencvenes években felmerült az igény kifejezetten az osztályozási jelzetek authority controlja számára létrehozott MARC formátumok kidolgozására is, melyek lehetőséget adnak az osztályozási jelzetek megjelenítési módjának és belső struktúrájának az elkülönítésére. Az első ilyen az USMARC (később MARC 21) formátumcsalád részeként készült el, majd ezt követte az UNIMARC osztályozási formátuma.

4.6.2.1.1 MARC 21 Osztályozási Formátum

Az első, határozottan osztályozási jelzetek leírására tervezett MARC formátum az USMARC volt.⁴³⁵ Később a formátum az USMARC és CAN/MARC összevonásával keletkezett MARC 21 formátumcsaládnak is része maradt.⁴³⁶

Mivel az LCC és DDC alapján készült, a formátum az egyszerű és paradigmatis, osztatlan számként kezelt intervallumok mellett a kiterjesztés eseti alkalmazásával (fő táblázati számokból) alkotott intervallumok kezelésére ad módot, illetve lehetővé teszi a hierarchikus linkelést is.⁴³⁷

A formátum főbb adatmezőiben alkalmazkodik a MARC 21 szabvány más formátumaihoz. Mivel authority rekordok számára készült, tartalmaz mezőket a „lásd” (253) és „lásd még” referenciák

⁴³⁰ „OWL 2 Web Ontology Language Document Overview”.

⁴³¹ Gnoli et. al. 2011.

⁴³² Zeng, Panzer és Salaba 2010 és Gnoli, De Santis és Pusterla 2015.

⁴³³ Cochrane és Markey; hiv. Slavić-Overfield 2005: 189, San Segundo 2009.

⁴³⁴ IFLA 2001, LOC 1999a: 080 – Universal Decimal Classification Number, „RUSMARC”, Rozman 2015, Ungváry 2010.

⁴³⁵ Guenther 1992; hiv. Slavić-Overfield 2005: 190.

⁴³⁶ „MARC 21: Harmonized USMARC and CAN/MARC”.

⁴³⁷ Slavić-Overfield 2005: 191.

sámára (353), és mivel osztályozási rendszerekből származó kódok leírására, ezért a séma pontos megnevezésre és a felhasznált kiadás számára is (084).⁴³⁸

A két legfontosabb mező a 153-as („osztályozási szám”) és a 765 („a szintetizált szám komponensei”).

A 153-as mező almezői lehetőséget adnak az osztályozási szám vagy intervallum (utóbbi esetben a végpontokat külön almezőkben lehet elhelyezni), illetve hierarchikus kapcsolatainak kifejezésére. Ez a mező tartalmazza a jelzet és hierarchikus kapcsolatai természetes nyelvű meghatározásait is.⁴³⁹

A 765-ös mező nem ad lehetőséget a szintetizált számok komponenseinek a leírására. A mező ismételhető: amennyiben egy jelzet előállításához több szabály is fel lett használva, mindegyiket külön mezőben kell elhelyezni. A mező elsődleges célja, hogy arról adjon információt, hogy „a szintetizált szám vagy a szintetizált szám egy része hogyan lett felépítve.”⁴⁴⁰

Azzal, hogy az intervallumok kivételével minden jelzetkapcsolási mód külön mezőbe kerül, *a formátum a DDC és LCC korlátozott szintetizációs lehetőségeihez alkalmazkodik*. Mivel ezekben a rendszerekben a facettákat jelölő számok jelentése helyenként változik, a 765-ös mező *almezői önmagukban nem alkalmasak információkeresésre*, inkább csak a könyvtárosok számára szolgáltatnak információt arról, hogy egy adott szám hogyan épül fel.⁴⁴¹

Tehát amennyiben a MARC 21 formátumot szeretnénk felhasználni kombinált ETO számok leírására, *le kellene mondanunk az általánosan közös alosztások leírásáról*, ez a formátum ugyanis csak a főtáblázati számok kiterjesztésének (a 153-as mezőben) és a korlátozottan közös alosztások (765-ös mezők) kifejezésére ad lehetőséget.

4.6.2.1.2 UNIMARC Osztályozási Formátum

Az USMARC osztályozási formátumának megjelenése után az IFLA úgy döntött, létrehoz egy hasonló formátumot, amely az Európában ekkor domináns UNIMARC formátumcsalád részeként fogja lehetővé tenni a kontinensen ekkor első számú osztályozási rendszernek számító Egyetemes Tizedes Osztályozás kezelését. Ennek érdekében 1994-ben létrehoztak egy a DDC, az ETO és a UNIMARC szakértőiből álló egyesített munkacsoportot, mely 1996-ban adta közre a munkáját lezáró jelentést.

Ebből kiderül, hogy a dokumentum kimondottan a meglévő USMARC formátum kiterjesztését célozta meg, és *nem vette kellő mértékben figyelembe az ETO speciális jellemzőit*. Ennek eredménye, hogy az új formátumot az osztályozási rendszer és az az alapján készült authority fájlok kezelésére egyaránt alkalmasnak szánták, figyelmen kívül hagyva, hogy egy analitikus-szintetikus rendszer esetén a kapcsolat jóval gyengébb a kettő között, mint az LLC-hez és DDC-hez hasonló enumerikus rendszerek esetében.

A formátumban a legtöbb változást a MARC 21 és a UNIMARC formátumok különbségei indukálták. Az osztályozási számok leírására szolgáló legfontosabb mezők belső szerkezetük meghagyása mellett, egyszerűen át lettek helyezve az UNIMARC családhhoz jobban illeszkedő helyekre (a 153 a 250-be és a 765 a 665-be). Az általánosan közös alosztások a 665-ös mező \$b almezőjében fejezhetők ki. Az ETO

⁴³⁸ LOC 2000: 084 - Classification Scheme and Edition.

⁴³⁹ LOC 2000: 153 - Classification Number.

⁴⁴⁰ LOC 2000: 765 - Synthesized Number Components.

⁴⁴¹ Slavić-Overfield 2005: 193.

alosztási jeleit, illetve az alosztások típusát csak a \$4 vezérlő almezővel lehet kifejezni, tehát a formátum nem tesz érdemi különbséget a különböző műveletek és alosztási táblázatok között.

A javaslat elkészülte után egy külön munkacsoport jött létre, amely több javaslatot is tett a formátum módosítására annak érdekében, hogy azt alkalmasabbá tegyék az ETO alapú információkeresésre.⁴⁴² A **javaslat** legfontosabb eleme a 250-es mező kettéválasztása volt: eszerint a 250-es mező megmaradt volna az enumeratív osztályozások kódjai számára, az analitikus-szintetikus rendszerek alapján létrehozott **szintetizált jelzetek** pedig a 251-es mezőbe kerültek volna. Utóbbi almezői tartalmazták volna a főtáblázati számok mellett az általánosan és korlátozottan közös alosztási táblázati számokat. A 675-ös mező a párhuzamos továbbosztásokat, illetve azonos jelentéssel, de más kóddal szereplő korlátozottan közös alosztásokat tartalmazta volna. Az alosztási jelek – a BC2 fázisrelációival kiegészülve – és a korlátozottan közös alosztások típusát továbbra is a \$4 tartalmazta volna. A legfontosabb változtatást az jelentette volna, hogy *a különböző típusú általánosan közös alosztások számára külön almezőket tartottak volna fent a 251-es mezőn belül: \$e az idő, \$f a népi, \$g a hely szerinti alosztások kifejezésére és így tovább. Ezzel a táblázat típusa kereséskor is érvényesíthetővé vált volna.*⁴⁴³

A javasolt módosítások, bár *nem fedik le az ETO teljes szintaktikai palettáját*, az összes komplex és több szintű jelzettelési lehetőséget, olyan formátumot eredményezhettek volna, amely kielégítő megoldást nyújthatna a prekoordinált ETO-jelzetek adatcseréjére és eltárolására. Sajnos azonban *nem kerültek elfogadásra*, és a formátum befejezetlen formában, mindössze egy rövidített leírással lett közreadva.

4.7 Irodalomjegyzék

„ASCII History.” ASCII World. <http://ascii-world.wikidot.com/history>. (megtekintve: 2017-11-01)

Babiczyk Béla. 1985. „Szemléletváltozás az ETO jelzetszerkesztésében.” *Könyvtári Figyelő* 31 (1): 17–27.

Babiczyk Béla, és Barátné Hajdu Ágnes. 1998. *Bevezetés az információkereső nyelvek elméletébe és gyakorlatába*. Budapest: Universitas.

Balíková, Marie. 2005. „Multilingual subject access to catalogues of national libraries (MSAC) : Czech Republic’s collaboration with Slovakia, Slovenia, Croatia, Macedonia, Lithuania and Latvia.” In *Libraries: a voyage of discovery. 71th IFLA general conference and council. Oslo, Norvégia 2005.08.14-18*. <http://www.ifla.org/IV/ifla71/papers/044e-Balikova.pdf>. (megtekintve: 2017-11-01)

Balíková, Marie. 2009. „The role of UDC classification in the Czech Subject Authority File.” *Extensions and Corrections to the UDC* 31: 253-274.

Balogh Béla. 2014. „ETO és Google.” Katalist (levelezőlista). <https://listserv.niif.hu/pipermail/katalist/2014-January/030127.html>. (megtekintve: 2017-11-01)

Bánki Zsolt István. 2005. „Osztályozási jelzetek és tárgyszavak konvertálhatóságának lehetőségei.” *Tudományos és műszaki tájékoztatás* 52 (11-12): 528-534. http://tmt-archive.omikk.bme.hu/show_news.html?id=4244&issue_id=467. (megtekintve: 2017-11-01)

⁴⁴² A javaslat az ETO mellett a BC2 facettás és relációs lehetőségeire is tekintettel volt.

⁴⁴³ A fejezet eddigi része teljes egészében Slavić-Overfield (2005: 194-203) alapján íródott.

- Barátné Hajdu Ágnes. 2006. „Usability and Responsibility.” *Extensions and Corrections to the UDC* 28: 46-55.
- Barátné Hajdu Ágnes. 2007a. „A magyar ETO-kiadás munkálatai. Hazai és nemzetközi tapasztalatok.” *Könyv, könyvtár, könyvtáros* 16 (2): 20-39. <http://epa.oszk.hu/01300/01367/00122/pdf/02muhelykerdesek.pdf>. (megtekintve: 2017-11-01)
- Barátné Hajdu Ágnes. 2007b. „A percepció és megjelenítés jelentősége az információkereső nyelvekben.” *Tudományos és műszaki tájékoztatás* 54 (10): 461-478. http://tmt-archive.omikk.bme.hu/show_news.html?id=4785&issue_id=487. (megtekintve: 2017-11-01)
- Barátné Hajdu Ágnes. 2009. „Integration of a thesaurus and Universal Decimal Classification (UDC) to improve subject access: the Hungarian experience.” *Extensions and Corrections to the UDC* 31: 197-207. <http://hdl.handle.net/10150/199891>. (megtekintve: 2017-11-01)
- Barki katalin, és Prejczér Paula. 2003. *Számítógépes katalogizálás hunmarc formátumban*. Szombathely: Berzsenyi Dániel Főiskola Központi Könyvtár, Berzsenyi Dániel Főiskola Könyvtár- és Információtudományi Tanszék. <http://mek.oszk.hu/05200/05287/html/index.htm>. (megtekintve: 2017-01-28)
- Bates, Marcia J. 1989. „The Design of Browsing and Berrypicking Techniques for the Online Search Interface.” *Online Review* 13 (5): 407-424. <https://pages.gseis.ucla.edu/faculty/bates/berrypicking.html>. (megtekintve: 2017-01-28)
- Berners-Lee, Tim, James Hendler, és Ora Lassila. 2001. „The Semantic Web” *Scientific American* 284 (5): 29-37.
- Bourne, Charles P., és Trudi Bellardo Hahn. 2003. *A History of Online Information Services, 1963-1976*. Cambridge, MA, and London: MIT Press
- Broughton, Vanda. 2010. „Concepts and Terms in the Faceted Classification: the Case of UDC.” *Knowledge Organization* 37 (4): 270-279.
- Buxton, Andrew. 1990. „Computer Searching of UDC Numbers.” *Journal of Documentation* 46 (3): 193-217.
- „Cancellations to the UDC.” UDCC (UDC Consortium). <http://udcc.org/index.php/site/page?view=cancellations>. (megtekintve: 2017-11-01)
- „CDS/ISIS Database Software.” UNESCO. <http://www.unesco.org/new/en/communication-and-information/information-society/open-source-and-low-cost-technologies/information-processing-tools/cdsisis-database-software/>. (megtekintve: 2017-11-01)
- Chatterjee, Amitabha. 2016. *Elements of Information Organization and Dissemination*. Amsterdam: Elsevier.
- Cochrane, Pauline Atherton, és Karen Markey. 1985. „Preparing for the Use of Classification in Online Cataloging Systems and in Online Catalogs.” *Information Technology and Libraries* 4 (2): 91-111
- „Creative Commons 3.0. Attribution-ShareAlike 3.0 International (CC BY-SA 3.0).” Creative Commons (CC) <https://creativecommons.org/licenses/by-sa/3.0/>. (megtekintve: 2017-11-01)
- Dahlberg, Ingetraut. 1972. „Az ETO újjáalakításának lehetőségeiről” [Fordította Horváth Magda]. *Könyvtári Figyelő* 18 (4): 412-430.

- Dancs Szabolcs. 2016. „Ismét az ETO hasznáról - az új elektronikus kiadás fényében.” *Könyv, könyvtár, könyvtáros* 25 (4): 3-8.
http://epa.oszk.hu/01300/01367/00277/pdf/EPA01367_3K_2016_04_003-008.pdf.
 (megtekintve: 2017-11-01)
- Dudás Anikó. 2013. „Nemcsak weben lenni, hanem webből lenni : A Funkcionális követelmények (FR) metaadatmodell-család névterei és a szemantikus web.” *Könyvtári figyelő* 59 (1): 45-64.
http://epa.oszk.hu/00100/00143/00086/pdf/EPA00143_konyvtari_figyelo_2013_1_045-064.pdf.
 (megtekintve: 2017-11-01)
- Ellis, David. 1989. „A Behavioural Approach to Information Retrieval System Design.” *Journal of Documentation* 45 (3): 171-212.
- „Az ETO kiadásának gépesítése.” 1965. *Könyvtári figyelő* 11 (1): 62.
http://epa.oszk.hu/00100/00143/00176/pdf/EPA00143_konyvtari_figyelo_1965_01_62.pdf.
 (megtekintve: 2017-11-01)
- Fejős László. 2007. „Az ETO adatbázis magyar kiadása és konkordanciája.” *Könyvtári levelező/lap* 19 (7): 3-6. <http://epa.oszk.hu/00300/00365/00068/pdf/00068.pdf>. (megtekintve: 2017-11-01)
- Fejős László. 2008a. „Adalékok az ETO új kiadásához és konkordanciájához.” *Könyv, könyvtár, könyvtáros* 17 (3): 19-28. <http://epa.oszk.hu/01300/01367/00187/pdf/02muhelykerdesek.pdf>.
 (megtekintve: 2017-11-01)
- Fejős László. 2008b. Adalékok az ETO új kiadásához és konkordanciájához (folytatás). *Könyv, könyvtár, könyvtáros* 17 (4): 26-32.
<http://epa.oszk.hu/01300/01367/00188/pdf/04muhelykerdesek.pdf>. (megtekintve: 2017-11-01)
- Fejős László. 2008c. Nyomdakész az ETO konkordancia. *Könyvtári Levelező/lap* 20 (12): 8, 20.
<http://www.epa.oszk.hu/00300/00365/00085/pdf/00085.pdf>. (megtekintve: 2017-11-01)
- Fejős László. 2009. ETO gondozás, konkordancia. *Könyvtári levelező/lap* 21 (2): 3-4.
http://www.epa.oszk.hu/00300/00365/00092/pdf/KLL_EPA00365_2009-02_003-004.pdf.
 (megtekintve: 2017-11-01)
- FID (International Federation de Documentation). 1983. *Az Egyetemes Tizedes Osztályozás (ETO) alapelvei, revíziójának és kiadásának szabályai (FID 603)*. Budapest: OMIKK.
- Frâncu, Victoria. 2003. „Multilingual access to information using an intermediate language.” PhD dissz., Antwerpen: Universiteit Antwerpen. <http://hdl.handle.net/10150/106319>. (megtekintve: 2017-11-01)
- Frâncu, Victoria, és Liviu-Iulian Dediu. 2015. „TinREAD – an integrative solution for subject authority control.” In *Classification & Authority Control: Expanding Resource Discovery : Proceedings of the International UDC Seminar 2015, 29-30 October 2015, Lisbon, Portugal*, szerk. Aida Slavić, és Maria Inês Cordeiro, 123-134. Würzburg: Ergon Verlag.
- Frâncu, Victoria, és Cosmin-Nicolae Sabo. 2010. „Implementation of a UDC-Based Multilingual Thesaurus in a Library Catalogue: The Case of BiblioPhil.” *Knowledge Organization* 37 (3): 209-215.
- Freeman, Robert R., és Pauline Atherton. 1967. *File Organization and Search Strategy Using the Universal Decimal Classification in Mechanized Retrieval Systems*. New York: American Institute of Physics (AIP).

- Frické, Martin H. 2017. „Faceted classification, analysis and search: some questions on their interrelations.” In *Faceted classification today: theory, technology and end users: proceedings of the International UDC Seminar 2017, London (UK), 14-15 September*, edited by Aida Slavić, and Claudio Gnoli, 25-41. Würzburg: Ergon Verlag.
- Gedam, Pranali B., és Ashwini Paradkar. 2013. „A study of web-based library classification schemes.” *International Journal of Library and Information Science* 5 (10): 386-393.
- Gendler-Szabó Zoltán. 2004. „Compositionality.” In *Stanford Encyclopedia of Philosophy*. Stanford, CA: Stanford University. <https://plato.stanford.edu/entries/compositionality>. (megtekintve: 2017-11-01)
- Gnoli, Claudio. 2009. „The UDC Philosophy Revision: First Report.” *Extensions and Corrections to the UDC* 31: 25-31. <http://hdl.handle.net/10150/200633>. (megtekintve: 2017-11-01)
- Gnoli, Claudio, Philippe Cousson, Tom Pullman, Gabriele Merli, és Rick Szostak. 2011. „Representing the structural elements of a freely faceted classification.” In *Classification and ontology: formal approaches and access to knowledge: Proceedings of the International UDC Seminar, 19-20 September 2011, The Hague, The Netherlands*, szerk. Aida Slavić, és Edgardo Civalero, 193-206. Würzburg: Ergon Verlag.
- Gnoli, Claudio, Rodrigo De Santis és Laura Pusterla. 2015. „Commerce, see also Rhetoric: cross-discipline relationships as authority data for enhanced retrieval.” In *Classification & Authority Control: Expanding Resource Discovery : Proceedings of the International UDC Seminar 2015, 29-30 October 2015, Lisbon, Portugal*, szerk. Aida Slavić, és Maria Inês Cordeiro, 151-162. Würzburg: Ergon Verlag.
- Guenther, Rebecca S. 1992. „The Development and Implementation of the USMARC Format for Classification Data.” *Information Technology and Libraries* 11 (2): 120-31.
- Harper, Corey, és Barbara B. Tillett. 2006. „Library of Congress controlled vocabularies and their application to the Semantic Web.” *Cataloging & Classification Quarterly* 43 (3/4): 47-68.
- Horváth Ádám. 2015. „Gondolatok a szemantikus webről és egyben az ALIADA szoftverről.” *Könyvtári Figyelő* 62 (3): 319-326. http://epa.oszk.hu/00100/00143/00304/pdf/EPA00143_konyvtari_figyelo_2015_3_319-326.pdf. (megtekintve: 2017-11-01)
- Hug, Hannes, és Rudolf Nöthiger. 1992. „ETHICS (ETH-Library information control system).” In *IATUL proceedings. 14th Biennial Conference of the International Association of Technological Libraries. Cambridge, MA, USA, July 8-12, 1991*, 124-127. London: British Library. <http://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1157&context=iatul>. (megtekintve: 2017-11-01)
- Hutchins, Williams John. 2001. „Indexelő és osztályozó nyelvek. Nyelvészeti tanulmány struktúrákról és funkciókról.” In *Osztályozás és információkeresés. Kommentált szöveggyűjtemény. Első kötet: Az osztályozás elmélete*, szerk. Ungváry Rudolf, és Orbán Éva, 388-469. Budapest: Országos Széchényi Könyvtár.
- International Federation of Library Associations (IFLA). 2000. „Concise UNIMARC Classification Format.” http://softsbgp.free.fr/bibliotheque/Unimarc_Format_US.pdf. (megtekintve: 2017-11-01)
- International Federation of Library Associations (IFLA). 2001. „UNIMARC Manual. Authorities Format.” <https://www.ifla.org/files/assets/uca/unimarc-authorities-format.pdf>. (megtekintve: 2017-11-01)

- International Federation of Library Associations (IFLA). 2008. „UNIMARC concise bibliographic Format.” <http://archive.ifla.org/VI/8/unimarc-concise-bibliographic-format-2008.pdf>. (megtekintve: 2017-11-01)
- International Federation of Library Associations (IFLA). 2010. „Functional Requirements for Subject Authority Data (FRSAD): Final Report.” <https://www.ifla.org/node/5849>. (megtekintve: 2017-11-01)
- „International UDC Seminar 2015.” UDCC (UDC Consortium). <http://seminar.udcc.org/2015/>. (megtekintve: 2017-11-01)
- „International UDC Seminars.” UDCC (UDC Consortium). <http://seminar.udcc.org/>. (megtekintve: 2017-11-01)
- „ISO 2709.” International Federation of Library Associations (IFLA). <http://www.ifla.org/best-practice-for-national-bibliographic-agencies-in-a-digital-age/node/8811>. (megtekintve: 2017-11-01)
- „ISSN Manual test.” International Standard Serial Number (ISSN) International Center. Utolsó módosítás 2010. <http://www.issn.org/issn-manual-test/>. (megtekintve: 2017-11-01)
- Kowalczyk, Agnieszka Maria, Łukasz Skonieczny, és Małgorzata Wornbard. 2015. „Visualisation of Warsaw University of Technology Main Library resources based on UDC.” Classification & authority control: expanding resource discovery: proceedings of the International UDC Seminar, 29-30 October 2015, Lisbon, Portugal. <http://seminar.udcc.org/2015/abstracts.php#kowalczyk>. (megtekintve: 2017-11-01)
- Kumar, T.V.R., és Parameswaran, M. 1999. „Chain indexing and UDC.” *Annals of Library Science and Documentation* 46 (2): 53–58.
- Library of Congress (LOC). 1999a. „MARC 21 Format for Authority Data : 1999 Edition (Update No. 24).” utolsó módosítás 2017. május. <http://www.loc.gov/marc/authority>. (megtekintve: 2017-11-01)
- Library of Congress (LOC). 1999b. „MARC 21 Format for Bibliographic Data : 1999 Edition (Update No. 24).” utolsó módosítás 2017. május. <http://www.loc.gov/marc/bibliographic>. (megtekintve: 2017-11-01)
- Library of Congress (LOC). 2000. „MARC 21 Format for Classification Data : 2000 Edition (Update No. 24).” utolsó módosítás 2017. május. <http://www.loc.gov/marc/classification/eccdhome.html>. (megtekintve: 2017-11-01)
- Lula, Paweł, és Urszula Cieraszewska. 2017. „Similarity measurement between UDC classmarks and its application.” In *Faceted classification today: theory, technology and end users: proceedings of the International UDC Seminar 2017, London (UK), 14-15 September*, edited by Aida Slavić, and Claudio Gnoli, 219-239. Würzburg: Ergon Verlag.
- Makke, Jane. 2016. „Who Needs UDC? Notes from Estonian Libraries.” *IFLA Metadata Newsletter* 2 (2): 34-37. http://www.ifla.org/files/assets/classification-and-indexing/ifla_metadata_newsletter_dec2016.pdf. (megtekintve: 2017-11-01) (megtekintve: 2017-11-01)
- „Major Changes to the UDC 1993-2013.” UDCC (UDC Consortium). http://www.udcc.org/index.php/site/page?view=major_revisions. (megtekintve: 2017-11-01)

- Mándy Gábor. 2011. „Régi ETO – új ETO” *Könyvtári Figyelő* 57 (2): 317-333.
http://epa.oszk.hu/00100/00143/00079/pdf/EPA00143_Konyvtari_Figyelo_2011_2_317-333.pdf.
 (megtekintve: 2017-11-01)
- Mándy Gábor. 2013. „A posztkoordináció esélyei az ETO-ban.” *Könyvtári figyelő* 59: 65-84.
http://epa.oszk.hu/00100/00143/00086/pdf/EPA00143_konyvtari_figyelo_2013_1_065-083.pdf.
 (megtekintve: 2017-11-01)
- „MARC 21: Harmonized USMARC and CAN/MARC.” Network Development and MARC Standards Office. Library of Congress. 1998. október 22. <https://www.loc.gov/marc/annmarc21.html>.
 (megtekintve: 2017-11-01)
- Marton József, és Prokné Palik Mária. 2009. „Adatbázis transzformációalapú kereshetőségjavítása: ETO-kereshetőség a BME OMIKK Aleph rendszerében.” *Tudományos és műszaki tájékoztatás* 56 (2): 65-77. http://tmt-archive.omikk.bme.hu/show_news.html?id=5057&issue_id=501.
 (megtekintve: 2017-11-01)
- McIlwaine, I. C. 1998. „The Universal Decimal Classification: some factors concerning its origins, development and influence.” In *Historical studies in information science*, szerk. Trudi Bellardo Hahn, és Michael Buckland, 94-106. Medford, NJ: Information Today.
- McIlwaine, I. C. 2006. „The new ecumenism: Exploration of a DDC/UDC view of religion.” *Extensions and Corrections to the UDC* 28: 9-16.
- McIlwaine, I. C., és Nancy Williamson. 2008. „Medicine and the UDC: the process of restructuring Class 61.” *Extensions and Corrections to the UDC* 30: 9-16.
- Mitchell, Joan S., Marcia Lei Zeng, és Maja Žumer. 2011. „Extending models for controlled vocabularies to classification systems: modelling DDC with FRSAD.” In *Classification and ontology: formal approaches and access to knowledge: Proceedings of the International UDC Seminar, 19-20 September 2011, The Hague, The Netherlands*, szerk. Aida Slavić, és Edgardo Civalero, 241-250. Würzburg: Ergon Verlag.
- Mitchell, Joan S., Marcia Lei Zeng, és Maja Žumer. 2012. „Modeling classification systems in multicultural and multilingual Contexts.” *Paper presented at the IFLA Satellite Post-Conference: Beyond Libraries – Subject Metadata in the Digital Environment and Semantic Web. 17-18 August 2012, Tallinn, Estonia*. http://www.nlib.ee/html/yritus/ifla_jarel/papers/3-3_Mitchell.docx.
 (megtekintve: 2017-11-01)
- Nyamisa, Mang'are Fridah, Waweru Mwangi, és Wilson Cheruiyot. 2017. „A Survey of Information Retrieval Techniques.” *Advances in Networks* 5 (2): 40-46.
- O'Neill, Edward T., Rick Bennett, és Kerre Kammerer. 2014. „Using Authorities to Improve Subject Searches.” *Cataloguing & Classification Quarterly* 52 (1): 6-19.
- Orosz Gábor. 1967. „Az ETO alkalmazása a gépesített dokumentációs és könyvtári munkákban.” *Tudományos és műszaki tájékoztatás* 14 (8): 541-560. http://tmt-archive.omikk.bme.hu/show_news.html?id=2474&issue_id=166. (megtekintve: 2017-11-01)
- Orosz Gábor. 1977. „Az Egyetemes Tizedes osztályozás tárgymutatójának összeállítása számítógéppel.” *Tudományos és műszaki tájékoztatás* 24 (7-8): 271-293.
- Országos Műszaki Információs Központ és Könyvtár (OMIKK), és Országos Széchényi Könyvtár Könyvtártudományi és Módszertani Központ (OSZK-KMK). 1990. *Egyetemes tizedes osztályozás. Rövidített kiadás. 1. kötet Táblázatok. FID Publication 691*. Budapest: OMIKK, OSZK-KMK.

- „OSZK Tezaurusz / Köztaurusz.” Országos Széchényi Könyvtár (OSZK). Utolsó módosítás 2013. szeptember. <http://mek.oszk.hu/adatbazis/thes.htm>. (megtekintve: 2017-11-01)
- „OWL 2 Web Ontology Language Document Overview (Second Edition).” World Wide Web Consortium (W3C). <https://www.w3.org/TR/owl2-overview/>. (megtekintve: 2017-11-01)
- Pethes Iván, Tolnai György, Turcsányi Tamás, és Varga Jánosné. 1976. „ETO-jelzetek számítógépes rendezése.” *Könyvtári figyelő* 22 (1-2): 66-87.
http://epa.oszk.hu/00100/00143/00228/pdf/EPA00143_konyvtari_figyelo_1976_01-02_066-087.pdf. (megtekintve: 2017-11-01)
- Pika, Jiri. 2007. „Universal Decimal Classification at the ETH-Bibliothek Zürich - a Swiss perspective.” *Extensions and Corrections to the UDC* 29: 229-251.
<http://arizona.openrepository.com/arizona/handle/10150/105111> (megtekintve: 2017-11-01)
- Pika, Jiri, és Milena Pika-Biolzi. 2015. „Multilingual subject access and classification-based browsing through authority control: the experience of the ETH-Bibliothek, Zürich.” In *Classification & Authority Control: Expanding Resource Discovery : Proceedings of the International UDC Seminar 2015, 29-30 October 2015, Lisbon, Portugal*, szerk. Aida Slavić, és Maria Inês Cordeiro, 177-194. Würzburg: Ergon Verlag.
- Piros Attila. 2014. „Az ETO alapú számítógépes információkeresés egy új megközelítése.” *Könyvtári Figyelő* 60 (2): 156-166.
http://epa.oszk.hu/00100/00143/00091/pdf/EPA00143_konyvtari_figyelo_2014_2_156-166.pdf. (megtekintve: 2017-11-01)
- Piros Attila. 2015. „Az ETO-jelzetek automatikus elemzésének kérdései.” *Tudományos és műszaki tájékoztatás* 62 (4): 127-142. http://tmt-archive.omikk.bme.hu/show_news.html?id=6025&issue_id=569. (megtekintve: 2017-11-01)
- Pollitt, Steven A. 1997. „The key role of classification and indexing in view-based searching.” In *Conference Programme and Proceedings, 63rd IFLA General Conference, Copenhagen, Denmark, 31 August - 5 Sept. 1997*. <https://archive.ifla.org/IV/ifla63/63polst.pdf>. (megtekintve: 2017-11-01)
- Prokné Palik Mária. 2013. „Google vs. ETO: Információkeresés a weben és a könyvtári katalógusban.” *Tudományos és műszaki tájékoztatás* 60 (6): 243-273.
http://epa.oszk.hu/03000/03071/00066/pdf/EPA03071_tmt_2013_06_243-273.pdf. (megtekintve: 2017-11-01)
- Rigby, Malcolm. 1974. *Computers and the UDC : a decade of progress 1963-1973. (FID 523)*. The Hague: International Federation of Documentation (FID)
- Riesthuis, Gerhard. 1997. „Decomposition of complex UDC notations.” *Extensions and Corrections to the UDC* 19: 13-19.
- Riesthuis, Gerhard. 1999. „Searching with words: re-use of subject indexing.” *Extensions & Corrections to the UDC* 21: 24-32.
- Riesthuis, Gerhard. 2003. „A Revised Format for the Master Reference File Database.” *Extensions and Corrections to the UDC* 25: 11-18.
- Rozman Darija. 2015. „Experience with UDC updates: the Slovenian perspective .” In *Classification & authority control: expanding resource discovery: proceedings of the International UDC Seminar*,

- 29-30 October 2015, Lisbon, Portugal, szerk. Aida Slavić, és Maria Inês Cordeiro, 227-232. Würzburg: Ergon Verlag.
- „RUSMARC.” Российская Национальная Библиотека.
<http://www.rusmarc.info/2017/auth/content.htm>. (megtekintve: 2017-11-01)
- San Segundo, Rosa. 2009. „Using MARC classification format for UDC and mappings to other KO systems for an enriched authority file.” *Extensions and Corrections to the UDC* 31: 285.291.
<http://arizona.openrepository.com/arizona/handle/10150/200631>. (megtekintve: 2017-11-01)
- Schallier, Wouter. 2004. „What a subject search interface can do.” *Extensions and Corrections to the UDC* 26: 18-28. <http://hdl.handle.net/10150/105315>. (megtekintve: 2017-11-01)
- „SKOS Simple Knowledge Organization System - Home Page.” Semantic Web. World Wide Web Consortium (W3C). <https://www.w3.org/2004/02/skos>. (megtekintve: 2017-11-01)
- Slavić-Overfield, Aida. 2005. „Classification Management and use in a Networked Environment: The Case of the Universal Decimal Classification.” PhD diss., London: University College London.
http://discovery.ucl.ac.uk/1334914/1/AidaSlavicOverfield_thesis_UCL2005.pdf. (megtekintve: 2017-11-01)
- Slavić, Aida, és Maria Inês Cordeiro. 2005. „Sharing and re-use of classification systems: the need for a common data model.” *Signum* 37 (8): 19-24. <http://hdl.handle.net/10150/105132>. (megtekintve: 2017-11-01)
- Slavić, Aida. 2006. „Interface to classification some objectives and options.” *Extensions and Corrections to the UDC* 28: 24-45.
- Slavić, Aida. 2008. „On presenting UDC Summary on the web.” Universal Decimal Classification (blog), UDC Consortium, 2008. szeptember 30.
<http://universaldecimalclassification.blogspot.hu/2008/09/on-presenting-of-udc-summary-on-web.html>. (megtekintve: 2017-11-01)
- Slavić, Aida, Maria Inês Cordeiro, és Gerhard Riesthuis. 2008. „Maintenance of the Universal Decimal Classification: Overview of the Past and Preparations for the Future.” *International Cataloging and Bibliographic Control* 37 (2): 23-29. <http://hdl.handle.net/10150/105220>. (megtekintve: 2017-11-01)
- Slavić, Aida. 2009a. „Multilingual UDC Summary – available.” Universal Decimal Classification (blog), UDC Consortium, 2009. december 19.
<http://universaldecimalclassification.blogspot.hu/2009/12/multilingual-udc-summary-available.html>. (megtekintve: 2017-11-01)
- Slavić, Aida. 2009b. „The round table 'UDC editorial perspectives': a report.” *Extensions and corrections to the UDC* 31: 19-23.
<http://arizona.openrepository.com/arizona/handle/10150/118425>. (megtekintve: 2017-11-01)
- Slavić, Aida, és Antoine Isaac. 2009. „Identifying management issues in networked KOS: examples from classification schemes.” NKOS 2009 workshop. 13th ECDL conference, Corfu, Greece. Thursday, October 1st 2009. https://at-web1.comp.glam.ac.uk/pages/research/hypermedia/nkos/nkos2009/presentations/slavic_isaac_NKOS2009_06.pdf. (megtekintve: 2017-11-01)
- Slavić, Aida, Chris Overfield, Gerhard Riesthuis, és Jiri Pika. 2009. „Multilingual UDC summary project: 2009 update.” *Extensions and corrections to the UDC* 31: 13-18.

- Slavić, Aida. 2012. „UDC as Linked Data.” Universal Decimal Classification (blog), UDC Consortium, 2012. augusztus 26. <http://universaldecimalclassification.blogspot.hu/2012/08/udc-as-linked-data.html>. (megtekintve: 2017-11-01)
- Slavić, Aida. 2013. „UDC Online – a new service for users of the scheme.” Universal Decimal Classification (blog), UDC Consortium, 2013. szeptember 22. <http://universaldecimalclassification.blogspot.hu/2013/09/udc-online-new-tool-for-users-of-scheme.html>. (megtekintve: 2017-11-01)
- Slavić, Aida, és Sylvie Davies. 2017. „Facet Analysis in UDC: questions of structure, functionality and data formality.” In *Faceted classification today: theory, technology and end users: proceedings of the International UDC Seminar 2017, London (UK), 14-15 September*, szerk. Aida Slavić, és Claudio Gnoli, 181-200. Würzburg: Ergon Verlag.
- Strachan, David P., és F. M. H. Oomes. 2001. „The UDC MRF Database Development and Design – a historical review.” UDCC (UDC Consortium). <http://www.udcc.org/mrf2.htm>. (megtekintve: 2017-11-01)
- Tartaglia, S. 2004. „Authority Control and Subject Indexing Languages.” *Cataloging & Classification Quarterly* 39 (1/2): 365-377.
- Thék György. 2014. „tárgyszó vagy ETO (is).” Katalist (levelezőlista). <https://listserv.niif.hu/pipermail/katalist/2014-April/030774.html>. (megtekintve: 2017-11-01)
- Tillett, Barbara B. 2001. „Authority Control on the Web.” In *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium: Confronting the Challenges of Networked Resources and the Web*. Washington, D.C.: Library of Congress. <http://www.loc.gov/catdir/bibcontrol/tillett.html>. (megtekintve: 2017-11-01)
- Tillett, Barbara B. 2015. „Complementarity of perspectives for resource descriptions.” In *Classification & Authority Control: Expanding Resource Discovery : Proceedings of the International UDC Seminar 2015, 29-30 October 2015, Lisbon, Portugal*, szerk. Aida Slavić, és Maria Inês Cordeiro, 19-30. Würzburg: Ergon Verlag.
- „UDC Licences.” UDCC (UDC Consortium). <http://udcc.org/index.php/site/page?view=licences>. (megtekintve: 2017-01-01)
- „UDC Master Reference File (MRF).” UDCC (UDC Consortium). <http://www.udcc.org/index.php/site/page?view=mrf>. (megtekintve: 2017-11-01)
- „UDC MRF Use Licences.” UDCC (UDC Consortium). http://udcc.org/index.php/site/page?view=mrf_licence. (megtekintve: 2017-11-01)
- „UDC Online.” UDCC (UDC Consortium). <http://www.udc-hub.com/>. (megtekintve: 2017-11-01)
- „UDC Online English – Prices.” UDCC (UDC Consortium). <http://www.udc-hub.com/en/prices.php>. (megtekintve: 2017-11-01)
- „UDC Publishing Licences.” UDCC (UDC Consortium). http://udcc.org/index.php/site/page?view=publishing_licence. (megtekintve: 2017-11-01)
- „UDC Summary. Multilingual Universal Decimal Classification Summary. UDCC Publication No. 088.” UDCC (UDC Consortium). <http://www.udcc.org/udcsummary/php/index.php>. (megtekintve: 2017-11-01)

- „UDC Summary Exports.” UDCC (UDC Consortium). <http://www.udcsummary.info/exports.htm>. (megtekintve: 2017-01-21)
- „UDC Summary Linked Data.” UDCC (UDC Consortium). <http://udcdata.info/>. (megtekintve: 2017-11-01)
- „UDC Summary Translations.” UDCC (UDC Consortium). <http://www.udcsummary.info/translation.htm>. (megtekintve: 2017-11-01)
- Ullah, Asim, Shah Khusro, és Irfan Ullah. 2017. „Bibliographic Classification in the Digital Age: Current Trends & Future Directions.” *Information Technology and Libraries* 36 (3): 48-77. doi:10.6017/ital.v36i3.8930
- UNESCO. 2015. *Interoperability and Retrieval*. Paris: UNESCO. <http://unesdoc.unesco.org/images/0023/002321/232199E.pdf>. (megtekintve: 2017-11-01)
- Ungváry Rudolf. 2013. „A Relex névtérkezelő.” (előadás) Networkshop 2013 konferencia, Sopron, 2013. március 26-28. <https://niif.videotorium.hu/hu/recordings/6195/relex-nevterkezo>. (megtekintve: 2017-11-01)
- Ungváry Rudolf, és Orbán Éva. 2001a. *Osztályozás és információkeresés. Kommentált szöveggyűjtemény. Első kötet: Az osztályozás elmélete*. Budapest: Országos Széchényi Könyvtár. <http://mek.oszk.hu/01600/01683/pdf/01683-1.pdf>. (megtekintve: 2017-11-01)
- Ungváry Rudolf, és Orbán Éva. 2001b. *Osztályozás és információkeresés. Kommentált szöveggyűjtemény. Második kötet: Az osztályozás elmélete*. Budapest: Országos Széchényi Könyvtár. <http://mek.oszk.hu/01600/01683/pdf/01683-2.pdf>. (megtekintve: 2017-11-01)
- Ungváry Rudolf, és Vajda Erik. 2002. *Könyvtári információkeresés*. Budapest: Typotex. <http://mek.oszk.hu/05000/05030/05030.pdf>. (megtekintve: 2017-11-01)
- Ungváry Rudolf. 2010. „MARC21/HUNMARC: a besorolási adatok metaadat-formátuma – Főbb jellemzők, fejlődés és problémák.” *Könyvtári Figyelő*. 56 (1). <http://ki.oszk.hu/kf/2010/10/marc21hunmarc-a-besorolasi-adatok-metaadat-formatuma-fobb-jellemzok-fejlodes-es-problemak/>. (megtekintve: 2017-11-01)
- „Universella Decimalklassifikationen. Svensk Elektronisk Utgåva.” Miguel Benito. <http://www.taranco.eu/udk/>. (megtekintve: 2017-11-01)
- Van der Linden, Erik-Jan, Jerke J. Van Wijk. 2007. „Visual Universal Decimal Classification?” *Extensions and Corrections to the UDC 29*: 297-300. <http://arizona.openrepository.com/arizona/handle/10150/105224>. (megtekintve: 2017-11-01)
- Vukadin, Ana. 2015. „Development of a classification-oriented authority control: the experience of the National and University Library in Zagreb.” In *Classification & Authority Control: Expanding Resource Discovery : Proceedings of the International UDC Seminar 2015, 29-30 October 2015, Lisbon, Portugal*, szerk. Aida Slavić, és Maria Inês Cordeiro, 111-120. Würzburg: Ergon Verlag.
- „What is UDC Sumary?” UDCC (UDC Consortium). <http://www.udcsummary.info/about.htm>. (megtekintve: 2017-11-01)
- Zeng, Marcia Lei, Michael Panzer, és Athena Salaba. 2010. „Expressing Classification Schemes with OWL 2 Web Ontology Language.” In *Paradigms and conceptual systems in KO: Proceedings of the ISKO 2010 International Conference. Sapienza University of Rome, February 23-26, 2010, Rome, Italy*, szerk. Claudio Gnoli, és Fulvio Mazocchi. 356-362. Ergon Verlag.

Zöldi Péter. 1993. „Az ETO az Országos Széchényi Könyvtár számítógépes rendszerében.” *Könyvtári figyelő* 39 (3): 384. http://epa.oszk.hu/00100/00143/00007/zoldi_h.html. (megtekintve: 2017-11-01)

5 Az ETO számok szerkezetének leírása

„Az analitikus-szintetikus osztályozások információkeresésre való felhasználása során az egyik legfontosabb követelmény a szintetizált jelzetek kezelése és minden szabványt, formátumot vagy jelölőrendszert aszerint kell megítélni, hogy mennyire képes a szintetizált számok struktúrájának kódolására.”⁴⁴⁴ (Aida Slavić)

5.1 Bevezetés

A korábbi fejezetekben áttekintettük, milyen feladatokkal és nehézségekkel kell szembenéznünk az ETO mint analitikus-szintetikus osztályozási rendszer gépesítése során. Láthattuk, hogy *a problémát elsősorban a szintaktikai kapcsolatok megfelelő kezelése*, a rugalmas jelzetszerkesztési szabályokra való reflektálás, a szintagmatikus fogalmak kifejezésének és visszakereshetőségének biztosítása okozza.

Azt is szóba került, hogy a problémák *már a jelzetek elmentése során megjelennek*, mivel a jelenleg elterjedt jelzettárolási módszerek nem képesek biztosítani azt a háttérrel, amely szükséges a további funkcionalitások megfelelő támogatásához, elsősorban azért, mert vagy nem is biztosítják a jelzetek elemei szerinti elérést, vagy mert nem veszik figyelembe azok belső, szintaktikai szerkezetét.

Paradigmatikus jelzetek, illetve megfelelően karbantartott, a gyűjteményhez folyamatosan igazított authority fájl esetén az *authority control elvben megfelelő* megoldást nyújthatna. Ennek azonban akadálya egy olyan megfelelő és elfogadott *jelzettárolási szabvány hiánya*, amely alkalmas lenne az ETO bonyolult jelzetszerkesztési szabályainak a kezelésére.

A MARC osztályozási formátumok jelen formájukban nem alkalmasak erre, ahogyan a SKOS szabvány sem. Az FRSAD ajánlás, illetve az OWL 2 ontológia leíró nyelv alkalmazhatóságát pedig egyelőre csak más, elsősorban enumerikus osztályozási rendszerekkel (DDC, CLC) kapcsolatban vizsgálták.

Mivel a **jelen kutatás** általános, távlati célja az ETO gépesítésére szolgáló, a korábbiaknál hatékonyabb módszerek kidolgozásának a támogatása, az elsődlegesen megoldandó feladat egy olyan **formátum** kidolgozása, amely **alkalmas a jelzetek belső szintaktikai szerkezetének megfelelő mélységű feltárására**. Másképpen fogalmazva a formátum célja az analitikus-szintetikus osztályozási rendszerek automatizálásában alapelvnek számító „függetlenség és integritás”⁴⁴⁵ lehető legteljesebb mértékű biztosítása azzal, hogy a jelzetnek nem csak valamennyi elemét, de teljes szintaktikai struktúráját is feltárjuk és eltároljuk.

A formátum elsősorban technikai célokat szolgál: célja az ETO-jelzetek átalakítása egy olyan alakra, ami automatikusan, **programok által is könnyen analízálható**, információvesztés nélkül **átalakítható további formátumokra**, illetve kereső és klaszterező **algoritmusok alapjául is szolgálhat**. A rendeltetése tehát nem az authority fájlok leváltása, hanem azok kiegészítése ⁴⁴⁶, illetve karbantartásuknak támogatása, például a jelzetek és formátumok közötti konverzió vagy - a megfelelő elméleti kutatások után - a szintaktikai kapcsolatokból automatikus kikövetkeztethető szemantikai kapcsolatok felismerésének lehetőségével.⁴⁴⁷

⁴⁴⁴ Slavić-Overfield 2005: 183.

⁴⁴⁵ Slavić és Cordeiro 2005.

⁴⁴⁶ vö. Vukadin 2015.

⁴⁴⁷ Piros 2015a.

Az információkereső rendszerekben való felhasználás mellett a jelzetek ilyen mélységű feltárása lehetőséget ad a jelzetelemzésre épülő más jellegű, például kvantitatív kutatások folytatására is. A formátum alapjául szolgáló elvek, illetve a grammatikai szabályok kezelésére nyújtott megoldások egyéb, szabványos formátumok létrehozásának alapjául is szolgálhatnak.

Ez a fejezet a fent említett **formátum mögötti alapelveket** és magának **a formátumnak a részletes leírását** tartalmazza. Mind az ismertetett alapelvek, mind a formátum **teljes egészében önálló kutatómunka eredményei**.⁴⁴⁸ A felhasznált ETO kiadások, illetve azok a munkák, amelyekre a kutatás során támaszkodtam, minden esetben fel vannak tüntetve, ahogyan a formátum elérhetősége és felhasználásának feltételei is.

5.2 A formátummal szemben támasztott követelmények

Ahogyan a bevezetőben említettük, az ismertetett *kutatás célja egy olyan jelzettárolási megoldás kidolgozása, amely alkalmas az ETO-jelzetek belső szintaktikai szerkezetének megfelelő mélységű feltárására*, az ETO számítógépesítésével kapcsolatos, az előző fejezetben ismertetett feladatok hatékony megoldásának támogatása érdekében.

Ehhez elsősorban egy olyan formátumot kell definiálnunk az ETO számok leírására, amely

- *ember és számítógép által egyaránt értelmezhető,*
- *alkalmas a jelzetek szintaktikai szerkezetének elemzésére és összehasonlítására, tehát kereső algoritmusok és klaszteranalízis alapjául is felhasználható.*

Továbbá

- *reprodukálható belőle a kiinduló jelzet (karaktersorozatként),*
- *alkalmas a jelzetelemek indexének előállítására, beleértve a jelzet egyszerű és több elemű részjelzeteit is,*
- *konvertálható tetszőleges adatcsere vagy a jelzetet és szerkezetét leíró egyéb formátumra,*
- *alkalmas a jelzetből kifejezhető hierarchikus kapcsolatok feltárására és*
- *amennyire lehet, támogatja az ETO verziók közötti konverziót.*

Az első két követelmény technológiai jellegű. A továbbiak megvalósításához szükség van arra, hogy **eltároljuk**

- *a jelzet elemeit, tehát az azt alkotó ETO számokat,*
- *a sorrendjüket,*
- *a közöttük fennálló szintaktikai relációk típusát és*
- *minden olyan egyéb járulékos információt, amely a jelzetből kinyerhető, és jelentőséggel bírhat a visszakeresés, böngészés, illetve a jelzet értelmezése során.*

5.3 A jelzetelemek fája

Az itt ismertetett formátumnak kimondott célja, hogy ne csak az összetett jelzetek esetén tegye lehetővé a jelzetelemek visszakeresését, hanem a fázisrelációk és agglomeráció, esetleg alosztások

⁴⁴⁸ A formátum 1.0. verziója vázlatosan bemutatásra került az Agria Media 2014 konferencián (Piros 2015b) és a 2015. évi ETO Szemináriumon (Piros 2015a), a 2.0. verzió pedig az Extensions and Corrections utolsó számában (Piros 2018). Az itt ismertetett 2.1-es verzió újdonságai röviden össze lettek foglalva a 2017. évi ETO Szemináriumon (Piros 2017). A formátumnak az itt leírthoz hasonló részletességű ismertetése eddig még nem jelent meg.

egymásba ágyazása segítségével alkotott, több szintű jelzeteket is megfelelően kezelve, ilyen esetekben is *megőrizzé a részek kapcsolódási módjára vonatkozó információt*.

Ehhez elsősorban a *több szintű jelzetek hierarchikus felépítését* kell megvizsgálnunk. A jelzeteknek ezt a belső szerkezeti hierarchiáját a facetta-, fázis- és agglomeráció relációk közötti, az első részben már említett precedencia sorrend határozza meg. Ez a fejezet azt mutatja be, hogy ez a precedencia sorrend hogyan érvényesül az ETO speciális relációi esetén, illetve milyen formában reprezentálja a jelzetek struktúráját.

Korábban már volt arról szó, hogy a *kiterjesztést egy intervallumként értelmezhetjük*, amely a relációkban az egyszerű táblázati számokkal azonos módon használható. Megfordítva, *az egyszerű számokat tekinthetjük egy olyan intervallumnak, amelynek a kezdő- és végpontjai megegyeznek*. Ez a megoldás analóg azzal, amelyet a korábban bemutatott MARC formátumok is alkalmaznak. Előnye, hogy ilyen tárolás esetén egy kereső algoritmus könnyen felkészíthető az intervallumba eső valamennyi szám és azok hierarchikus leszármazottainak, illetve a részintervallumok visszakeresésére is.

Az ETO jelenlegi jelzetszerkesztési szemlélete a fő táblázati számokat és a független általánosan közös alosztásokat egyenrangúnak tekinti.⁴⁴⁹ Az ilyen *alosztások facettaként pontosíthatják a fő táblázati szám tartalmát, de önállóan állva is kifejezhetnek fogalmakat*, és nem kell szükségszerűen az általuk pontosított fogalom mögött állniuk, hanem akár az elé is kerülhetnek, a besorolás céljának megfelelően. Egy így alkotott *egyszerű vagy összetett tárgykör* – amely tehát legfeljebb egy fő táblázati számot vagy kiterjesztést és tetszőleges számú általánosan közös alosztást tartalmaz - mint önálló egység *operandusa lehet tetszőleges relációnak, illetve az összekötéssel jelölt agglomerációnak*. A több elemű jelzeten belül az aposztrófós jelzetszintézis és a csoportképzéssel (szögletes zárójel) alkotott jelzetek a fő táblázati számokkal és intervallumokkal azonos módon viselkednek a fogalmak alkotórészeiként.

A további összetett tárgykörök a *nem független általánosan, illetve korlátozottan közös alosztásoknak a táblázati számhoz (intervallumhoz) való kapcsolásával ábrázolhatók*.⁴⁵⁰

Korábban már láthattuk, hogy az általánosan közös *alosztások tartalmazhatnak összekötést, relációt és szintézist, illetve további táblázati számokat és intervallumokat*, amelyekhez újabb korlátozottan (esetenként általánosan) közös alosztások kapcsolódhatnak.

A fentiek szerint létrehozott egyszerű és összetett tárgykörökből *viszonyítással, illetve sorrendrögzítéssel komplex jelzetek állíthatók elő*. Az egyszerű, összetett és illetve komplex jelzetek pedig *összekötéssel agglomerált tárgykörökké vonhatók össze*. Ez elméletben a *hierarchia legmagasabb szintje*.

Összefoglalva, **az ETO alosztási jeleinek precedencia sorrendje** a gyengébbtől az erősebb felé haladva a következő:⁴⁵¹

1. Összekötés
2. Viszonyítás
3. Sorrendrögzítés

⁴⁴⁹ Babiczky 1985.

⁴⁵⁰ Ez megfelel Ranganathan már említett „kép-fal” elvének (Ranganathan 1967: 425-426). Ugyanígy ábrázolható a csoportképzés régebbi típusú, mellérendelő viszonyításnak megfelelő alkalmazása is (Babiczky 1975: 45-46).

⁴⁵¹ Piros 2015.

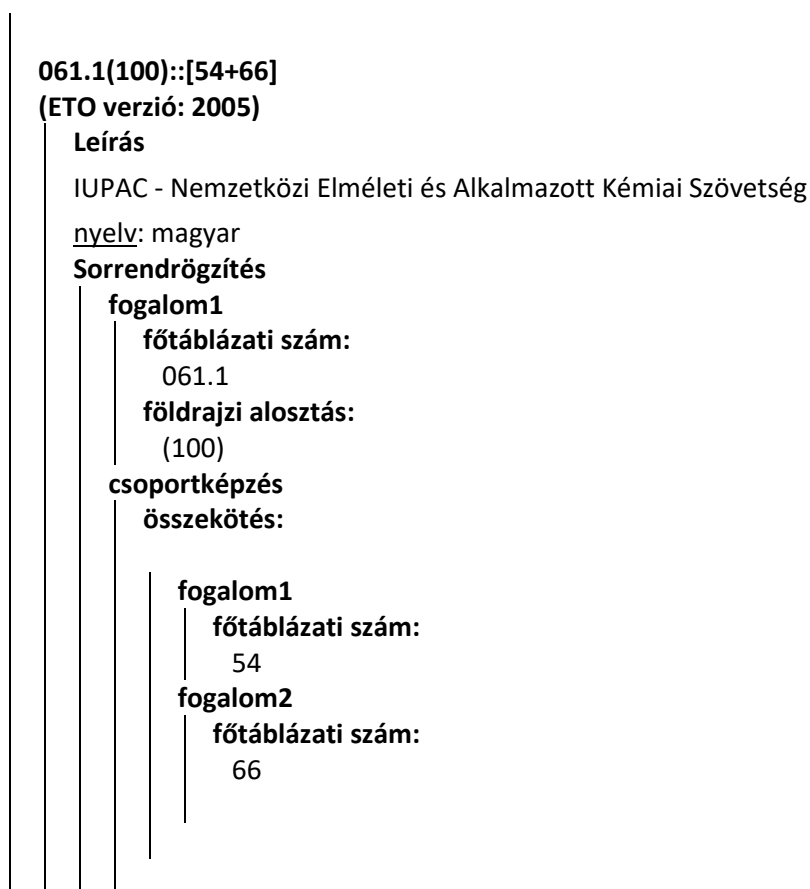
4. Csoportképzés

5. Jelzetszintézis

A *csoportképzés szerepe kicsit eltér* a többi műveletétől, mivel az első három relációban tényezőként is előfordulhat, miközben tartalmazhatja is azokat.

A *viszonyítás és sorrendrögzítés közötti sorrendet a szimmetria határozza meg*: mivel a viszonyítás szimmetrikus reláció, az általa képzett összetett jelzetek jelentése a sorrend változásával nem változik. Ha azonban egy kettőnél több elemből álló jelzetben csak – azonos szinten álló – viszonyítás és sorrendrögzítés fordul elő, akkor a sorrendrögzítés által összekötött elemek sorrendje nem változtatható, míg az összes többié igen. Tehát a sorrendrögzítés által alkotott összetett jelzet egységként, a viszonyítás egyik operandusaként viselkedik.

A fenti precedencia sorrend meghatároz egy **hierarchiát** is a műveletek között, melynek magasabb szintjein a gyengébb, alacsonyabb szintjein pedig az erősebb precedenciával rendelkező műveletek állhatnak. Ez alapján az alosztási jelek olyan fákként ábrázolhatók, melyek csúcsai az alosztási jelek, levelei pedig az egyszerű és összetett fogalmak. Ezek a fogalmak szintén ábrázolhatók fákként, melyeknek csúcsai az alosztások, csoportképzések és egyéb olyan jelzetelemek, amelyek további műveleteket is tartalmazhatnak. Ezeknek a fáknak a levelei már minden esetben táblázati számok, illetve a belőlük képzett intervallumok. A *teljes jelzet a részfák által alkotott teljes faként reprezentálható*, ahogy az a 7. ábrán lévő példán látható.⁴⁵²



⁴⁵² A példa valamennyi felhasznált ETO kiadásban megtalálható a csoportképzés bemutatására (ld. például OSZK KI 2005: 12).



7. ábra A 061.1(100)::[54+66] jelzet fa reprezentációja

5.4 Az XML nyelv előnyei

Ahogy az előző fejezetben láthattuk, az egyszerű és összetett ETO-jelzetek mint fák, egy hierarchikus adatszerkezetben ábrázolhatók. Ez egyszersmind azt is jelenti, hogy mint a fákhhoz általában, az ETO-jelzetekhez is készíthető egy-egy XML⁴⁵³ dokumentum, mely megőrzi azok szerkezetét és elemeit.

Az XML mint formátum **általános előnyei** a következők:⁴⁵⁴

- *önleíró* formátum
 - *mind gép, mind ember számára olvasható*
 - *rugalmas* és könnyen módosítható, mivel az adatok jelentését elsősorban nem az adat, hanem annak jelölői dokumentumon belüli elhelyezkedése határozza meg
- *komplex adatstruktúrák is leírhatók* benne
 - alkalmas a legáltalánosabb adatstruktúrák (rekord, lista, fa) ábrázolására és azok rugalmas egymásba ágyazására
- *biztosítja az adatintegritást*
 - a dokumentumnak meg kell felelnie az *XML szabályoknak*
 - ezen felül *további szabályok definiálhatók* például a Document Type Definition (DTD) és az XML Schema (XSD) nyelvek segítségével
- *platform független*
 - *sima szövegfájlban* valósul meg, licencektől és korlátozásoktól mentesen
 - *internetes szabványokon* alapul
 - támogatja a Unicode-ot, így alkalmas bármely információ bármilyen nyelven történő közzétételére
- *elterjedt*
 - széles körben és sokoldalúan használt formátum
 - *legtöbb programozási nyelv jól dokumentált támogatást ad* XML dokumentumok létrehozására és feldolgozására (beépítve vagy külső csomagokon keresztül)
 - szoftveresen erősen támogatott az XML fájlok megtekintése és manuális szerkesztése is.

Az XML-t a fenti előnyök különösen alkalmassá teszik arra, hogy az ETO-jelzeteket leírjuk vele a korábban lefektetett követelményeket kielégítő módon. Általánosabban pedig **definiálhatunk egy XML alapú leíró nyelvet**, mely meghatározza ezt a leírást, *valamennyi ETO-jelzetre egységesen alkalmazható* módon.

⁴⁵³ „Extensible Markup Language”.

⁴⁵⁴ Hiroshi (2007), Murell 2010: 114-118.

5.5 XML séma definíció ETO-jelzetek leírására

A korábbi fejezetekben láthattuk, hogy az ETO számítógépes felhasználását nagy mértékben megkönnyítene a jelzetek szerkezetének feltárása és rögzítése, illetve meghatároztuk a szükséges formátummal szembeni követelményeket. Az is kiderült, hogy az XML nyelv alkalmas egy ilyen formátum definiálására. A következő lépés *egy olyan XML formátum létrehozása, amely reflektál az ETO legspeciálisabb jelzetszerkesztési megoldásaira is, figyelembe véve az ETO táblázataiban és szabályaiban történt változásokat*. Ez a fejezet egy ilyen **formátum**⁴⁵⁵ **részletes bemutatását** tartalmazza.

A leírás az **XML Séma Definíció**⁴⁵⁶ ajánlás felhasználásával készült, amely különösen alkalmas egy ilyen nyelv leírására, mivel részletesen *definiálhatók benne a szükséges XML leírások elemei*, azok attribútumai, számuk, sorrendjük stb. Ezen kívül *validálási lehetőségei* lehetővé teszik az ETO-jelzetekben található esetleges szintaktikai hibák felismerését a jelzetek XML alapú leírásainak ellenőrzésével. Az *adatkötés*⁴⁵⁷ lehetősége pedig programozási szempontból jelent nagy segítséget a jelzetek leírásához, illetve a formátum feldolgozásához.

A séma definíció elkészítéséhez több különböző nyomtatott és online *ETO kiadás*⁴⁵⁸, *osztályozási tankönyvek*⁴⁵⁹, illetve a szakirodalomban megjelenő *cikkek*⁴⁶⁰ szolgáltak alapul. A *változások követésére* a fenti források mellett az „Extensions and Corrections to the UDC”⁴⁶¹, illetve az ETO konzorcium honlapján elérhető információk⁴⁶² lettek feldolgozva.

5.5.1 Áttekintés

A séma szerint az **ETO-jelzetet leíró dokumentum** egyetlen, ETO számmal leírt fogalmat reprezentáló elemet tartalmaz. Az elem **attribútumai** a fogalmat leíró *ETO-jelzet*, illetve az *ETO kiadás*, ami alapján az indexelő a jelzetet összeállította. Ezen kívül **elemként** tartalmazhatja az ETO-jelzet *természetes nyelvű leírásait, tetszőleges számban és nyelven*.

⁴⁵⁵ Az itt bemutatott séma definíció letölthető az alábbi URL-ről: <http://piros.udc-interpretor.hu/#xsd> és felhasználható Creative Commons Nevezd meg! - Ne add el! – Így add tovább! 4.0 Nemzetközi Licenz („CC BY-NC-ND 4.0”) szabályai szerint.

⁴⁵⁶ „XML Schema”.

⁴⁵⁷ Bourret 2011.

⁴⁵⁸ A munka során a következő ETO kiadások lettek felhasználva: a legutóbbi magyar nyelvű nyomtatott kiadások (OMIKK és OSZK-KMK 1990 és OSZK KI 2005), a British Standard Institute nyomtatott angol nyelvű szabvány kiadása (BSI 2005), a UDC Summary („UDC Summary”) és az angol nyelvű UDC Online („UDC Online”). A továbbiakban pontos helymegjelölést azoknak az ETO osztályoknak az esetében adunk, melyek nem fordulnak elő valamennyi itt felsorolt kiadásban.

⁴⁵⁹ A felhasznált tankönyvek: Babiczky Béla: Osztályozás és szakkatalógus szerkesztés (1975), Babiczky B. és Barátné Hajdu Ágnes: Bevezetés az információkereső nyelvek elméletébe és gyakorlatába (1998), Eszenyiné Borbély Mária, és Salgáné Medveczki Marianna. 2010. Komplex könyvtári feldolgozó munka gyakorlata (2010) és Pálvolgyi Mihály: Információfeldolgozás, információkereső nyelvek (1999).

⁴⁶⁰ A cikkek a felhasznált információ kapcsán vannak hivatkozva, Claudio Gnoli (2011) és Mándy Gábor (2011) írásai azonban külön is említést érdemelnek.

⁴⁶¹ „Extensions and Corrections to the UDC”.

⁴⁶² „Cancellations to the UDC” és „Major Changes to the UDC 1993-2013”.

Az **ETO szám szerkezetét** egy további elem írja le, melynek *típusa attól függ, hogy agglomerált, összetett, komplex vagy egyszerű* jelzetről van szó. Más szóval *ez az elem reprezentálja az ETO-jelzetet leíró fát*, az alosztási jelek korábbiakban leírt precedenciája alapján.

Az **egyszerű jelzetek** legfeljebb egy *főtáblázati számot* (intervallumot, aposztrófós jelzetszintézist vagy csoportképzést), és tetszőleges számú *önálló általánosan közös alosztást* reprezentáló elemet tartalmaznak, melyek további elemként tartalmazhatják a *hozzájuk kapcsolódó további alosztásokat*. Amennyiben egy alosztás további alosztási jelet tartalmaz, akkor az őt reprezentáló elem az ennek megfelelő részfat tartalmazza. Az ilyen **speciális megoldások minden esetben önálló típussal vannak leírva**.

A **kiterjesztéseket** leíró elemek *az intervallum végpontjait* tartalmazzák. A táblázati számok speciális kiterjesztésként vannak kezelve, amelyeknek kezdő és végpontjuk azonos, így elég az egyik attribútumnak értéket adni.

Az elemek attribútumként tartalmazhatnak olyan további adatokat, amelyek a kereséskor, feldolgozáskor, megjelenítéskor hasznosak lehetnek, például a jelzetben elfoglalt helyükre vagy a bevezető karakterekre vonatkozóan.

A fentiek szemléltetéséhez tekintsük a következő jelzetet⁴⁶³:

[792+82-2](091)(086.7) A színházi előadás és a dráma története. Hangfelvétel.

A fenti jelzet egy csoportképzést tartalmaz, azon belül egy összekötést, melynek második operandusához egy korlátozottan közös alosztás is kapcsolódik. A csoportképzéshez két formai alosztás tartozik, egyik a dokumentum „belső” (történelmi tárgyalásmód), másik a „külső” (hanganyag) formát határozza meg, ezért a sorrendjüknek is lehet jelentősége.

A fentiek alapján a jelzetet az alábbi XML-lel reprezentálhatjuk:

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE xml>
<ns:udc_concept
  xsi:schemaLocation="http://piros.udc-interpreter.hu/#xsd udc.xsd"
  xmlns:ns="http://piros.udc-interpreter.hu/#xsd"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  udc_edition="2005"
  notation="[792+82-2] (091) (086.7)">
  <ns:description xml:lang="EN">
    The history of theatre performances and dramas. Sound recording.
  </ns:description>
  <ns:description xml:lang="HU">
    A színházi előadás és a dráma története. Hangfelvétel.
  </ns:description>
  <ns:main_concept>
  <ns:main_table_subgrouping>
    <ns:main_table_addition>
      <ns:main_concept order="1">
        <ns:main_table_number number1="792"/>
      </ns:main_concept>
      <ns:main_concept order="2">
        <ns:main_table_number
          number1="82">
          <ns:special_auxiliary order="1">
            <ns:special_auxiliary_number xsi:type="ns:special_auxiliary_hyphen_number"
              number1="-2" />
          </ns:special_auxiliary>
        </ns:main_table_number>
      </ns:main_concept>
    </ns:main_table_addition>
  </ns:main_table_subgrouping>
</ns:main_concept>
```

⁴⁶³ A példa forrása a BSI angol nyelvű standard ETO kiadása (BSI 2005: 17)

```

        </ns:special_auxiliary>
    </ns:main_table_number>
</ns:main_concept>
</ns:main_table_addition>
<ns:common_auxiliary_independent
    xsi:type="ns:common_auxiliary_of_form" order="1">
    <ns:common_auxiliary_of_form_number
        number1="(091) " />
</ns:common_auxiliary_independent>
<ns:common_auxiliary_independent
    xsi:type="ns:common_auxiliary_of_form" order="2">
    <ns:common_auxiliary_of_form_number
        number1="(086.7) " />
</ns:common_auxiliary_independent>
</ns:main_table_subgrouping>
</ns:main_concept>
</ns:udc_concept>

```

5.5.2 Az ETO szám leírása XML dokumentumként

Ahogy az előző fejezetben láttuk, egy **ETO-jelzetet** az XML dokumentum egyetlen eleme reprezentál:

```

<xsd:element
    name="udc_concept"
    type="udc:udc_concept">
</xsd:element>

```

Az elemet leíró **komplex típus** az alábbi:

```

<xsd:complexType name="udc_concept">
<xsd:sequence>
    <xsd:element name="validation_errors"
        type="udc:validation_errors"
        minOccurs="0" maxOccurs="1">
    </xsd:element>
    <xsd:element name="description"
        type="udc:description"
        minOccurs="0"
        maxOccurs="unbounded">
    </xsd:element>
    <xsd:choice minOccurs="1" maxOccurs="1">
        <xsd:element name="main_concept"
            type="udc:main_concept"
            minOccurs="1" maxOccurs="1">
        </xsd:element>
        <xsd:element name="main_table_orderfixing"
            type="udc:main_table_orderfixing"
            minOccurs="1" maxOccurs="1"/>
        <xsd:element name="main_table_relation"
            type="udc:main_table_relation"
            minOccurs="1" maxOccurs="1"/>
        <xsd:element name="main_table_addition"
            type="udc:main_table_addition"
            minOccurs="1" maxOccurs="1"/>
    </xsd:choice>
</xsd:sequence>
<xsd:attribute name="notation"
    type="udc:udc_concept_notation_string"
    use="required">
</xsd:attribute>
<xsd:attribute name="udc_edition"
    type="udc:udc_edition_year_string">
</xsd:attribute>

```



```
</xsd:complexType>
```

A típus először is tartalmazhat tetszőleges számú **leírást**, ezeket az alábbi típus definiálja:

```
<xsd:complexType name="description">
  <xsd:simpleContent>
    <xsd:extension base="xsd:string">
      <xsd:attribute
        ref="xml:lang"
        use="required">
      </xsd:attribute>
    </xsd:extension>
  </xsd:simpleContent>
</xsd:complexType>
```

A fogalmat leíró típus további elemei az ETO-jelzetek már korábban kifejtett **hierarchikus szerkezetét** tükrözik. Ennek megfelelően egy egyszerű⁴⁶⁴ vagy összetett fogalom legfeljebb egy fő táblázati számot és a hozzá kapcsolódó alosztásokat, egy komplex vagy agglomerált fogalom pedig a képzéséhez használt legkülső általánosan (jelzetszintézis esetén korlátozottan) közös alosztási jelet leíró elemet tartalmaz.

Ahogy a típusdefinícióban is látható, a fogalmat leíró elem attribútumai az eredeti **ETO számot** és az **ETO kiadás évét** tartalmazzák, melyeket az alábbi egyszerű típusok írnak le:

```
<xsd:simpleType name="udc_concept_notation_string">
  <xsd:restriction base="xsd:string">
    <xsd:minLength value="1"/>
  </xsd:restriction>
</xsd:simpleType>

<xsd:simpleType name="udc_edition_year_string">
  <xsd:restriction base="xsd:string">
    <xsd:pattern value="( (19) | (2\d) ) \d{2}"/>
  </xsd:restriction>
</xsd:simpleType>
```

Az ETO kiadás évszámát elsősorban a különböző kiadások közötti különbségtétel miatt szükséges eltárolni, melyek egy jelzet jelentését, vagy akár érvényességét is meghatározhatják. A legpontosabb megoldás ebben az esetben az adatbázis azonosító lenne⁴⁶⁵, hiszen nem egyértelmű, hogy az évszám a bevezetés, az adatbázis kiadásának, esetleg az adott kiadás megjelenésének az évét jelöli⁴⁶⁶. Ehhez azonban az indexelőnek pontosan ismernie kell ezt az azonosítót.

5.5.3 Az egyszerű fogalmakat leíró típus

Ahogy azt korábban említettük, az ETO jelenlegi jelzetszerkesztési szemlélete a fő táblázati számokat és az önálló általánosan közös alosztásokat azonos szinten kezeli. Ezért *az egyszerű fogalmak esetén lehetővé kell tenni azt is, hogy fő táblázati számok vagy önálló általánosan közös alosztások önállóan, illetve együtt is szerepelhessenek az egyszerű fogalmakban.* Ezt az alábbi típussal írhatjuk le:

⁴⁶⁴ A korábbiak alapján az egyszerű fogalmak közé soroljuk a kiterjesztéssel kapott intervallumokat (agglomerált tárgyköröket) is.

⁴⁶⁵ Slavić és Isaac 2009.

⁴⁶⁶ Például a 2005. évi magyar kiadás a UDC Master Reference File 2000-ben lezárt kiadása alapján készült (OSZK KI 2005: 1/3), így számos eltérést tartalmaz a szintén 2005-ben kiadott, de a 2004-ben lezárt MRF anyaga alapján készült BSI Standard kiadáshoz képest (BSI 2005: vii).

```

<xsd:complexType name="main_concept">
  <xsd:sequence>
    <xsd:choice minOccurs="0" maxOccurs="1">
      <xsd:element name="main_table_number"
        type="udc:main_table_number"/>
      <xsd:element name="main_table_synthesis"
        type="udc:main_table_synthesis"/>
      <xsd:element name="main_table_subgrouping"
        type="udc:main_table_subgrouping" minOccurs="1" maxOccurs="1"/>
    </xsd:choice>
    <xsd:element name="common_auxiliary_independent"
      type="udc:common_auxiliary_independent"
      minOccurs="0" maxOccurs="unbounded"/>
    </xsd:element>
  </xsd:sequence>
  <xsd:attribute name="order" type="xsd:int" use="optional"/>
</xsd:complexType>

```

Látható, hogy a típus tartalmaz egy attribútumot a **sorrend** eltárolására is. A sorrendrögzítés és a szintézis hivatalosan nem kommutatív, tehát ezekben az esetekben a sorrendnek mindenképpen van jelentése. Az egyéb esetekben elsősorban a jelzet reprodukálhatósága miatt van erre szükség, de korábban már láthattuk, hogy az általánosan közös alosztások esetében is lehet a sorrendnek jelentést megkülönböztető szerepe.

5.5.4 A táblázati számokat leíró osztályok őssztályai

Az összes olyan típus, amely tetszőleges számú korlátozottan közös **alosztást** és névalosztást, illetve egy külső forrásból (tehát az ETO-n kívülről) származó jelzetet tartalmazhat, leszármaztatható az alábbi **őssztályból**, melynek a felsoroltak az elemei.

```

<xsd:complexType name="auxiliary_container_abstract" abstract="true">
  <xsd:sequence>
    <xsd:element name="special_auxiliary"
      type="udc:special_auxiliary" minOccurs="0" maxOccurs="unbounded"/>
    <xsd:element name="nonudc_notation"
      type="udc:nonudc_notation" minOccurs="0" maxOccurs="1"/>
    <xsd:element name="alphabetical_specification"
      type="udc:alphabetical_specification"
      minOccurs="0" maxOccurs="unbounded"/>
  </xsd:sequence>
</xsd:complexType>

```

A *fő- és általánosan közös alosztás* táblázati számokon kívül a fenti típus leszármazottja lesz a *jelzetszintézis* és a *csoportképzés* is.

A következő típus az előzőnek egy olyan *leszármazottja*, amely két attribútumot tartalmaz a *táblázati számok számára*. Ha a jelzet kiterjesztést (azaz intervallumot) tartalmaz, a két attribútum az **kiterjesztés** kezdő- és végpontját jelenti. Mivel a kiterjesztést jelölő jelzet többnyire csak az első számot tartalmazza teljes formában, a másodikat csak rövidítve, az utóbbit *ki kell egészítenünk*, mielőtt az XML-be kerülne, hogy a későbbi feldolgozó algoritmusok a valódi, teljes számok alapján működhessenek (illetve ne a feldolgozás fázisában kelljen minden egyes alkalommal a kiegészítésről gondoskodni). *Egyszerű táblázati szám esetén pedig csak az első attribútumot kell kitölteni*, a második üresen hagyható.

Ennek a típusnak a leszármazottai a fő- és az általánosan közös alosztás **táblázati számok**.

```

<xsd:complexType name="table_number" abstract="true">
  <xsd:complexContent>
    <xsd:extension base="udc:auxiliary_container_abstract">

```

```

    <xsd:attribute name="number1" type="xsd:string" use="optional">
  </xsd:attribute>
  <xsd:attribute name="number2" type="xsd:string" use="optional">
  </xsd:attribute>
</xsd:extension>
</xsd:complexContent>
</xsd:complexType>

```

5.5.5 A fő táblázati számokat leíró típusok

Fő táblázati számok leírásához a fenti őosztályokat először *úgy kell kiterjesztenünk*, hogy tartalmazhassanak **nem önálló általánosan közös alosztásokat** is (*beleértve a csoportképzést, amely a régi gyakorlat szerint állhat közvetlenül a fő táblázati számhoz kapcsolva is*).

```

<xsd:complexType name="main_table_number_abstract" abstract="true">
  <xsd:complexContent>
    <xsd:extension base="udc:table_number">
      <xsd:sequence>
        <xsd:element name="common_auxiliary_dependent"
          type="udc:common_auxiliary_dependent" minOccurs="0" maxOccurs="unbounded"/>
        <xsd:element name="main_table_subgrouping"
          type="udc:main_table_subgrouping" minOccurs="0" maxOccurs="unbounded"/>
      </xsd:sequence>
      <xsd:attribute name="parameter" type="xsd:string">
      </xsd:attribute>
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>

```

Az így kapott absztrakt típust a számok ellenőrizhetősége érdekében **szűkíthetjük** úgy, hogy **csak a megfelelő formátumú fő táblázati számokat** tartalmazhassa.

```

<xsd:complexType name="main_table_number">
  <xsd:complexContent>
    <xsd:restriction base="udc:main_table_number_abstract">
      <xsd:sequence>
        <xsd:element name="special_auxiliary"
          type="udc:special_auxiliary" minOccurs="0" maxOccurs="unbounded"/>
        <xsd:element name="nonudc_notation"
          type="udc:nonudc_notation" minOccurs="0" maxOccurs="1"/>
        <xsd:element name="alphabetical_specification"
          type="udc:alphabetical_specification" minOccurs="0" maxOccurs="unbounded"/>
        <xsd:element name="common_auxiliary_dependent"
          type="udc:common_auxiliary_dependent" minOccurs="0" maxOccurs="unbounded"/>
        <xsd:element name="main_table_subgrouping"
          type="udc:main_table_subgrouping" minOccurs="0" maxOccurs="unbounded"/>
      </xsd:sequence>
      <xsd:attribute name="number1" type="udc:main_table_number_string"
        use="required"/>
      <xsd:attribute name="number2" type="udc:main_table_number_string"
        use="optional"/>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>

```

Látható, hogy a táblázati számok egy, a **főtáblázati számok leírására** készült karakterlánc típus példányai. Ezt az **egyszerű típust** az alábbi formában definiálhatjuk:⁴⁶⁷

```
<xsd:simpleType name="main_table_number_string">
  <xsd:restriction base="xsd:string">
    <xsd:pattern value="([0-3]|[5-9])(\d{0,1}|(\d{2}(\.[0-9]\d{2})*)|(\.[0-9]\d{2}))?" />
  </xsd:restriction>
</xsd:simpleType>
```

Az alábbi osztály a fentiek olyan kiterjesztése, amely tartalmazza a *sorrendet* is:

```
<xsd:complexType name="main_table_number_ordered">
  <xsd:complexContent>
    <xsd:extension base="udc:main_table_number">
      <xsd:attribute name="order" type="xsd:integer" use="required"/>
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>
```

Előfordulhatnak olyan esetek is (például a csoportképzésen belül szereplő számokban), amikor egy fő táblázati szám *nem tartalmazhat csoportképzéssel hozzáfűzött másik fő táblázati számot*. Amennyiben egy ilyen számot alosztási jel köt össze egy másikkal, ebben az esetben is szükséges lehet a sorrend rögzítése. Ezt az esetet írja le az alábbi típus:

```
<xsd:complexType name="main_table_number_without_subgrouping">
  <xsd:complexContent>
    <xsd:restriction base="udc:main_table_number">
      <xsd:sequence>
        <xsd:element name="special_auxiliary"
          type="udc:special_auxiliary"
          minOccurs="0" maxOccurs="unbounded"/>
        <xsd:element name="nonudc_notation"
          type="udc:nonudc_notation"
          minOccurs="0"/>
        <xsd:element name="alphabetical_specification"
          type="udc:alphabetical_specification"
          minOccurs="0" maxOccurs="unbounded"/>
        <xsd:element name="common_auxiliary_dependent"
          type="udc:common_auxiliary_dependent"
          minOccurs="0" maxOccurs="unbounded"/>
      </xsd:sequence>
      <xsd:attribute name="number1" type="udc:main_table_number_string"
        use="optional"/>
      <xsd:attribute name="number2" type="udc:main_table_number_string"
        use="optional"/>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>
```

Végül fel kell készülnünk arra az esetre is, amikor a fő táblázati szám egyáltalán nem tartalmazhat általánosan közös alosztást. Erre tipikus példa az, amikor a fő táblázati szám maga is egy általánosan közös alosztáson belül áll, pl. kettősponttal kapcsolódva a (0) formai alosztáshoz. Ennek az esetnek is

⁴⁶⁷ A négyes főosztály – egyelőre – továbbra is betöltetlen, így ezzel a számjeggyel nem kezdődhet fő táblázati szám. A .0 karaktorsor hivatalosan korlátozottan közös alosztások bevezetésére szolgál, a gyakorlatban azonban – elsősorban a párhuzamos továbbosztások miatt – fő táblázati számok belsejében is állhat.

van olyan alosete, amikor szükséges eltárolni a *sorrendet* (például, ha szintézissel összekapcsolt számokról van szó), így két *típust* kell definiálnunk:

```
<xsd:complexType
name="main_table_number_without_common_auxiliaries_and_subgrouping">
  <xsd:complexContent>
    <xsd:restriction base="udc:main_table_number_without_subgrouping">
      <xsd:sequence>
        <xsd:element name="special_auxiliary"
          type="udc:special_auxiliary"
          minOccurs="0" maxOccurs="unbounded"/>
        <xsd:element name="nonudc_notation"
          type="udc:nonudc_notation"
          minOccurs="0"/>
        <xsd:element name="alphabetical_specification"
          type="udc:alphabetical_specification"
          minOccurs="0" maxOccurs="unbounded"/>
        <xsd:element name="common_auxiliary_independent"
          type="udc:common_auxiliary_independent"
          minOccurs="0" maxOccurs="0"/>
        <xsd:element name="common_auxiliary_dependent"
          type="udc:common_auxiliary_dependent"
          minOccurs="0" maxOccurs="0"/>
      </xsd:sequence>
      <xsd:attribute name="number1"
        type="udc:main_table_number_string"
        use="required"/>
      <xsd:attribute name="number2"
        type="udc:main_table_number_string"
        use="optional"/>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>

<xsd:complexType
name="main_table_number_without_auxiliaries_and_subgrouping_ordered">
  <xsd:complexContent>
    <xsd:restriction base="udc:main_table_number_ordered">
      <xsd:sequence>
        <xsd:element name="special_auxiliary"
          type="udc:special_auxiliary"
          minOccurs="0" maxOccurs="0"/>
        <xsd:element name="nonudc_notation"
          type="udc:nonudc_notation"
          minOccurs="0" maxOccurs="1"/>
        <xsd:element name="alphabetical_specification"
          type="udc:alphabetical_specification"
          minOccurs="0" maxOccurs="unbounded"/>
        <xsd:element name="common_auxiliary_independent"
          type="udc:common_auxiliary_independent"
          minOccurs="0" maxOccurs="0"/>
        <xsd:element name="common_auxiliary_dependent"
          type="udc:common_auxiliary_dependent"
          minOccurs="0" maxOccurs="0"/>
      </xsd:sequence>
      <xsd:attribute name="number1"
        type="udc:main_table_number_string" use="required"/>
      <xsd:attribute name="number2"
        type="udc:main_table_number_string" use="optional"/>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>
```

5.5.6 A korlátozottan közös alosztásokat leíró típusok

Az egyszerű, enumerikus, illetve párhuzamos továbbosztással képezhető korlátozottan közös alosztásokat (facettákat) az alábbi típusból származtatott osztályokkal írhatjuk le:

```
<xsd:complexType name="special_auxiliary_root">
  <xsd:attribute name="order" type="xsd:integer" use="required"/>
</xsd:complexType>
```

Bár a korlátozottan közös alosztások *sorrendje* elméletileg nem hordoz jelentést⁴⁶⁸, a gyakorlatban az ajánlástól való eltérés esetén „a fogalmi láncban nyomatékosabb jelentést kapnak az előrehozott elemek”⁴⁶⁹, ezért célszerű azt is elmenteni.

Az ilyen korlátozottan közös alosztások *egy táblázati számot vagy intervallumot tartalmazznak*, melyeket a fő- és általánosan közös alosztási táblázatokhoz hasonlóan írhatunk le:

```
<xsd:complexType name="special_auxiliary">
  <xsd:complexContent>
    <xsd:extension base="udc:special_auxiliary_root">
      <xsd:sequence>
        <xsd:element name="special_auxiliary_number"
          type="udc:special_auxiliary_number"/>
      </xsd:sequence>
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>

<xsd:complexType name="special_auxiliary_number" abstract="true">
  <xsd:attribute name="number1"
    type="xsd:string" use="optional">
  </xsd:attribute>
  <xsd:attribute name="number2"
    type="xsd:string" use="optional">
  </xsd:attribute>
</xsd:complexType>
```

A fenti típusokat tovább szűkíthetjük a korlátozottan közös alosztás típusának megfelelően.

Az enumerikus **kötőjeles** korlátozottan közös alosztások⁴⁷⁰ leírhatók az alábbi típussal:

```
<xsd:complexType name="special_auxiliary_hyphen">
  <xsd:complexContent>
    <xsd:restriction base="udc:special_auxiliary">
      <xsd:sequence>
        <xsd:element name="special_auxiliary_number"
          type="udc:special_auxiliary_hyphen_number"/>
      </xsd:sequence>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>

<xsd:complexType name="special_auxiliary_hyphen_number">
  <xsd:complexContent>
```

⁴⁶⁸ Vannak kivételek, például az irodalmi művek leírásánál kötött a betűjeles és a korlátozottan közös alosztások sorrendje (OSZK KI 2005: 2/529., illetve BSI 2005: 890.).

⁴⁶⁹ OSZK KI 2005: 1/7.

⁴⁷⁰ Az egyetlen nem enumerikus kötőjeles korlátozottan közös alosztás a [355/359]-5/-9, amelyről a jelzetszintézis kapcsán lesz szó.

```

<xsd:restriction base="udc:special_auxiliary_number">
  <xsd:attribute name="number1"
    type="udc:special_auxiliary_hyphen_number_string"
    use="required"/>
  <xsd:attribute name="number2"
    type="udc:special_auxiliary_hyphen_number_string"
    use="optional"/>
</xsd:restriction>
</xsd:complexContent>
</xsd:complexType>

<xsd:simpleType name="special_auxiliary_hyphen_number_string">
<xsd:restriction base="udc:special_auxiliary_number_string">
  <xsd:minLength value="2"/>
  <xsd:pattern value="-\d(\d{0,1}|(\d{2}(\.[1-9]\d{2}))*(\.[1-9]\d{0,2}))?" />
</xsd:restriction>
</xsd:simpleType>

```

A következő típusokat használhatjuk a **pont-nullás** speciális alosztások leírására:

```

<xsd:complexType name="special_auxiliary_pointnought">
  <xsd:annotation>
    <xsd:documentation xml:lang="en">
      Use this complex type to describe special auxiliaries which are initiated with
      character literal '.0'.
    </xsd:documentation>
  </xsd:annotation>
  <xsd:complexContent>
    <xsd:restriction base="udc:special_auxiliary">
      <xsd:sequence>
        <xsd:element name="special_auxiliary_number"
          type="udc:special_auxiliary_pointnought_number"/>
      </xsd:sequence>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>

<xsd:complexType name="special_auxiliary_pointnought_number">
  <xsd:complexContent>
    <xsd:restriction base="udc:special_auxiliary_number">
      <xsd:attribute name="number1"
        type="udc:special_auxiliary_pointnought_number_string"
        use="required"/>
      <xsd:attribute name="number2"
        type="udc:special_auxiliary_pointnought_number_string"
        use="optional"/>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>

<xsd:simpleType name="special_auxiliary_pointnought_number_string">
<xsd:restriction base="udc:special_auxiliary_number_string">
  <xsd:minLength value="2"/>
  <xsd:pattern value="\.0(\d{0,2}(\.[1-9]\d{2}))*(\.[1-9]\d{0,2}))?" />
</xsd:restriction>
</xsd:simpleType>

```

Az **aposztróf**os korlátozottan közös alosztások esetén az enumerikus esetekben használható az alábbi típus:

```

<xsd:complexType name="special_auxiliary_apostrophe">
  <xsd:complexContent>

```

```

<xsd:restriction base="udc:special_auxiliary">
  <xsd:sequence>
    <xsd:element name="special_auxiliary_number"
      type="udc:special_auxiliary_apostrophe_number"/>
  </xsd:sequence>
</xsd:restriction>
</xsd:complexContent>
</xsd:complexType>

<xsd:complexType name="special_auxiliary_apostrophe_number">
  <xsd:complexContent>
    <xsd:restriction base="udc:special_auxiliary_number">
      <xsd:attribute name="number1"
        type="udc:special_auxiliary_apostrophe_number_string" use="required"/>
      <xsd:attribute name="number2"
        type="udc:special_auxiliary_apostrophe_number_string" use="optional"/>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>

<xsd:simpleType name="special_auxiliary_apostrophe_number_string">
  <xsd:restriction base="udc:special_auxiliary_number_string">
    <xsd:minLength value="2"/>
    <xsd:pattern value="'\d(\d{0,1}|(\d{2}(\.[1-9]\d{2}))*(\.[1-9]\d{0,2}))'"/>
  </xsd:restriction>
</xsd:simpleType>

```

Legvégül a közvetlen **számvégződéses** alosztások leírására szolgáló típusok következnek.⁴⁷¹

```

<xsd:complexType name="special_auxiliary_finaldigits">
  <xsd:complexContent>
    <xsd:restriction base="udc:special_auxiliary">
      <xsd:sequence>
        <xsd:element name="special_auxiliary_number"
          type="udc:special_auxiliary_finaldigits_number"/>
      </xsd:sequence>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>

<xsd:complexType name="special_auxiliary_finaldigits_number">
  <xsd:complexContent>
    <xsd:restriction base="udc:special_auxiliary_number">
      <xsd:attribute name="number1"
        type="udc:special_auxiliary_finaldigits_number_string" use="required"/>
      <xsd:attribute name="number2"
        type="udc:special_auxiliary_finaldigits_number_string" use="optional"/>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>

<xsd:simpleType name="special_auxiliary_finaldigits_number_string">
  <xsd:restriction base="udc:special_auxiliary_number_string">
    <xsd:minLength value="1"/>
    <xsd:pattern value="\d{1}|\d{2}|(\d{3}(\.[1-9]\d{2}))*(\.[1-9]\d{0,2})'"/>
  </xsd:restriction>

```

⁴⁷¹ Az ilyen alosztások a táblázatok legtöbb helyén már kivonásra kerültek („Revised UDC tables 2000”), éppen a számítógépes kezelésük nehézségei miatt.

</xsd:simpleType>

Láthatjuk, hogy a fenti, enumerikus korlátozottan közös alosztásokat leíró típusok lényegében *csak a táblázati számok formátumában térnek el egymástól*. A jelzetek összekapcsolását lehetővé tevő korlátozottan közös alosztás, az úgynevezett jelzetszintézis azonban teljesen eltérő megközelítést igényel.

5.5.6.1 Jelzetszintézis

Ahogy korábban már szó volt róla, a jelzetszintézis olyan korlátozottan közös alosztás amely *összetett jelzetek képzését teszi lehetővé a táblázat egy adott pontján álló számok között*.⁴⁷²

Az így alkotott fogalom a *fő táblázati számokhoz hasonlóan* pontosítható az alosztásokkal és a relációval, illetve bővíthető az összekötés segítségével, ezért az azt leíró típusnak is hasonló elemeket kell tartalmaznia.

A jelzetszintézist a következő típussal írhatjuk le:

```
<xsd:complexType name="main_table_synthesis">
  <xsd:complexContent>
    <xsd:extension base="udc:auxiliary_container_abstract">
      <xsd:sequence>
        <xsd:element name="main_table_number"
          type="udc:main_table_number_without_auxiliaries_and_subgrouping_ordered"
          minOccurs="2" maxOccurs="unbounded"/>
        <xsd:element name="special_auxiliary"
          type="udc:special_auxiliary"
          minOccurs="0" maxOccurs="unbounded"/>
        <xsd:element name="alphabetical_specification"
          type="udc:alphabetical_specification"
          minOccurs="0"/>
        <xsd:element name="nonudc_notation"
          type="udc:nonudc_notation"
          minOccurs="0"/>
        <xsd:element name="common_auxiliary_independent"
          type="udc:common_auxiliary_independent"
          minOccurs="0" maxOccurs="unbounded"/>
        <xsd:element name="common_auxiliary_dependent"
          type="udc:common_auxiliary_dependent"
          minOccurs="0" maxOccurs="unbounded"/>
        <xsd:element name="main_table_subgrouping"
          type="udc:main_table_subgrouping"
          minOccurs="0" maxOccurs="unbounded"/>
      </xsd:sequence>
      <xsd:attribute name="punctuation_mark" use="required">
        <xsd:simpleType>
          <xsd:restriction base="xsd:string">
            <xsd:length value="1"/>
            <xsd:pattern value="('|-)"/>
          </xsd:restriction>
        </xsd:simpleType>
      </xsd:attribute>
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>
```

⁴⁷² Az aposztrófós jelzetszintézis jelentősége az utóbbi időben nagy mértékben csökkent. Míg a 2005. évi magyar kiadás még tartalmazza valamennyi itt említett alosztást (OSZK KI 2005: 239, 315, illetve 435), a 2005. évi BSI kiadásban csak a politikai pártoknál és szervetlen kémiai vegyületeknél szerepelnek ilyen alosztások (BSI 2005: 240), a UDC Online pedig már csak az utóbbiak esetében teszi lehetővé ezt a típusú jelzetkapcsolást („UDC Online”). Az aposztrófós jelzetszintézis kérdéseit részletesen tárgyalja Wimmer (2000).

```

        </xsd:complexContent>
    </xsd:complexType>

    <xsd:complexType
name="main_table_number_without_auxiliaries_and_subgrouping_ordered">
    <xsd:complexContent>
        <xsd:restriction base="udc:main_table_number_ordered">
            <xsd:sequence>
                <xsd:element name="special_auxiliary"
                    type="udc:special_auxiliary"
                    minOccurs="0" maxOccurs="0"/>
                <xsd:element name="nonudc_notation"
                    type="udc:nonudc_notation"
                    minOccurs="0" maxOccurs="1"/>
                <xsd:element name="alphabetical_specification"
                    type="udc:alphabetical_specification"
                    minOccurs="0" maxOccurs="unbounded"/>
                <xsd:element name="common_auxiliary_independent"
                    type="udc:common_auxiliary_independent"
                    minOccurs="0" maxOccurs="0"/>
                <xsd:element name="common_auxiliary_dependent"
                    type="udc:common_auxiliary_dependent"
                    minOccurs="0" maxOccurs="0"/>
            </xsd:sequence>
            <xsd:attribute name="number1"
                type="udc:main_table_number_string"
                use="required"/>
            <xsd:attribute name="number2"
                type="udc:main_table_number_string" use="optional"/>
        </xsd:restriction>
    </xsd:complexContent>
</xsd:complexType>

```

Az alosztások mellett a típusnak *elemként* tartalmaznia kell kettő vagy több fő táblázati számot, melyek a **szintézis által összekötött számokat** jelölik.

A szintézissel összekapcsolt fő táblázati számok *csak külső forrásból származó, illetve névalosztásokat tartalmazhatnak, korlátozottan közösek nem*. Ezért egy külön osztályt kell létrehoznunk a számukra, amely a korlátozások mellett egy, a *sorrendet* tároló attribútummal is ki van egészítve. Erre itt is elsősorban az eredeti jelzet reprodukálásához van szükség, illetve a fogalmi láncon belüli sorrend esetleges jelentést árnyaló szerepe miatt (például a politikai pártok leírásánál).

A szintézis, akárcsak a kiterjesztés, többnyire csak az első számot tartalmazza teljes formában, a többi csak rövidítve, ezért ezeknek a *számoknak a kiegészítéséről* is gondoskodni kell az XML összeállítása során.

A fentiek mellett még egy attribútum érdemel említést. A *központozási jel* eltárolására („punctuation_mark”) azért van szükség, mert van olyan kivételes eset, amikor a szintézist nem aposztróf, hanem kötőjel használatával lehet képezni.⁴⁷³

5.5.7 Az általánosan közös alosztásokat leíró típusok

5.5.7.1 Önálló általánosan közös alosztások

Korábban már említettük, hogy az ETO jelenlegi szabályai szerint az önálló általánosan közös alosztások – szemben a nem önálló általánosan közös alosztásokkal – *fő táblázati számok nélkül*,

⁴⁷³ Az említett megoldásra példa a [355/359]-5/-9 korlátozottan közös alosztás (OSZK KI 2005: 1/315).

önállóan is szerepelhetnek az ETO-jelzetekben. Az utóbbi eset azt jelenti, hogy a „main_concept” objektum nem tartalmaz fő táblázati számot leíró elemet, csak egy vagy több alosztást.

Az általánosan közös alosztásokon belül egyszerű táblázati számok, komplex, összetett és agglomerált fogalmak egyaránt szerepelhetnek. Bizonyos esetekben különböző típusú alosztásokból álló, esetleg fő táblázati számokat is tartalmazó összetett fogalmak kifejezésére is van lehetőség egy alosztási jelen belül.

Az ilyen lehetőségek miatt külön kell választanunk az általánosan közös alosztás és alosztási szám fogalmát: előbbi az alosztáson belüli fogalmat jelöli, utóbbi pedig azokat a táblázati számokat, melyekből az felépül.

Mivel a lehetséges jelzetszerkesztési megoldások alosztási táblázatonként jelentősen eltérnek, az egyes alosztások számára önálló típusokat kell definiálnunk, melyek az alább látható őstípusból vannak származtatva. A sorrend elmentésére ebben az esetben is a jelzet reprodukálhatósága, illetve az esetleges jelentésbeli különbségek miatt van szükség.⁴⁷⁴ Más attribútumra ezen a szinten nincs szükség.

```
<xsd:complexType name="common_auxiliary_independent" abstract="true">
  <xsd:attribute name="order" type="xsd:int" use="required"/>
</xsd:complexType>
```

A következő absztrakt típus pedig az alosztási számokat leírók közös őse:

```
<xsd:complexType name="common_auxiliary_number_abstract" abstract="true">
  <xsd:complexContent>
    <xsd:restriction base="udc:table_number">
      <xsd:sequence>
        <xsd:element name="special_auxiliary"
          type="udc:special_auxiliary"
          minOccurs="0" maxOccurs="unbounded"/>
        <xsd:element name="nonudc_notation"
          type="udc:nonudc_notation"
          minOccurs="0" maxOccurs="1"/>
        <xsd:element name="alphabetical_specification"
          type="udc:alphabetical_specification"
          minOccurs="0" maxOccurs="1"/>
      </xsd:sequence>
      <xsd:attribute name="number1" type="xsd:string" use="required"/>
      <xsd:attribute name="number2" type="xsd:string" use="optional"/>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>
```

A típus a „table_number” leszármazottja, ennek megfelelően tartalmazza a táblázati számot vagy intervallumot attribútumként, emellett pedig tetszőleges számú korlátozottan közös alosztást, és opcionálisan egy-egy külső forrásból származó jelzetet, illetve névalosztást.

5.5.7.1.1 Nyelvi általánosan közös alosztások

A nyelvi általánosan közös alosztások között *nincsenek definiálva relációk*, így egy ilyen alosztás mindig egy **táblázati számot vagy intervallumot** tartalmaz, a hozzá esetlegesen kapcsolódó **korlátozottan közös alosztásokkal** (melyek bizonyos esetekben tartalmazhatnak máshonnan származó számokat). Ezért egy **nyelvi alosztás** az alábbi *összetett típussal* írható le:

⁴⁷⁴ Robinson 2002, Robinson 2003 és BSI 2005: 17.

```

<xsd:complexType name="common_auxiliary_of_language">
  <xsd:complexContent>
    <xsd:extension base="udc:common_auxiliary_independent">
      <xsd:sequence>
        <xsd:element name="common_auxiliary_of_language_number"
          type="udc:common_auxiliary_of_language_number"
          minOccurs="1" maxOccurs="1"/>
      </xsd:sequence>
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>

```

Amennyiben egy dokumentum **többnyelvű** vagy valamilyen más okból több nyelvi alosztást tartalmaz, azok *mindegyikét külön elemmel kell leírni*.

Az alosztáson belül szereplő **táblázati számot** leíró típust az alábbi módon definiálhatjuk:

```

<xsd:complexType name="common_auxiliary_of_language_number_root">
  <xsd:annotation>
    <xsd:documentation xml:lang="en">
      Complex type to describe parts of auxiliaries of language.
    </xsd:documentation>
  </xsd:annotation>
  <xsd:complexContent>
    <xsd:restriction base="udc:common_auxiliary_number_abstract">
      <xsd:sequence>
        <xsd:element name="special_auxiliary"
          type="udc:special_auxiliary"
          minOccurs="0" maxOccurs="0"/>
      </xsd:sequence>
      <xsd:attribute name="number1"
        type="udc:common_auxiliary_of_language_number_string"
        use="required"/>
      <xsd:attribute name="number2"
        type="udc:common_auxiliary_of_language_number_string"
        use="optional"/>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>

<xsd:complexType name="common_auxiliary_of_language_number">
  <xsd:complexContent>
    <xsd:extension base="udc:common_auxiliary_of_language_number_root">
      <xsd:sequence>
        <xsd:element name="special_auxiliary_language"
          type="udc:special_auxiliary_language"
          minOccurs="0" maxOccurs="unbounded"/>
      </xsd:sequence>
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>

```

Az első típusban a használható számok formátumát pontosítottuk és letiltottuk a hagyományos korlátozottan közös alosztásokat. A második típusra a *csak a nyelvi alosztásokban használható korlátozottan közös alosztások* bevezetéséhez van szükség.

A nyelvi alosztás **táblázataiban szereplő számokat** az alábbi egyszerű típus írja le:

```

<xsd:simpleType name="common_auxiliary_of_language_number_string">
  <xsd:restriction base="xsd:string">
    <xsd:pattern value="=00"/>
  </xsd:restriction>
</xsd:simpleType>

```

```

<xsd:pattern value="03(0)?\.\d{0,1}(\d{2}(\.[1-9]\d{2}))*(\.[1-9]\d{0,2})?)?"/>
<xsd:pattern value="[0-9](\d{0,1}(\d{2}(\.[1-9]\d{2}))*(\.[1-9]\d{0,2})?)?"/>
</xsd:restriction>
</xsd:simpleType>

```

A „=00” a soknyelvű dokumentumokat jelenti, a „=30” (illetve korábban „=3”) használatával egy meghatározott nyelvből **fordított művek** részére képezhetők komplex jelzetek. Ez utóbbiak egységes számként szerepelnek, mint a párhuzamos továbbosztással képzett számok általában.

A fordítások jelölése és az enumerikus (kötőjeles és aposztrófós) korlátozottan közös alosztások mellett az =...’276/’282 speciális korlátozottan közös alosztás használható a „gyakorlati nyelvhasználat, **dialektusok és nyelvváltozatok**” jelölésére.

Ezekben az alosztásokban *földrajzi és etnikai alosztások* (a dialektusok és tájnyelvek leírásakor), *személyek szerinti alosztások* (szociális dialektusok, szlengek, zsargonok leírásakor) illetve, kettősponttal bevezetve akár *főttáblázati számok* is megjelenhetnek.⁴⁷⁵ A fordításokkal szemben ezek nem párhuzamos alosztásokkal képzett egységes számok, hanem heterogén számok speciális jelzetkapcsolásai. Ezért itt *célszerű feltárni az alosztás teljes felépítését*, az elemeivel és azok kapcsolódási módjaival.

A fentiek alapján a nyelvek **korlátozottan közös alosztásait** leíró komplex típus a következő:

```

<xsd:complexType name="special_auxiliary_language">
  <xsd:complexContent>
    <xsd:extension base="udc:special_auxiliary_root">
      <xsd:choice>
        <xsd:element name="special_auxiliary_number"
          type="udc:special_auxiliary_number"/>
        <xsd:element name="special_auxiliary_language_dialect_relation"
          type="udc:special_auxiliary_language_dialect_relation"/>
        <xsd:sequence>
          <xsd:element name="special_auxiliary_dialect_number"
            type="udc:special_auxiliary_language_dialect_number"/>
          <xsd:choice minOccurs="1" maxOccurs="1">
            <xsd:element name="common_auxiliary_of_place"
              type="udc:common_auxiliary_of_place"
              minOccurs="1" maxOccurs="1"/>
            <xsd:element name="common_auxiliary_of_ethnics"
              type="udc:common_auxiliary_of_ethnics"
              minOccurs="1" maxOccurs="1"/>
            <xsd:element name="common_auxiliary_of_persons"
              type="udc:common_auxiliary_of_persons_number"
              minOccurs="1" maxOccurs="1"/>
          </xsd:choice>
        </xsd:sequence>
      </xsd:choice>
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>

```

⁴⁷⁵ Ld. <http://udcdata.info/000016> („UDC Summary”), BSI 2005: 5. és „UDC Online”.

Egyes szlengek leírására *viszonyítást*⁴⁷⁶ kell használni. Egy ilyen viszonyítást a következő összetett típussal írhatunk le⁴⁷⁷:

```
<xsd:complexType name="special_auxiliary_language_dialect_relation">
  <xsd:sequence>
    <xsd:element name="special_auxiliary_number"
      type="udc:special_auxiliary_language_dialect_number"/>
    <xsd:choice minOccurs="1" maxOccurs="1">
      <xsd:element name="main_table_number"
        type="udc:main_table_number_without_common_auxiliaries_and_subgrouping"
        minOccurs="1" maxOccurs="1"/>
      <xsd:element name="main_table_synthesis"
        type="udc:main_table_synthesis_without_common_auxiliaries_and_subgrouping"
        minOccurs="1" maxOccurs="1"/>
      <xsd:element name="main_table_subgrouping"
        type="udc:main_table_subgrouping_without_auxiliaries"/>
    </xsd:choice>
  </xsd:sequence>
</xsd:complexType>
```

Amennyiben a dialektus egy másik általánosan közös (például földrajzi, etnikai vagy személyek szerinti) alosztással van kifejezve, azt egy korlátozottan közös *alosztási szám és megfelelő alosztás szekvenciájával* írhatjuk le a nyelvi alosztást leíró osztályon belül:

```
<xsd:sequence>
  <xsd:element name="special_auxiliary_dialect_number"
    type="udc:special_auxiliary_language_dialect_number"/>
  <xsd:choice minOccurs="1" maxOccurs="1">
    <xsd:element name="common_auxiliary_of_place"
      type="udc:common_auxiliary_of_place"
      minOccurs="1" maxOccurs="1"/>
    <xsd:element name="common_auxiliary_of_ethnics"
      type="udc:common_auxiliary_of_ethnics"
      minOccurs="1" maxOccurs="1"/>
    <xsd:element name="common_auxiliary_of_persons"
      type="udc:common_auxiliary_of_persons_number"
      minOccurs="1" maxOccurs="1"/>
  </xsd:choice>
</xsd:sequence>
```

A nyelvváltozatok leírásakor a viszonyításon, illetve a *szekvencián belüli korlátozottan közös alosztási szám* leírásához az alábbi komplex típust használtuk.

```
<xsd:complexType name="special_auxiliary_language_dialect_number">
  <xsd:complexContent>
    <xsd:restriction base="udc:special_auxiliary_apostrophe_number">
      <xsd:attribute name="number1"
        type="udc:special_auxiliary_language_dialect_number_string"
        use="required"/>
      <xsd:attribute name="number2"
        type="udc:special_auxiliary_language_dialect_number_string"
        use="optional"/>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>
```

⁴⁷⁶ Ld. =...’272.6 („UDC Summary”) és <http://udcdata.info/076293> („UDC Summary”).

⁴⁷⁷ Ugyan a szabályok részletesen nem térnek ki rá, de természetesen fel kell készülnünk arra, hogy a viszonyítás az alosztást egy csoportképzéssel vagy egy szintézissel köti össze. Előbbire tetszőleges heterogén csoport, utóbbira egy politikai csoportosulás által használt egyéni nyelvezet kifejezése is példa lehet.

```
</xsd:complexType>
```

A fenti típusban használt **számformátumot** pedig az aposztrófes korlátozottan közös alosztásokból származtatott egyszerű típus írja le:

```
<xsd:simpleType name="special_auxiliary_language_dialect_number_string">
  <xsd:restriction base="udc:special_auxiliary_apostrophe_number_string">
    <xsd:minLength value="4"/>
    <xsd:pattern value="'2(76|82)((\.[1-9]\d{2})*(\.[1-9]\d{0,2})?)"/>
  </xsd:restriction>
</xsd:simpleType>
```

5.5.7.1.2 Formai általánosan közös alosztások

A formai általánosan közös alosztások leírása az alábbi **összetett típussal** történhet meg:

```
<xsd:complexType name="common_auxiliary_of_form">
  <xsd:complexContent>
    <xsd:extension base="udc:common_auxiliary_independent">
      <xsd:sequence>
        <xsd:choice minOccurs="1" maxOccurs="1">
          <xsd:element name="common_auxiliary_of_form_addition"
            type="udc:common_auxiliary_of_form_addition"
            minOccurs="1" maxOccurs="1"/>
          <xsd:element name="common_auxiliary_of_form_relation"
            type="udc:common_auxiliary_of_form_relation"
            minOccurs="1" maxOccurs="1"/>
          <xsd:element name="common_auxiliary_of_form_number"
            type="udc:common_auxiliary_of_form_number"
            minOccurs="1" maxOccurs="1"/>
        </xsd:choice>
      </xsd:sequence>
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>
```

Amennyiben az alosztás egyetlen **táblázati számot** tartalmaz (a hozzá tartozó korlátozottan közös alosztásokkal, külső forrásból származó, illetve névalosztással), akkor az azt leíró elem az alábbi típus egy példányát kell, hogy tartalmazza:

```
<xsd:complexType name="common_auxiliary_of_form_number">
  <xsd:complexContent>
    <xsd:restriction base="udc:common_auxiliary_number_abstract">
      <xsd:sequence>
        <xsd:element name="special_auxiliary"
          type="udc:special_auxiliary"
          minOccurs="0" maxOccurs="unbounded"/>
        <xsd:element name="nonudc_notation"
          type="udc:nonudc_notation"
          minOccurs="0" maxOccurs="1"/>
        <xsd:element name="alphabetical_specification"
          type="udc:alphabetical_specification"
          minOccurs="0" maxOccurs="1"/>
      </xsd:sequence>
      <xsd:attribute name="number1"
        type="udc:common_auxiliary_of_form_number_string"
        use="required"/>
      <xsd:attribute name="number2"
        type="udc:common_auxiliary_of_form_number_string"
        use="optional"/>
    </xsd:restriction>
  </xsd:complexContent>
```

```
</xsd:complexType>
```

A formai alosztások táblázatain belül egy speciális, szintagmatikus jelzetszetkesztési mód áll rendelkezésre olyan dokumentum formátumok kifejezésére, amelyek nem szerepelnek a táblázatokban. Az ilyen esetekben a zárójelen belül álló nullás számjegyet kell **viszonyítással** összekötni a szükséges fő táblázati számmal (esetleg intervallummal vagy szintézissel).

Egy ilyen jelzetet az alábbi típussal fejezhetünk ki:

```
<xsd:complexType name="common_auxiliary_of_form_relation">
  <xsd:sequence>
    <xsd:element name="common_auxiliary_of_form_number"
      type="udc:common_auxiliary_of_form_number_nought"/>
    <xsd:choice>
      <xsd:element name="main_table_number"
        type="udc:main_table_number_without_common_auxiliaries_and_subgrouping"/>
      <xsd:element name="main_table_synthesis"
        type="udc:main_table_synthesis_without_common_auxiliaries_and_subgrouping"/>
    </xsd:choice>
    <xsd:element name="special_auxiliary"
      type="udc:special_auxiliary"
      minOccurs="0" maxOccurs="unbounded"/>
    <xsd:element name="nonudc_notation"
      type="udc:nonudc_notation"
      minOccurs="0"/>
    <xsd:element name="alphabetical_specification"
      type="udc:alphabetical_specification"
      minOccurs="0"/>
  </xsd:sequence>
</xsd:complexType>
```

Azt az esetet, amikor egy zárójelen belül, **pluszjellel** összekapcsolva szerepel több formai alosztási szám az alábbi összetett típussal írhatjuk le:

```
<xsd:complexType name="common_auxiliary_of_form_addition">
  <xsd:choice minOccurs="2" maxOccurs="unbounded">
    <xsd:element name="common_auxiliary_of_form_number"
      type="udc:common_auxiliary_of_form_number"
      minOccurs="1" maxOccurs="1"/>
    <xsd:element name="common_auxiliary_of_form_relation"
      type="udc:common_auxiliary_of_form_relation"
      minOccurs="1" maxOccurs="1"/>
  </xsd:choice>
</xsd:complexType>
```

Legvégül a táblázati **számok formátumát** az alábbi **egyszerű típus** írja le:

```
<xsd:simpleType name="common_auxiliary_of_form_number_string">
  <xsd:restriction base="xsd:string">
    <xsd:pattern value="\ (0(\d|(\d{2}(\.[1-9]\d{2}))*(\.[1-9]\d{0,2})?)?)\ \d{2}"/>
  </xsd:restriction>
</xsd:simpleType>
```


5.5.7.1.3 Hely szerinti (földrajzi) általánosan közös alosztások

A földrajzi hely szerinti alosztások tartalmazzák a legtöbb táblázati számot és speciális jelzetalkotási lehetőséget.⁴⁷⁸

A teljes alosztást leíró összetett típus a következő:

```
<xsd:complexType name="common_auxiliary_of_place">
  <xsd:complexContent>
    <xsd:extension base="udc:common_auxiliary_independent">
      <xsd:choice>
        <xsd:element name="common_auxiliary_of_place_number"
          type="udc:common_auxiliary_of_place_number"/>
        <xsd:element name="common_auxiliary_of_geography_addition"
          type="udc:common_auxiliary_of_geography_addition"/>
        <xsd:element name="common_auxiliary_of_place_relation"
          type="udc:common_auxiliary_of_place_relation"/>
        <xsd:element name="common_auxiliary_of_place_addition"
          type="udc:common_auxiliary_of_place_addition"/>
      </xsd:choice>
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>
```

Az alosztáson belül a legegyszerűbb esetben egyetlen táblázati szám áll.

Egy konkrét földrajzi területen élő adott népcsoportot a közvetlenül a földrajzi helyet kifejező számhoz kapcsolódó etnikai alosztási számmal kell kifejezni. Ezért opcionálisan egy etnikai alosztási szám, illetve több, pluszjellel összekötött etnikai alosztási szám megadását is lehetővé kell tenni az alosztás valamennyi elemének visszakereshetővé tétele érdekében.

```
<xsd:complexType name="common_auxiliary_of_place_number">
  <xsd:complexContent>
    <xsd:extension base="udc:auxiliary_container_abstract">
      <xsd:choice minOccurs="0">
        <xsd:element name="common_auxiliary_of_ethnics_number"
          type="udc:common_auxiliary_of_ethnics_number"/>
        <xsd:element name="common_auxiliary_of_ethnics_addition"
          type="udc:common_auxiliary_of_ethnics_addition"/>
      </xsd:choice>
      <xsd:attribute name="number1"
        type="udc:common_auxiliary_of_place_number_string" use="required"/>
      <xsd:attribute name="number2"
        type="udc:common_auxiliary_of_place_number_string" use="optional"/>
      <xsd:attribute name="order" type="xsd:integer" use="optional"/>
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>
```

A földrajzi alosztások táblázati számait általánosságban az alábbi egyszerű típus írja le:

```
<xsd:simpleType name="common_auxiliary_of_place_number_string">
  <xsd:restriction base="xsd:string">
    <xsd:pattern value="\ ([1-9] (\d| (\d{2}) (\.[1-9] \d{2})) * (\.[1-9] \d{0,2})) ?\) "/>
  </xsd:restriction>
</xsd:simpleType>
```

⁴⁷⁸ Emellett a nemzeti kiadások által tartalmazott bővítések, az ETO alapú egyéb osztályozások által tartalmazott módosítások (pl. Gilbert-Lane 2008), illetve a politikai változások is gyakran érintik ezeket az alosztásokat.

```
</xsd:simpleType>
```

Az egyes földrajzi egységek közötti kapcsolatok leírására a szabályok szerint a - zárójelen belül szereplő – kettőspontos viszonyítást kell alkalmazni, amely az alábbi módon írható le:

```
<xsd:complexType name="common_auxiliary_of_place_relation">
  <xsd:sequence>
    <xsd:choice minOccurs="2" maxOccurs="unbounded">
      <xsd:element name="common_auxiliary_of_place_number"
        type="udc:common_auxiliary_of_place_number"/>
      <xsd:element name="common_auxiliary_of_geography_addition"
        type="udc:common_auxiliary_of_geography_addition"/>
    </xsd:choice>
  </xsd:sequence>
</xsd:complexType>
```

Legvégül a **pluszjeles összekötés**re kell felkészülnünk. A földrajzi alosztások ebből a szempontból egyedi esetet képeznek, itt ugyanis kétféle összekötésről beszélhetünk. A *szintagmatikus összekötés* a táblázati számok eseti összekapcsolása, mely jelentését tekintve megegyezik a földrajzi alosztások külön zárójelben történő egymás után írásával. Ilyen esetben a számok egy alosztáson belüli alkalmazását csak és kizárólag a jelzet szerkezetének reprodukálhatósága indokolja, információ visszakeresési szempontból a jelzet külön alosztásokra tagolása is kielégítő eredményt nyújtana.

A másik, *paradigmatikus összekötés*re olyan, az irodalomban gyakran előforduló földrajzi helyek leírásakor van szükség, melyek nem fejezhetők ki egyetlen táblázati számmal. Mivel a földrajzi alosztások táblázatainak továbbosztásai a politikai földrajzi kontinenshatárokat, illetve az aktuális országhatárokat követik, ilyen számokra akkor lehet szükség, ha egy ország több kontinensre is kiterjed, vagy valamilyen más – általában történelmi, illetve politikai – okból nem illik bele ebbe a hierarchiába. Az ilyen számok egy részét maguk a táblázatok is tartalmazzák (pl. Oroszország, Szovjetunió), egy másik részüket pedig a használt authority fájlok vagy egy adott könyvtár vagy ország osztályozási hagyományai határozzák meg.

A szintagmatikus összekötés leírására a következő típus használható:

```
<xsd:complexType name="common_auxiliary_of_place_addition">
  <xsd:choice minOccurs="2" maxOccurs="unbounded">
    <xsd:element name="common_auxiliary_of_place_number"
      type="udc:common_auxiliary_of_place_number"/>
    <xsd:element name="common_auxiliary_of_place_relation"
      type="udc:common_auxiliary_of_place_relation"/>
  </xsd:choice>
</xsd:complexType>
```

A földrajzi táblázati számok *paradigmatikus összekötése* azért különleges, mert miközben az így képzett jelzet *az összes komponense alapján visszakeresendő* (és azok kapcsolatát is jelölni kell), minden egyéb szempontból *úgy viselkedik, mint egy egyszerű táblázati szám*. Ezért leszámaztatni is a táblázati számok őstípusából célszerű.

```
<xsd:complexType name="common_auxiliary_of_geography_addition">
  <xsd:complexContent>
    <xsd:extension base="udc:auxiliary_container_abstract">
      <xsd:sequence>
        <xsd:element name="common_auxiliary_of_place_number"
          type="udc:common_auxiliary_of_geography_number"
          minOccurs="2" maxOccurs="unbounded"/>
        <xsd:choice minOccurs="0">
          <xsd:element name="common_auxiliary_of_ethnics_number"
            type="udc:common_auxiliary_of_ethnics_number"/>
        </xsd:choice>
      </xsd:sequence>
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>
```

```

    <xsd:element name="common_auxiliary_of_ethnics_addition"
      type="udc:common_auxiliary_of_ethnics_addition"/>
  </xsd:choice>
  <xsd:element name="special_auxiliary"
    type="udc:special_auxiliary"
    minOccurs="0" maxOccurs="unbounded"/>
  <xsd:element name="nonudc_notation"
    type="udc:nonudc_notation"
    minOccurs="0" maxOccurs="1"/>
  <xsd:element name="alphabetical_specification"
    type="udc:alphabetical_specification"
    minOccurs="0" maxOccurs="unbounded"/>
</xsd:sequence>
</xsd:extension>
</xsd:complexContent>
</xsd:complexType>

```

A paradigmatiszus összekötés *operandusai* számára is célszerű önálló típust létrehozni, amely nem tartalmazhat speciális vagy külső jelzetből származó alosztásokat (amennyiben ilyen előfordul, annak a teljes összekötéshez kell kapcsolódnia).

```

<xsd:complexType name="common_auxiliary_of_geography_number">
  <xsd:complexContent>
    <xsd:restriction base="udc:common_auxiliary_of_place_number">
      <xsd:sequence>
        <xsd:element name="special_auxiliary"
          type="udc:special_auxiliary"
          minOccurs="0" maxOccurs="0"/>
        <xsd:element name="alphabetical_specification"
          type="udc:alphabetical_specification"
          minOccurs="0" maxOccurs="0"/>
        <xsd:element name="nonudc_notation"
          type="udc:nonudc_notation"
          minOccurs="0" maxOccurs="0"/>
      </xsd:sequence>
      <xsd:attribute name="number1"
        type="udc:common_auxiliary_of_geography_number_string"
        use="required"/>
      <xsd:attribute name="number2"
        type="udc:common_auxiliary_of_geography_number_string"
        use="prohibited"/>
      <xsd:attribute name="order"
        type="xsd:integer"
        use="required"/>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>

```

Még szükséges definiálnunk a földrajzi táblázati számok azon részhalmazát, amely a **politikai földrajzi helyek** leírását tartalmazza. *Paradigmatikus összekötés csak ezek között a számok között lehetséges* (a számformátumot a táblázatokban elfoglalt hely határozza meg).

```

<xsd:simpleType name="common_auxiliary_of_geography_number_string">
  <xsd:restriction base="udc:common_auxiliary_of_place_number_string">
    <xsd:pattern value="\ ([3-9] (\d| (\d{2}) (\.[1-9]\d{2}) * (\.[1-9]\d{0,2}) ?) ) ? \) " />
  </xsd:restriction>
</xsd:simpleType>

```

Végül két speciális jelzetalkotási lehetőségről kell szólnunk. A (161/164) számok a **földi helyek** és területek négy kvadráns szerinti, *fokbeosztás segítségével történő meghatározására* szolgálnak. Ebben az esetben legegyszerűbb számjelzetes korlátozottan közös alosztásokat használni a pontos

koordináták meghatározásához. Hasonlóan járhatunk el a **térbeli méretek** kifejezéséhez a (18) alatt. Ebben a speciális esetben névalosztással határozhatjuk meg a hosszúság, terület, illetve térfogat *mértékegységét*, melyben az érték ki van fejezve.⁴⁷⁹

5.5.7.1.4 Az emberek származása, etnikai csoportok és nemzetiség szerinti (etnikai) általánosan közös alosztások

A népi általánosan közös alosztásokat kerek zárójel határolja, a táblázati számokat pedig egyenlőségjel vezeti be. A zárójelen *belül egyszerű, komplex vagy agglomerált tárgykörök is állhatnak*. Ezekon kívül még a *kettőspontos viszonyítás egy speciális formájára* kell felkészülnünk.

A népi alosztásokat leíró összetett típus a következő:

```
<xsd:complexType name="common_auxiliary_of_ethnics">
  <xsd:complexContent>
    <xsd:extension base="udc:common_auxiliary_independent">
      <xsd:sequence>
        <xsd:choice>
          <xsd:element name="common_auxiliary_of_ethnics_number"
            type="udc:common_auxiliary_of_ethnics_number"/>
          <xsd:element name="common_auxiliary_of_ethnics_addition"
            type="udc:common_auxiliary_of_ethnics_addition"/>
          <xsd:element name="common_auxiliary_of_ethnics_relation"
            type="udc:common_auxiliary_of_ethnics_relation"/>
        </xsd:choice>
      </xsd:sequence>
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>
```

Az **összekötés** a fő táblázati számokhoz hasonló módon fejezhető ki:

```
<xsd:complexType name="common_auxiliary_of_ethnics_addition">
  <xsd:choice minOccurs="2" maxOccurs="unbounded">
    <xsd:element name="common_auxiliary_of_ethnics_number"
      type="udc:common_auxiliary_of_ethnics_number"/>
    <xsd:element name="common_auxiliary_of_ethnics_relation"
      type="udc:common_auxiliary_of_ethnics_relation"/>
  </xsd:choice>
</xsd:complexType>
```

A táblázati számokat az alábbi összetett típus írja le:

```
<xsd:complexType name="common_auxiliary_of_ethnics_number">
  <xsd:complexContent>
    <xsd:restriction base="udc:common_auxiliary_number_abstract">
      <xsd:sequence>
        <xsd:element name="special_auxiliary"
          type="udc:special_auxiliary"
          minOccurs="0" maxOccurs="unbounded"/>
        <xsd:element name="nonudc_notation"
          type="udc:nonudc_notation"
          minOccurs="0"/>
        <xsd:element name="alphabetical_specification"
```

⁴⁷⁹ OSZK KI 2005: 37-38, BSI 2005: 26. Az utóbbi revíziók ezeket a speciális eseteket is érintették, a jelenlegi szabályok szerint ugyanezek a helyeken, de külső forrásból származó alosztásokkal kell képezni az ilyen jelzeteket („UDC Online”).

```

        type="udc:alphabetical_specification"
        minOccurs="0"/>
    </xsd:sequence>
    <xsd:attribute name="number1"
        type="udc:common_auxiliary_of_ethnics_number_string"
        use="required"/>
    <xsd:attribute name="number2"
        type="udc:common_auxiliary_of_ethnics_number_string"
        use="optional"/>
</xsd:restriction>
</xsd:complexContent>
</xsd:complexType>

```

A pontos számformátumot pedig az alábbi egyszerű típus írja le:

```

<xsd:simpleType name="common_auxiliary_of_ethnics_number_string">
    <xsd:restriction base="xsd:string">
        <xsd:pattern value="\=(1.\d|\d)(\d{0,1}|(\d{2}(\.[1-9]\d{2}))*(\.[1-9]\d{0,2})?)\)" />
    </xsd:restriction>
</xsd:simpleType>

```

A politikai nemzethez egy földrajzi területen élő **közösséghez tartozás** a (=1) etnikai alosztási szám és a megfelelő földrajzi alosztás kettősponttal ⁴⁸⁰ (a régebbi kiadásokban ponttal ⁴⁸¹) való összekapcsolásával fejezhető ki⁴⁸². Ez a megoldás az alábbi összetett típussal írható le⁴⁸³:

```

<xsd:complexType name="common_auxiliary_of_ethnics_relation">
    <xsd:sequence>
        <xsd:element name="common_auxiliary_of_ethnics_number"
            type="udc:common_auxiliary_of_ethnics_number_in_relation"/>
        <xsd:element name="common_auxiliary_of_place"
            type="udc:common_auxiliary_of_place"/>
    </xsd:sequence>
    <xsd:attribute name="punctuation_mark" use="required">
        <xsd:simpleType>
            <xsd:restriction base="xsd:string">
                <xsd:length value="1"/>
                <xsd:pattern value="(:|\.)" />
            </xsd:restriction>
        </xsd:simpleType>
    </xsd:attribute>
</xsd:complexType>

```

A validáció miatt a *viszonyításban szereplő etnikai alosztási szám* leírására is célszerű saját típusokat definiálni:

```

<xsd:complexType name="common_auxiliary_of_ethnics_number_in_relation">
    <xsd:complexContent>
        <xsd:restriction base="udc:common_auxiliary_of_ethnics_number">
            <xsd:sequence>
                <xsd:element name="special_auxiliary"
                    type="udc:special_auxiliary"
                    minOccurs="0" maxOccurs="0"/>
            </xsd:sequence>
        </xsd:restriction>
    </xsd:complexContent>
</xsd:complexType>

```

⁴⁸⁰ BSI 2005: 130 és „UDC Online” (2013, a (=...) alatt) és <http://udcdata.info/011755> („UDC Summary”).

⁴⁸¹ OMIKK és OSZK-KMK 1990: 38 és OSZK KI 2005: 1/97.

⁴⁸² Például a (=1:439) jelzet jelöli Magyarország lakosait, a (=1:6=013) pedig Afrika fekete bőrű lakosságát

⁴⁸³ A két szám közötti központosítási jelnek információkeresési szempontból nincs jelentősége, de az eredeti jelzet reprodukálhatósága érdekében célszerű ezt is elmentenünk.

```

<xsd:element name="alphabetical_specification"
  type="udc:alphabetical_specification"
  minOccurs="0" maxOccurs="0"/>
<xsd:element name="nonudc_notation"
  type="udc:nonudc_notation"
  minOccurs="0" maxOccurs="0"/>
</xsd:sequence>
<xsd:attribute name="number1"
  type="udc:common_auxiliary_of_ethnics_number_in_relation_string"
  use="required"/>
<xsd:attribute name="number2"
  type="udc:common_auxiliary_of_ethnics_number_string"
  use="prohibited"/>
</xsd:restriction>
</xsd:complexContent>
</xsd:complexType>

<xsd:simpleType name="common_auxiliary_of_ethnics_number_in_relation_string">
  <xsd:restriction base="udc:common_auxiliary_of_ethnics_number_string">
    <xsd:pattern value="\ (=1\)" />
  </xsd:restriction>
</xsd:simpleType>

```

5.5.7.1.5 Idő szerinti általánosan közös alosztások

Az idő szerinti általánosan közös alosztások *sok szempontból eltérnek a többi alosztási táblázattól*. A földrajzi alosztásokhoz hasonlóan, melyek valamennyi fizikai helyhez köthető jellemző leírására igyekeznek lehetőséget nyújtani, ezek az alosztások is *úgy lettek összeállítva, hogy bármilyen, az idővel összefüggő aspektust ki lehessen fejezni velük*. Ebből következően, akárcsak a földrajzi alosztások, az idő szerinti alosztások táblázatai is nagyon összetettek.

Az idő alosztásoknak **két fő fajtája** van.⁴⁸⁴

- *kronológiai, időszámítás szerinti* idő alosztások, melyek a szokásos időszámítások és időegységek szerinti osztályozást teszik lehetővé
- *idő-fenomenológiai* idő alosztások, melyek a konkrét időszámítástól függetlenül lefolyó *általános időjelenségeket* jelölnek.

A **kronológiai idő** alosztások nem enumerikusan vannak felsorolva, csak a dátumok számára fenntartott helyek és a *dátumok leírási módjára vonatkozó szabályok* vannak meghatározva. Például a „0/2” osztályok a Gregorián naptár szerinti dátumok leírására vannak lefoglalva: „0001” jelöli a Krisztus utáni első évet, „001” az első évtizedet, „01” az első évszázadot, „1” pedig az első évezredet.

Az éven belüli pontos dátumok, illetve *időpontok* szintén a „0/2” alatt, „éééé.hh.nn.óó.pp.ss” formában fejezhetők ki.⁴⁸⁵

A fentiekből következik, hogy *a három számjegyenkénti pont szabálya a kronológiai idő alosztások esetén nem érvényes*.⁴⁸⁶

⁴⁸⁴ Babiczky 1974: 34.

⁴⁸⁵ Például "1898.12.07.15.46.03" jelöli 1898. december 7. 15 óra 46 perc 3 másodpercet mint időpontot.

⁴⁸⁶ A korábbi kiadásokban az idő-fenomenológiai alosztásokban is előfordultak olyan esetek, amikor a harmadik szám után pont nélkül áll a következő továbbosztás, például a "344" vagy a "366" alatt. A 2003. évi revízió során ez mindenütt korrigálva lett, és valamennyi ilyen esetben a helyére került a pont („Revised UDC Tables 2003”).

Dátumok esetén lehetőség van a *mínusz-* és *pluszjel* használatára annak kifejezésére, hogy időszámításunk előtti vagy utáni dátumról van szó. Ezek a jelek önmagukban is használhatók az időszámításunk előtti és utáni időszak kifejezésére.

A felsoroltokon kívül még egy speciális karaktersor említendő meg: *ha egy intervallum kezdete vagy vége nem ismert*, azt három ponttal fejezhetjük ki (pl. "19/..." jelenti a XIX. század utáni időszakot).

Az előzőek szerint az **idő alosztás számainak leírására** az alábbi *egyszerű típust* definiálhatjuk:

```
<xsd:simpleType name="common_auxiliary_of_time_number_string">
  <xsd:restriction base="xsd:string">
    <xsd:pattern value="'\.\.\.'"/>
    <xsd:pattern value="'(-|\+)?[0-2]\d{0,3}'"/>
  </xsd:pattern>
    <xsd:pattern value="'(-|\+)?([0-2]\d{3}(\.\d{2}(\.\d{2}(\.\d{2}(\.\d{2}(\.\d{2})?)?)?)?)?)'" />
  </xsd:pattern>
    <xsd:pattern value="'[3-9]((\.)?\d{1,4})(\.\d{1,4}) *'" />
  </xsd:restriction>
</xsd:simpleType>
```

Az idő alosztások táblázatai *korlátozottan közös alosztást nem tartalmaznak*. Azonban vannak olyan esetek, melyeket a legpraktikusabban számjelzetes alosztás használatával írhatunk le.⁴⁸⁷

A fentieket összegezve, az **idő alosztásban szereplő számokat** a következő *komplex típus* írja le:

```
<xsd:complexType name="common_auxiliary_of_time_number" abstract="false">
  <xsd:complexContent>
    <xsd:restriction base="udc:common_auxiliary_number_abstract">
      <xsd:sequence>
        <xsd:element name="special_auxiliary"
          type="udc:special_auxiliary"
          minOccurs="0" maxOccurs="1"/>
        <xsd:element name="nonudc_notation"
          type="udc:nonudc_notation"
          minOccurs="0"/>
        <xsd:element name="alphabetical_specification"
          type="udc:alphabetical_specification"
          minOccurs="0"/>
      </xsd:sequence>
      <xsd:attribute name="number1"
        type="udc:common_auxiliary_of_time_number_string"
        use="required"/>
      <xsd:attribute name="number2"
        type="udc:common_auxiliary_of_time_number_string"
        use="optional"/>
    </xsd:restriction>
  </xsd:complexContent>
```

⁴⁸⁷ Például ilyen megoldások voltak az évfelosztások ("3") és az időtartamok ("4") alatt (Egyetemes Tizedes Osztályozás 1990 pp. 39-40, illetve 2005, 1/1. pp. 98. és pp. 100.). A már említett 2003. évi revízió (Revised UDC Tables, 2003) ezeket a megoldásokat is érintette: ilyen esetekben most már az intervallum hosszát külső forrásból származó jelzetként, azaz csillaggal bevezetve kell jelölni (BSI 2005. pp. 133-134., ill. UDC Online 2013), ami nagymértékben megkönnyíti az ilyen jelzetek felbontását és értelmezését. Ez a módosítás egyben azt is jelenti, hogy az idő alosztásokban kiemelt jelentőséget kaptak a külső forrásból származó jelzetek.

```
</xsd:complexType>
```

A **periodicitást, időszakosságot** a "5" alatti osztályokkal fejezhetjük ki. Ezen belül a "52/55" számok használatával, a "3" alatti számokkal való kettőspontos viszonyítással fejezhetjük ki a nap, hét, hónap vagy év megadott időpontjaiban történő ismétlődést.⁴⁸⁸ Ezt az alábbi összetett és egyszerű típusokkal írhatjuk le:

```
<xsd:complexType name="common_auxiliary_of_time_relation">
  <xsd:sequence>
    <xsd:element name="common_auxiliary_of_time_periods_number"
      type="udc:common_auxiliary_of_time_periods_number"
      minOccurs="1" maxOccurs="1"/>
    <xsd:element name="common_auxiliary_of_time_dimensions_number"
      type="udc:common_auxiliary_of_time_dimensions_number"
      minOccurs="1" maxOccurs="1"/>
  </xsd:sequence>
</xsd:complexType>
```

```
<xsd:complexType name="common_auxiliary_of_time_dimensions_number">
  <xsd:complexContent>
    <xsd:restriction base="udc:common_auxiliary_of_time_number">
      <xsd:attribute name="number1"
        type="udc:common_auxiliary_of_time_dimensions_number_string"
        use="required"/>
      <xsd:attribute name="number2"
        type="udc:common_auxiliary_of_time_dimensions_number_string"
        use="optional"/>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>
```

```
<xsd:complexType name="common_auxiliary_of_time_periods_number">
  <xsd:complexContent>
    <xsd:restriction base="udc:common_auxiliary_of_time_number">
      <xsd:sequence>
        <xsd:element name="special_auxiliary"
          type="udc:special_auxiliary"
          minOccurs="0" maxOccurs="0"/>
        <xsd:element name="nonudc_notation"
          type="udc:nonudc_notation"
          minOccurs="0" maxOccurs="0"/>
        <xsd:element name="alphabetical_specification"
          type="udc:alphabetical_specification"
          minOccurs="0" maxOccurs="0"/>
      </xsd:sequence>
      <xsd:attribute name="number1"
        type="udc:common_auxiliary_of_time_periods_number_string"
        use="required"/>
      <xsd:attribute name="number2"
        type="udc:common_auxiliary_of_time_periods_number_string"
        use="prohibited"/>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>
```

```
<xsd:simpleType name="common_auxiliary_of_time_dimensions_number_string">
  <xsd:restriction base="udc:common_auxiliary_of_time_number_string">
```

⁴⁸⁸ Például "52:345" jelöli az éjszakai órákban, "53:329" a hét egy adott napján, "55:327" az év egy adott hónapjában történő ismétlődést.


```

    <xsd:pattern value="'3(\.)?\d{1,4})(\.\d{1,4})*'" />
  </xsd:restriction>
</xsd:simpleType>

<xsd:simpleType name="common_auxiliary_of_time_periods_number_string">
  <xsd:restriction base="udc:common_auxiliary_of_time_number_string">
    <xsd:length value="4" />
    <xsd:pattern value="'5[2-5]'" />
  </xsd:restriction>
</xsd:simpleType>

```

Akárcsak a többi alosztás esetén, az idő alosztásokban is definiálhatunk egy összetett típust több szám egy alosztáson belül, **pluszjellel történő összekötésére**.

```

<xsd:complexType name="common_auxiliary_of_time_addition">
  <xsd:choice minOccurs="2" maxOccurs="unbounded">
    <xsd:element name="common_auxiliary_of_time_number"
type="udc:common_auxiliary_of_time_number"/>
    <xsd:element name="common_auxiliary_of_time_relation"
type="udc:common_auxiliary_of_time_relation"/>
  </xsd:choice>
</xsd:complexType>

```

Ezzel meghatároztuk az idő alosztás lehetséges részeit. Magát az **alosztást** a következő **összetett típus** írja le:

```

<xsd:complexType name="common_auxiliary_of_time">
  <xsd:complexContent>
    <xsd:extension base="udc:common_auxiliary_independent">
      <xsd:choice>
        <xsd:element name="common_auxiliary_of_time_number"
type="udc:common_auxiliary_of_time_number"/>
        <xsd:element name="common_auxiliary_of_time_relation"
type="udc:common_auxiliary_of_time_relation"/>
        <xsd:element name="common_auxiliary_of_time_addition"
type="udc:common_auxiliary_of_time_addition"/>
      </xsd:choice>
      <xsd:sequence>
        <xsd:element name="common_auxiliary_of_time_number_1"
type="udc:common_auxiliary_of_time_number"/>
        <xsd:choice>
          <xsd:element name="common_auxiliary_of_place"
type="udc:common_auxiliary_of_place"/>
          <xsd:element name="common_auxiliary_of_ethnics"
type="udc:common_auxiliary_of_ethnics"/>
        </xsd:choice>
        <xsd:choice>
          <xsd:element name="common_auxiliary_of_time_number"
type="udc:common_auxiliary_of_time_number"/>
          <xsd:element name="nonudc_notation"
type="udc:nonudc_notation"/>
        </xsd:choice>
      </xsd:sequence>
    </xsd:choice>
  </xsd:extension>
</xsd:complexContent>
</xsd:complexType>

```

A hely, illetve népi általánosan közös alosztások szerepeltetésére a **sajátos időszámítások** kifejezése során van szükség: ilyen esetekben a "68" számot egy földrajzi vagy etnikai alosztás, majd egy kronológiai idő alosztási szám követi, itt azonban figyelemmel kell lennünk a séma változásaira is,

mivel a legújabb kiadásokban ilyen esetekben a konkrét dátumot már külső forrásból származó alosztással kell jelölni.⁴⁸⁹

5.5.7.1.6 Külső forrásból vett jelzetek és A/Z névalosztások

A külső forrásból származó alosztást leíró *komplex és egyszerű típusok* a következők:

```
<xsd:complexType name="nonudc_notation">
  <xsd:attribute name="code1" type="udc:nonudc_notation_string" use="required">
  </xsd:attribute>
  <xsd:attribute name="code2" type="udc:nonudc_notation_string" use="optional">
  </xsd:attribute>
  <xsd:attribute name="indexing_language" type="xsd:string">
  </xsd:attribute>
  <xsd:attribute name="order" type="xsd:int" use="required"/>
</xsd:complexType>

<xsd:simpleType name="nonudc_notation_string">
  <xsd:restriction base="xsd:string">
    <xsd:minLength value="1"/>
    <xsd:pattern value="\*.\+"/>
    <xsd:pattern value="#.\+"/>
  </xsd:restriction>
</xsd:simpleType>
```

Látható, hogy a jelzet formátumára vonatkozó *egyetlen megkötés a bevezető karakterre vonatkozik*. Arra is figyelemmel kell lennünk, hogy mivel a külső forrásból származó alosztás valamilyen más osztályozási vagy egyéb kódrendszerből származik, tartalmazhat **intervallumot** is, és az egyértelműség érdekében lehetőséget kell adni a forrásul szolgáló **kódrendszer** („indexing_language”) attribútumként történő megadására is.

Az alábbi típusokat a **betűjeles névalosztások** leírására használhatjuk:

```
<xsd:complexType name="alphabetical_specification">
  <xsd:attribute name="text" type="udc:alphabetical_specification_string"
  use="required"/>
  <xsd:attribute name="standard" type="xsd:string">
  </xsd:attribute>
  <xsd:attribute name="order" type="xsd:int" use="required"/>
</xsd:complexType>

<xsd:simpleType name="alphabetical_specification_string">
  <xsd:restriction base="xsd:string">
    <xsd:minLength value="1"/>
  </xsd:restriction>
</xsd:simpleType>
```

A névalosztás *intervallumot nem tartalmazhat*, a karakterekre viszont nem tehetünk érdemi megkötést, mivel valamennyi nyelven lehetővé kell tennünk a nevek leírását.⁴⁹⁰ Ezzel szemben célszerű lehetőséget adnunk a *szabvány* meghatározására, amely alapján a név leírásra került.

⁴⁸⁹ „UDC Online”.

⁴⁹⁰ Az ETO fordításokról készített 2004. évi felmérés (Slavić 2004) eredménye azt mutatja, hogy az (esetenként speciális karaktereket is tartalmazó) latin abécét használó nyelvek mellett különböző méretű táblázatok jelentek meg különböző cirill, arab, kínai, grúz, héber, japán (hiragana/katakana/kanji), koreai (hangeul/hanja) és vietnami

5.5.7.2 Nem önálló általánosan közös alosztások

A nem önálló általánosan közös alosztásokat *nem lehet önállóan alkalmazni, csak a fő táblázati jelzetek kiegészítőiként*. Ezek az alosztások tehát *nem állhatnak önmagukban, ahogy egy összetett jelzet első tagjaként sem*.

Két nagy csoportjuk van: a (már megszűnt) **szempont szerinti alosztások** és az **általános ismérvek szerinti alosztások**.

A nem önálló általánosan közös alosztásokat leíró típusokat az alábbi absztrakt **őstípus**ból származtatjuk:

```
<xsd:complexType name="common_auxiliary_dependent">
  <xsd:attribute name="order" type="xsd:int" use="required"/>
</xsd:complexType>
```

5.5.7.2.1 Szempont szerinti általánosan közös alosztások

A szempont szerinti általánosan közös alosztások⁴⁹¹ már kivonásra kerültek⁴⁹², de a régi jelzetek feldolgozhatósága miatt fel kell készülni a leírásukra. Ez az alábbi **komplex típus**sall történhet:

```
<xsd:complexType name="common_auxiliary_of_point_of_view">
  <xsd:complexContent>
    <xsd:extension base="udc:common_auxiliary_dependent">
      <xsd:sequence>
        <xsd:element name="common_auxiliary_of_point_of_view_number"
          type="udc:common_auxiliary_of_point_of_view_number"/>
      </xsd:sequence>
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>
```

A nyelvi alosztásokhoz hasonlóan, *a szempont szerinti alosztások között sincsenek definiálva relációk*; egy alosztás egyetlen **alosztási szám**ot vagy intervallumot tartalmazhat. Ezek típusa a következő:

```
<xsd:complexType name="common_auxiliary_of_point_of_view_number">
  <xsd:complexContent>
    <xsd:restriction base="udc:common_auxiliary_number_abstract">
      <xsd:sequence>
        <xsd:element name="special_auxiliary"
          type="udc:special_auxiliary"
          minOccurs="0" maxOccurs="0"/>
        <xsd:element name="nonudc_notation"
          type="udc:nonudc_notation"
          minOccurs="0"/>
        <xsd:element name="alphabetical_specification"
          type="udc:alphabetical_specification"
          minOccurs="0"/>
      </xsd:sequence>
      <xsd:attribute name="number1"
        type="udc:common_auxiliary_of_point_of_view_number_string"
        use="required"/>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>
```

(Quóe Ngű) betűket használó nyelveken is. A UDC Summary 2600 osztálya 57 nyelven érhető el, köztük olyanokon is, amelyek a fentiektől eltérő abécét használnak („UDC Summary Translations”). Ha még hozzátesszük, hogy az egyetlen nyelv, amelyen még továbbra is teljes nyelvű ETO kiadások jelennek meg is cirill betűket használ (Arszkij 2008), látható, hogy mennyire fontos a speciális karakterek támogatása.

⁴⁹¹ OMIKK és OSZK-KMK 1990: 44.

⁴⁹² „Cancellations to the UDC”.

```

    <xsd:attribute name="number2"
      type="udc:common_auxiliary_of_point_of_view_number_string"
      use="optional"/>
  </xsd:restriction>
</xsd:complexContent>
</xsd:complexType>

```

Legvégül ezúttal is a **szám formátumát** leíró *egyszerű típust* határozzuk meg:

```

<xsd:simpleType name="common_auxiliary_of_point_of_view_number_string">
  <xsd:restriction base="xsd:string">
    <xsd:minLength value="3"/>
    <xsd:pattern value="\.00\d{0,2}(\.[1-9]\d{2})*(\.[1-9]\d{0,2})?" />
  </xsd:restriction>
</xsd:simpleType>

```

5.5.7.2.2 Általános ismérvek szerinti általánosan közös alosztások

Ez az alosztás csoport a fogalmak olyan *általános ismérvek szerinti pontosítását* teszi lehetővé, mint a tulajdonságok, anyagok, a kapcsolatok és folyamatok vagy személyek.

Bár ezek az alosztások *nem tartalmaznak korlátozottan közös alosztásokat*⁴⁹³, az ide tartozó osztályokat egyedi jellemzőik és a rájuk jellemző *speciális jelzetszerkesztési módok* miatt külön-külön kell tárgyalnunk.

5.5.7.2.2.1 Tulajdonságok általánosan közös alosztásai

Az általános ismérvek szerinti alosztások első nagy csoportját a tulajdonságokat és attribútumokat leíró -02 karaktersorral kezdődő táblázati számok alkotják.⁴⁹⁴

Az ide tartozó számokat általánosan az alábbi **egyszerű és összetett típusokkal** írhatjuk le:

```

<xsd:complexType name="common_auxiliary_of_properties_number">
  <xsd:complexContent>
    <xsd:restriction base="udc:common_auxiliary_number_abstract">
      <xsd:sequence>
        <xsd:element name="special_auxiliary"
          type="udc:special_auxiliary_pointnought" minOccurs="0" maxOccurs="unbounded"/>
        <xsd:element name="nonudc_notation" type="udc:nonudc_notation" minOccurs="0"
          maxOccurs="1"/>
        <xsd:element name="alphabetical_specification"
          type="udc:alphabetical_specification" minOccurs="0" maxOccurs="1"/>
      </xsd:sequence>
      <xsd:attribute name="number1"
        type="udc:common_auxiliary_of_properties_number_string" use="required"/>
      <xsd:attribute name="number2"
        type="udc:common_auxiliary_of_properties_number_string" use="optional"/>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>

<xsd:simpleType name="common_auxiliary_of_properties_number_string">
  <xsd:restriction base="xsd:string">
    <xsd:pattern value="-02(\d{1}(\.[1-9]\d{2}))*(\.[1-9]\d{0,2})?" />
  </xsd:restriction>
</xsd:simpleType>

```

⁴⁹³ Ez következik abból is, hogy a jelzetelemek ajánlott sorrendje a korlátozottan közös alosztások elé sorolja őket, tehát a közvetlen utánuk következő korlátozottan közös alosztásoknak a módosított fő táblázati számra kell vonatkozniuk (OSZK KI 2005: 1/7), bár ez a -029 alatt kapcsolt fő táblázati számok esetében kétértelműséghez vezethet.

⁴⁹⁴ A -02-es táblázatok 1999-ben lettek bevezetve („Major Changes to the UDC 1993-2013”).

```

</xsd:restriction>
</xsd:simpleType>

```

A hagyományos korlátozottan közös alosztások mellett egy speciális jelzetszerkesztési megoldással új tulajdonságok is származtathatók a meglévő **fő táblázati számokból**, amennyiben azokat **kettősponttal a -029** jelzethez kapcsoljuk.

```

<xsd:complexType name="common_auxiliary_of_properties_relation">
  <xsd:sequence>
    <xsd:element name="common_auxiliary_of_properties_number"
      type="udc:common_auxiliary_of_properties_number_in_relation"/>
    <xsd:choice>
      <xsd:element name="main_table_number"
        type="udc:main_table_number_without_common_auxiliaries_and_subgrouping"/>
      <xsd:element name="main_table_synthesis"
        type="udc:main_table_synthesis_without_common_auxiliaries_and_subgrouping"/>
      <xsd:element name="main_table_subgrouping"
        type="udc:main_table_subgrouping_without_auxiliaries"/>
    </xsd:choice>
  </xsd:sequence>
</xsd:complexType>

```

A második paraméter egy *fő táblázati szám* vagy *aposztrófos jelzetszintézissel*, illetve *csoportképzéssel alkotott fogalom* lehet, amelyhez kapcsolódhat korlátozottan közös vagy külső forrásból származó alosztás is.

Az első paramétert az alábbi típusok írják le:

```

<xsd:complexType name="common_auxiliary_of_properties_number_in_relation">
  <xsd:complexContent>
    <xsd:restriction base="udc:common_auxiliary_number_abstract">
      <xsd:sequence>
        <xsd:element name="alphabetical_specification"
          type="udc:alphabetical_specification"
          minOccurs="0" maxOccurs="0"/>
        <xsd:element name="nonudc_notation"
          type="udc:nonudc_notation"
          minOccurs="0" maxOccurs="0"/>
        <xsd:element name="special_auxiliary"
          type="udc:special_auxiliary"
          minOccurs="0" maxOccurs="0"/>
      </xsd:sequence>
      <xsd:attribute name="number1"
        type="udc:common_auxiliary_of_properties_number_string"
        use="required"/>
      <xsd:attribute name="number2"
        type="udc:common_auxiliary_of_properties_number_string"
        use="optional"/>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>

<xsd:simpleType name="common_auxiliary_of_properties_number_in_relation_string">
  <xsd:restriction base="udc:common_auxiliary_of_properties_number_string">
    <xsd:pattern value="-029"/>
    <xsd:pattern value="-026.613"/>
  </xsd:restriction>
</xsd:simpleType>

```

A -026.613 karaktersorra azért kell felkészülni, mert a korábbi kiadások⁴⁹⁵ a színek leírását ennek a számnak és a megfelelő főtáblázati számnak (535-2) a kettőspontos összekapcsolásával tették lehetővé, ami a későbbi kiadásokban az ugyanazon főtáblázati korlátozottan közös alosztások szerinti – párhuzamos továbbosztássá változott.⁴⁹⁶

Miután az alosztás valamennyi lehetséges eleméhez definiálva vannak a megfelelő típusok, magát a **tulajdonságok alosztásait** a következőképpen írhatjuk le:

```
<xsd:complexType name="common_auxiliary_of_properties">
  <xsd:complexContent>
    <xsd:extension base="udc:common_auxiliary_dependent">
      <xsd:choice>
        <xsd:element name="common_auxiliary_of_properties_number"
          type="udc:common_auxiliary_of_properties_number"/>
        <xsd:element name="common_auxiliary_of_properties_relation"
          type="udc:common_auxiliary_of_properties_relation"/>
      </xsd:choice>
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>
```

5.5.7.2.2 Anyagok szerinti alosztások

A -03-al kezdődő táblázati számok az *anyagok mint tulajdonságok* leírására használhatók.

Ezek az alosztások sem tartalmaznak korlátozottan közös alosztásokat, ezért a **táblázati számok** leírhatók az alábbi típusokkal:

```
<xsd:complexType name="common_auxiliary_of_materials_number">
  <xsd:complexContent>
    <xsd:restriction base="udc:common_auxiliary_number_abstract">
      <xsd:sequence>
        <xsd:element name="special_auxiliary"
          type="udc:special_auxiliary"
          minOccurs="0" maxOccurs="0"/>
        <xsd:element name="nonudc_notation"
          type="udc:nonudc_notation"
          minOccurs="0" maxOccurs="1"/>
        <xsd:element name="alphabetical_specification"
          type="udc:alphabetical_specification"
          minOccurs="0" maxOccurs="1"/>
      </xsd:sequence>
      <xsd:attribute name="number1"
        type="udc:common_auxiliary_of_materials_number_string"
        use="required"/>
      <xsd:attribute name="number2"
        type="udc:common_auxiliary_of_materials_number_string"
        use="optional"/>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>
```

⁴⁹⁵ Ld. Egyetemes Tizedes Osztályozás 2005, 1/1. pp. 115.

⁴⁹⁶ BSI 2005: 143 és „UDC Online” (a -026.613 alatt). Érdekesség, hogy ez a fejlesztés ellentmond a UDCC törekvéseinek, amelyek az enumerikus táblázatok és a párhuzamos továbbosztások facettás átszervezését tűzték ki célul.

```
<xsd:simpleType name="common_auxiliary_of_materials_number_string">
  <xsd:restriction base="xsd:string">
    <xsd:pattern value="-03(\d{1}(\.[1-9]\d{2}))*(\.[1-9]\d{0,2})?" />
  </xsd:restriction>
</xsd:simpleType>
```

A kémiai vegyületeket leíró főtáblázati számokhoz hasonlóan az összetett anyagok, vegyületek stb. itt is **aposztrófós szintézissel** jelölhetők, ezt az alábbi *komplex típus* írja le:⁴⁹⁷

```
<xsd:complexType name="common_auxiliary_of_materials_synthesis">
  <xsd:sequence>
    <xsd:element name="common_auxiliary_of_materials_number"
      type="udc:common_auxiliary_of_materials_number"
      maxOccurs="unbounded"/>
  </xsd:sequence>
</xsd:complexType>
```

A – fenti elemeket tartalmazó – **alosztás leírása** pedig a következő *összetett típus*sal történhet:

```
<xsd:complexType name="common_auxiliary_of_materials">
  <xsd:complexContent>
    <xsd:extension base="udc:common_auxiliary_dependent">
      <xsd:choice>
        <xsd:element name="common_auxiliary_of_materials_number"
          type="udc:common_auxiliary_of_materials_number"/>
        <xsd:element name="common_auxiliary_of_materials_synthesis"
          type="udc:common_auxiliary_of_materials_synthesis"/>
      </xsd:choice>
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>
```

5.5.7.2.2.3 Kapcsolatok, folyamatok és műveletek közös alosztásai

Az általános ismérvek szerinti alosztások -04-gyel kezdődő altáblázatai a *fogalmak közötti kapcsolatok és az általános folyamatok, cselekvések, műveletek leírására* van elkülönítve.⁴⁹⁸

Ezek a táblázatok nem tartalmaznak korlátozottan közös alosztásokat, illetve egyéb speciális jelzetalkotási megoldást sem, ezért egyszerűen **leírhatók** az alábbi típusok használatával:

```
<xsd:complexType name="common_auxiliary_of_processes">
  <xsd:complexContent>
    <xsd:extension base="udc:common_auxiliary_dependent">
      <xsd:sequence>
        <xsd:element name="common_auxiliary_of_processes_number"
          type="udc:common_auxiliary_of_processes_number"/>
      </xsd:sequence>
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>

<xsd:complexType name="common_auxiliary_of_processes_number">
  <xsd:complexContent>
    <xsd:restriction base="udc:common_auxiliary_number_abstract">
```

⁴⁹⁷ A korábbi kiadásokban ez a megoldás szerepel (BSI 2005: 146, OMIKK és OSZK-KMK 1990: 47, ill. OSZK KI 2005: 1/119), a jelenlegi szabályok viszont már az alosztások egymás után írását javasolják, ami az analitikus-szintetikus szemlélet felől a – gépesítést megkönnyítő, de az ETO kifejezésbeli lehetőségeit csökkentő - teljes facettázás felé mutat („UDC Online”, a -03 alatt).

⁴⁹⁸ A -04-es alosztások 2003-ban kerültek bevezetésre („Major changes to the UDC 1993-2013”).

```

<xsd:sequence>
  <xsd:element name="special_auxiliary"
    type="udc:special_auxiliary"
    minOccurs="0" maxOccurs="0"/>
  <xsd:element name="nonudc_notation"
    type="udc:nonudc_notation"
    minOccurs="0" maxOccurs="1"/>
  <xsd:element name="alphabetical_specification"
    type="udc:alphabetical_specification"
    minOccurs="0" maxOccurs="1"/>
</xsd:sequence>
<xsd:attribute name="number1"
  type="udc:common_auxiliary_of_processes_number_string"
  use="required"/>
<xsd:attribute name="number2"
  type="udc:common_auxiliary_of_processes_number_string"
  use="optional"/>
</xsd:restriction>
</xsd:complexContent>
</xsd:complexType>

```

A számok formátumát az alábbi egyszerű típus írja le:

```

<xsd:simpleType name="common_auxiliary_of_processes_number_string">
  <xsd:restriction base="xsd:string">
    <xsd:pattern value="-04(\d{1}(\.[1-9]\d{2}))*(\.[1-9]\d{0,2})?" />
  </xsd:restriction>
</xsd:simpleType>

```

5.5.7.2.2.4 Személyek szerinti alosztások

A -05-tel kezdődő altáblázatok a *személyeket és az egyes személyes emberi jellemzőket* leíró alosztásokat tartalmazzák.

Akárcsak a kapcsolatok és folyamatok alosztásai, ezek a táblázatok sem tartalmazznak korlátozottan közös alosztásokat, ezért a leírásuk is hasonlóan történhet az alábbi összetett típusokkal:

```

<xsd:complexType name="common_auxiliary_of_persons">
  <xsd:complexContent>
    <xsd:extension base="udc:common_auxiliary_dependent">
      <xsd:sequence>
        <xsd:element name="common_auxiliary_of_persons_number"
          type="udc:common_auxiliary_of_persons_number"/>
      </xsd:sequence>
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>

<xsd:complexType name="common_auxiliary_of_persons_number">
  <xsd:complexContent>
    <xsd:restriction base="udc:common_auxiliary_number_abstract">
      <xsd:sequence>
        <xsd:element name="special_auxiliary"
          type="udc:special_auxiliary"
          minOccurs="0" maxOccurs="0"/>
        <xsd:element name="nonudc_notation"
          type="udc:nonudc_notation"
          minOccurs="0" maxOccurs="1"/>
        <xsd:element name="alphabetical_specification"
          type="udc:alphabetical_specification"
          minOccurs="0" maxOccurs="1"/>
      </xsd:sequence>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>

```



```

<xsd:attribute name="number1"
  type="udc:common_auxiliary_of_persons_number_string"
  use="required"/>
<xsd:attribute name="number2"
  type="udc:common_auxiliary_of_persons_number_string"
  use="optional"/>
</xsd:restriction>
</xsd:complexContent>
</xsd:complexType>

```

A számok formátumát leíró egyszerű típus is hasonló:

```

<xsd:simpleType name="common_auxiliary_of_persons_number_string">
  <xsd:restriction base="xsd:string">
    <xsd:pattern value="-05(\d{1}(\.[1-9]\d{2}))*(\.[1-9]\d{0,2})?" />
  </xsd:restriction>
</xsd:simpleType>

```

5.5.7.3 Jelzetek összekapcsolására szolgáló általánosan közös alosztási jelek

5.5.7.3.1 Összekötés

Az *agglomerációt* kifejező összekötés a **legmagasabb hierarchiaszintet** képviseli az ETO szintaktikai relációi között, mivel egyszerű, komplex és összetett tárgyköröket is összeköthet.

```

<xsd:complexType name="main_table_addition">
  <xsd:choice minOccurs="2" maxOccurs="unbounded">
    <xsd:element name="main_concept" type="udc:main_name"/>
    <xsd:element name="main_table_orderfixing"
      type="udc:main_table_orderfixing_ordered"/>
    <xsd:element name="main_table_relation"
      type="udc:main_table_relation_ordered"/>
  </xsd:choice>
</xsd:complexType>

```

Az összekötés **kommutatív**, azaz az elemek sorrendjének nincs jelentése, az tetszőlegesen változtatható, amennyiben a besorolási vagy más megfontolások⁴⁹⁹ azt indokolják. Az elemek sorrendjének elmentése elsősorban az eredeti jelzet helyreállítása miatt lehet hasznos.

Összekötés állhat **szögletes zárójelek között** is, ekkor azonban *nem tartalmazhat további csoportképzést*, ezért erre az esetre is fel kell készülnünk.

```

<xsd:complexType name="main_table_addition_without_subgrouping">
  <xsd:complexContent>
    <xsd:restriction base="udc:main_table_addition">
      <xsd:choice minOccurs="2" maxOccurs="unbounded">
        <xsd:element name="main_concept"
          type="udc:main_concept_without_subgrouping"/>
        <xsd:element name="main_table_orderfixing"
          type="udc:main_table_orderfixing_without_subgrouping"/>
        <xsd:element name="main_table_relation"
          type="udc:main_table_relation_without_subgrouping"/>
      </xsd:choice>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>

```

⁴⁹⁹ Ilyen lehet például a böngészés megkönnyítése, a keresési eredmények megfelelő formátumban történő megjelenítése, vagy a logikailag azonos jelzetek csoportosítása.

5.5.7.3.2 Viszonyítás

A viszonyítás közvetlen az agglomeráció után következik a prioritást tekintve. Ezért a **viszonyítást leíró komplex típus**nak opcionális elemként kell tartalmaznia sorrendrögzítést, illetve csoportképzést is a fő táblázati számok mellett.

```
<xsd:complexType name="main_table_relation">
  <xsd:choice minOccurs="2" maxOccurs="unbounded">
    <xsd:element name="main_concept" type="udc:main_concept"/>
    <xsd:element name="main_table_orderfixing"
      type="udc:main_table_orderfixing_ordered"/>
  </xsd:choice>
</xsd:complexType>

<xsd:complexType name="main_table_relation_ordered">
  <xsd:complexContent>
    <xsd:extension base="udc:main_table_relation">
      <xsd:attribute name="order" type="xsd:integer" use="required"/>
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>
```

A viszonyítás is **kommutatív**, a sorrend elmentését ugyanaz indokolja, mint az összekötés esetében.

Ebben az esetben is definiálnunk kell egy **külön típust a csoportképzés esetére**.

```
<xsd:complexType name="main_table_relation_without_subgrouping">
  <xsd:complexContent>
    <xsd:restriction base="udc:main_table_relation_ordered">
      <xsd:sequence>
        <xsd:choice minOccurs="2" maxOccurs="unbounded">
          <xsd:element name="main_concept"
            type="udc:main_concept_without_subgrouping"/>
          <xsd:element name="main_table_orderfixing"
            type="udc:main_table_orderfixing_without_subgrouping"/>
        </xsd:choice>
      </xsd:sequence>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>
```

5.5.7.3.3 Sorrendrögzítés

A **sorrendrögzítés a prioritási sorrendben eggyel a viszonyítás alatt van**, ezért a két típus közötti alapvető különbség, hogy az előbbi **nem tartalmazhat elemként sorrendrögzítést**.

```
<xsd:complexType name="main_table_orderfixing">
  <xsd:sequence>
    <xsd:element name="main_concept"
      type="udc:main_concept"
      minOccurs="2" maxOccurs="unbounded"/>
  </xsd:sequence>
</xsd:complexType>

<xsd:complexType name="main_table_orderfixing_ordered">
  <xsd:complexContent>
    <xsd:extension base="udc:main_table_orderfixing">
      <xsd:attribute name="order" type="xsd:integer" use="required"/>
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>
```

Az összekötéssel és a viszonyítással szemben az **elemek sorrendjének** itt *kiemelt jelentősége van*, hiszen a teljes jelzet jelentését és besorolását meghatározza.

A csoportképzésen belül álló sorrendrögzítést az alábbi típus írja le:

```
<xsd:complexType name="main_table_orderfixing_without_subgrouping">
  <xsd:complexContent>
    <xsd:restriction base="udc:main_table_orderfixing_ordered">
      <xsd:sequence>
        <xsd:element name="main_concept"
          type="udc:main_concept_without_subgrouping"
          minOccurs="2" maxOccurs="unbounded"/>
      </xsd:sequence>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>
```

5.5.7.3.4 Csoportképzés

A csoportképzésen belül tetszőleges reláció állhat, kivéve magát a csoportképzést. Ezért a csoportképzést a következő komplex típussal lehet leírni:

```
<xsd:complexType name="main_table_subgrouping">
  <xsd:complexContent>
    <xsd:extension base="udc:auxiliary_container_abstract">
      <xsd:sequence>
        <xsd:choice>
          <xsd:element name="main_concept"
            type="udc:main_concept_without_subgrouping"/>
          <xsd:element name="main_table_orderfixing"
            type="udc:main_table_orderfixing_without_subgrouping"/>
          <xsd:element name="main_table_relation"
            type="udc:main_table_relation_without_subgrouping"/>
          <xsd:element name="main_table_addition"
            type="udc:main_table_addition_without_subgrouping"/>
        </xsd:choice>
        <xsd:element name="special_auxiliary"
          type="udc:special_auxiliary"
          minOccurs="0" maxOccurs="unbounded"/>
        <xsd:element name="alphabetical_specification"
          type="udc:alphabetical_specification"
          minOccurs="0" maxOccurs="unbounded"/>
        <xsd:element name="nonudc_notation"
          type="udc:nonudc_notation"
          minOccurs="0"/>
        <xsd:element name="common_auxiliary_independent"
          type="udc:common_auxiliary_independent"
          minOccurs="0" maxOccurs="unbounded"/>
        <xsd:element name="common_auxiliary_dependent"
          type="udc:common_auxiliary_dependent"
          minOccurs="0" maxOccurs="unbounded"/>
      </xsd:sequence>
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>

<xsd:complexType name="main_table_subgrouping_ordered">
  <xsd:complexContent>
    <xsd:extension base="udc:main_table_subgrouping">
      <xsd:attribute name="order" type="xsd:integer" use="required"/>
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>
```

A csoportképzésen belüli számot leíró elem *nem tartalmazhat csoportképzést elemként*, az egymásba ágyazás elkerülése miatt. A további elemek a zárójelen belüli összetett számot egységesen módosító, tehát *a zárójelen kívül álló alosztások* leírására szolgálnak.

A csoportképzés egy **speciális esete**, amikor az így képzett komplex jelzet *egy alosztáson belül áll* (például kettősponttal bevezetve). Ilyenkor a csoportképzéshez nem kapcsolódhatnak további alosztások:

```
<xsd:complexType name="main_table_subgrouping_without_auxiliaries">
  <xsd:complexContent>
    <xsd:restriction base="udc:main_table_subgrouping">
      <xsd:sequence>
        <xsd:choice>
          <xsd:element name="main_concept"
            type="udc:main_concept_without_subgrouping"/>
          <xsd:element name="main_table_orderfixing"
            type="udc:main_table_orderfixing_without_subgrouping"/>
          <xsd:element name="main_table_relation"
            type="udc:main_table_relation_without_subgrouping"/>
          <xsd:element name="main_table_addition"
            type="udc:main_table_addition_without_subgrouping"/>
        </xsd:choice>
        <xsd:element name="special_auxiliary"
          type="udc:special_auxiliary"
          minOccurs="0" maxOccurs="0"/>
        <xsd:element name="alphabetical_specification"
          type="udc:alphabetical_specification"
          minOccurs="0" maxOccurs="unbounded"/>
        <xsd:element name="nonudc_notation"
          type="udc:nonudc_notation"
          minOccurs="0"/>
        <xsd:element name="common_auxiliary_independent"
          type="udc:common_auxiliary_independent"
          minOccurs="0" maxOccurs="0"/>
        <xsd:element name="common_auxiliary_dependent"
          type="udc:common_auxiliary_dependent"
          minOccurs="0" maxOccurs="0"/>
      </xsd:sequence>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>
```

5.6 Példák az XML alkalmazására

Az első példa egy **illusztrált, orosz nyelvű könyv a szoftverkarbantartásról és az integrációs tesztek és a modularizáció kapcsolatáról online rendszerekben** ETO-jelzetének XML reprezentációját mutatja be.

Az tesztek és a modularizáció kapcsolattát egy viszonyítás (004.415.23:004.415.53) fejezi ki, melyet pluszjellel kötünk össze a karbantartás jelzetével (004.416). Ez az összekötés, mivel az alosztások egységként módosítják, szögletes zárójelek közé kerül, az így kialakított csoportképzéshez pedig egy korlátozottan közös alosztás (.031.4, online), és egy - szintén korlátozottan közös alosztással pontosított – formai és egy nyelvi alosztás kapcsolódik.

A teljes jelzet **XML reprezentációját** az alábbi példa mutatja be.

```
<ns:udc_concept
  xsi:schemaLocation="http://piros.udc-interpreter.hu/#xsd udc.xsd"
  xmlns:ns="http://piros.udc-interpreter.hu/#xsd"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" udc_edition="2017"
```

```

notation="004.416+004.415.23:004.415.53].031.4(02.025.2)=161.1">
<ns:description xml:lang="EN">
  Software maintenance. The relationship between modularization and integration
  testing (book in Russian, with illustrations).
</ns:description>
<ns:description xml:lang="HU">
  Szoftverkarbantartás. A modularizáció és az integrációs tesztek kapcsolata.
  (illusztrált könyv orosz nyelven).
</ns:description>
<ns:main_concept>
  <ns:main_table_subgrouping>
    <ns:main_table_addition>
      <ns:main_concept order="1">
        <ns:main_table_number number1="004.416"/>
      </ns:main_concept>
      <ns:main_table_relation order="2">
        <ns:main_concept order="1">
          <ns:main_table_number number1="004.415.23"/>
        </ns:main_concept>
        <ns:main_concept order="2">
          <ns:main_table_number number1="004.415.53"/>
        </ns:main_concept>
      </ns:main_table_relation>
    </ns:main_table_addition>
  </ns:main_table_subgrouping>
  <ns:special_auxiliary
    xsi:type="ns:special_auxiliary_pointnought" order="1">
    <ns:special_auxiliary_number
      xsi:type="ns:special_auxiliary_pointnought_number" number1=".031.4" />
    </ns:special_auxiliary>
  </ns:special_auxiliary>
  </ns:main_table_subgrouping>
  <ns:common_auxiliary_independent
    xsi:type="ns:common_auxiliary_of_form"
    order="1">
    <ns:common_auxiliary_of_form_number number1="(02)">
      <ns:special_auxiliary
        xsi:type="ns:special_auxiliary_pointnought"
        order="1">
      </ns:special_auxiliary>
      <ns:special_auxiliary_number
        xsi:type="ns:special_auxiliary_pointnought_number"
        number1=".025.2"/>
      </ns:special_auxiliary>
    </ns:common_auxiliary_of_form_number>
    </ns:common_auxiliary_independent>
  </ns:common_auxiliary_independent>
  <ns:common_auxiliary_independent
    xsi:type="ns:common_auxiliary_of_language"
    order="2">
    <ns:common_auxiliary_of_language_number number1="=161.1"/>
    </ns:common_auxiliary_independent>
  </ns:common_auxiliary_independent>
</ns:main_concept>
</ns:udc_concept>

```

A másik példa *Stuart Pearce 1990. évi labdarúgó világbajnokságon elhibázott büntetőrúgásának* a leírására készült:

796.332.093.46(100)"1990":796.332.093.434(430:410.1)"1990":796.332.012.574.45-027.233.4-051:796.332(410.1).056.22Pearce

A jelzet XML reprezentációja a következő:

```

<ns:udc_concept xsi:schemaLocation="http://piros.udc-interpreter.hu/#xsd udc.xsd"
  xmlns:ns="http://piros.udc-interpreter.hu/#xsd"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  udc_edition="2005"

```

notation="796.332.093.46(100)"1990";796.332.093.434(430:410.1)"1990\";796.332.012.574.45-027.233.4-051:796.332(410.1).056.22Pearce">

```

<ns:main_table_relation>
  <ns:main_concept>
    <ns:main_table_number number1="796.332">
      <ns:special_auxiliary
        xsi:type="ns:special_auxiliary_pointnought" order="1">
          <ns:special_auxiliary_number
            xsi:type="ns:special_auxiliary_pointnought_number" number1=".093.46" />
          </ns:special_auxiliary>
        </ns:main_table_number>
        <ns:common_auxiliary_independent
          xsi:type="ns:common_auxiliary_of_place" order="1">
            <ns:common_auxiliary_of_place_number number1="(100)" />
          </ns:common_auxiliary_independent>
          <ns:common_auxiliary_independent
            xsi:type="ns:common_auxiliary_of_time" order="2">
              <ns:common_auxiliary_of_time_number
                number1="'1990'" />
            </ns:common_auxiliary_independent>
          </ns:main_concept>
        <ns:main_concept>
          <ns:main_table_number number1="796.332">
            <ns:special_auxiliary
              xsi:type="ns:special_auxiliary_pointnought" order="1">
                <ns:special_auxiliary_number
                  xsi:type="ns:special_auxiliary_pointnought_number" number1=".093.434" />
                </ns:special_auxiliary>
              </ns:main_table_number>
              <ns:common_auxiliary_independent
                xsi:type="ns:common_auxiliary_of_place" order="1">
                  <ns:common_auxiliary_of_place_relation>
                    <ns:common_auxiliary_of_place_number number1="(430)" />
                    <ns:common_auxiliary_of_place_number number1="(410.1)" />
                  </ns:common_auxiliary_of_place_relation>
                </ns:common_auxiliary_independent>
                <ns:common_auxiliary_independent
                  xsi:type="ns:common_auxiliary_of_time" order="2">
                    <ns:common_auxiliary_of_time_number
                      number1="'1990'" />
                    </ns:common_auxiliary_independent>
                  </ns:main_concept>
                <ns:main_concept>
                  <ns:main_table_number number1="796.332">
                    <ns:special_auxiliary
                      xsi:type="ns:special_auxiliary_pointnought" order="1">
                        <ns:special_auxiliary_number
                          xsi:type="ns:special_auxiliary_pointnought_number"
                            number1=".012.574.45" />
                        </ns:special_auxiliary>
                      <ns:common_auxiliary_dependent
                        xsi:type="ns:common_auxiliary_of_properties" order="1">
                          <ns:common_auxiliary_of_properties_number number1="-027.233.4" />
                        </ns:common_auxiliary_dependent>
                      <ns:common_auxiliary_dependent
                        xsi:type="ns:common_auxiliary_of_persons" order="1">
                          <ns:common_auxiliary_of_persons_number number1="-051" />
                        </ns:common_auxiliary_dependent>
                      </ns:main_table_number>
                    </ns:main_concept>
                  <ns:main_concept>
                    <ns:main_table_number number1="796.332">
                      <ns:special_auxiliary

```

```

      xsi:type="ns:special_auxiliary_pointnought" order="1">
      <ns:special_auxiliary_number
      xsi:type="ns:special_auxiliary_pointnought_number" number1=".056.22" />
    </ns:special_auxiliary>
    <ns:alphabetical_specification order="2" text="Pearce"/>
  </ns:main_table_number>
  <ns:common_auxiliary_independent
  xsi:type="ns:common_auxiliary_of_place" order="1">
    <ns:common_auxiliary_of_place_number number1="(410.1)" />
  </ns:common_auxiliary_independent>
</ns:main_concept>
</ns:main_table_relation>
</ns:udc_concept>

```

8. ábra A 796.332.093.46(100)"1990":796.332.093.434(430:410.1)"1990":796.332.012.574.45-027.233.4-051:796.332(410.1).056.22Pearce jelzet XML reprezentációja

5.7 Irodalomjegyzék

- Arszkij, Jurij M. 2008. „UDC activities in Russia: A VINITI report.” *Extensions and Corrections to the UDC 30*: 49-52. <http://arizona.openrepository.com/arizona/handle/10150/105158>. (megtekintve: 2017-11-01)
- Babiczyk Béla. 1975. *Könyvtári feldolgozó munka: A tanítóképző intézet II. évfolyama számára (népművelő-könyvtáros szak). 2. rész. Osztályozás és szakkatalógus szerkesztés*. Budapest: Tankönyvkiadó.
- Babiczyk Béla. 1985. „Szemléletváltozás az ETO jelzetszerkesztésében.” *Könyvtári Figyelő* 31 (1): 17–27.
- Babiczyk Béla, és Barátné Hajdu Ágnes. 1998. *Bevezetés az információkereső nyelvek elméletébe és gyakorlatába*. Budapest: Universitas.
- British Standards Institution (BSI). 2005. *Universal Decimal Classification: standard edition: volume 1: systematic tables*. London: British Standards Institution.
- Bourret, Ronald. 2011. „XML Data Binding Resources.” Legutóbb módosítva 2011. október 7-én. <http://www.rpbourret.com/xml/XMLDataBinding.htm>. (megtekintve: 2017-11-01)
- „Cancellations to the UDC.” UDCC (UDC Consortium). <http://udcc.org/index.php/site/page?view=cancellations>. (megtekintve: 2017-11-01)
- „Creative Commons (CC) 4.0. Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0).” Creative Commons (CC) <https://creativecommons.org/licenses/by-nc-nd/4.0/>. (megtekintve: 2017-11-01)
- Eszenyiné Borbély Mária, és Salgáné Medveczki Marianna. 2010. *Komplex könyvtári feldolgozó munka gyakorlata*. Budapest: Kempelen Farkas Hallgatói Információs Központ. http://www.tankonyvtar.hu/hu/tartalom/tamop425/0046_komplex_konyvtari_feldolgozo_munka_gyakorlata/index.html. (megtekintve: 2017-11-01)
- „Extensible Markup Language (XML).” World Wide Web Consortium (W3C). Legutóbb módosítva 2016. október 11-én. <https://www.w3.org/XML/>. (megtekintve: 2017-11-01)
- „Extensions and Corrections to the UDC.” UDCC (UDC Consortium). <http://udcc.org/index.php/site/page?view=ec>. (megtekintve: 2017-11-01)

- Gilbert, Mark, és Heather Lane. 2008. „Forty-five numbers for snow: a brief introduction to the UDC for Polar libraries.” *Extensions and Corrections to the UDC* 30: 23-25.
<http://hdl.handle.net/10150/105332>. (megtekintve: 2017-11-01)
- Gnoli, Claudio. 2011. „Facets in UDC: a review of current situation.” *Extensions and Corrections to the UDC* 33: 19-36. <http://arizona.openrepository.com/arizona/handle/10150/236511>.
 (megtekintve: 2017-11-01)
- Hiroshi. 2007. „Advantages & Disadvantages Of XML.” <http://www.techmynd.com/advantages-disadvantages-of-xml/> (megtekintve: 2017-11-01)
- „Major Changes to the UDC 1993-2013.” UDCC (UDC Consortium).
http://www.udcc.org/index.php/site/page?view=major_revisions. (megtekintve: 2017-11-01)
- Mándy Gábor. 2013. „A posztkoordináció esélyei az ETO-ban.” *Könyvtári figyelő* 59: 65-84.
http://epa.oszk.hu/00100/00143/00086/pdf/EPA00143_konyvtari_figyelo_2013_1_065-083.pdf.
 (megtekintve: 2017-01-29)
- Murell, Paul. 2009. *Introduction to Data Technologies*. London: Chapman & Hall / CRC
<https://www.stat.auckland.ac.nz/~paul/ltDT/> (megtekintve: 2017-11-01)
- Országos Műszaki Információs Központ és Könyvtár (OMIKK), és Országos Széchényi Könyvtár Könyvtártudományi és Módszertani Központ (OSZK-KMK). 1990. *Egyetemes tizedes osztályozás. Rövidített kiadás. 1. kötet Táblázatok. FID Publication 691*. Budapest: OMIKK, OSZK-KMK.
- Országos Széchényi Könyvtár Könyvtári Intézet (OSZK KI). 2005. *Egyetemes tizedes osztályozás. 1. kötet Táblázatok 1-2. Rész. UDCC Publication P057*. Budapest: Országos Széchényi Könyvtár Könyvtári Intézet.
- Pálvölgyi Mihály. 1998. *Információfeldolgozás, információkereső nyelvek*. Szombathely: Berzsenyi Dániel Tanárképző Főiskola.
- Piros Attila. 2015a. „Automatic interpretation of complex UDC numbers: towards support for library systems.” In *Classification & authority control: expanding resource discovery: proceedings of the International UDC Seminar, 29-30 October 2015, Lisbon, Portugal*, szerk. Aida Slavić, és Maria Inês Cordeiro, 177-193. Würzburg: Ergon Verlag. „Revised UDC tables 2000”. *Extensions and Corrections to the UDC* 22.
- Piros Attila. 2015b. „A New Approach to Universal Decimal Classification as the Indexing Language for a Mechanized Information Retrieval System.” In *Agria Media 2014: XI. Információtechnikai és Oktatástechnológiai Konferencia és Kiállítás. Eszterházy Károly Főiskola - 2014. október 8-10.*, szerk. Dr. Nádasi András, 228-235. Eger: Eszterházy Károly Főiskola Médiainformatikai Intézet.
- Piros Attila. 2017. „The thought behind the symbol: about the automatic interpretation and representation of UDC numbers.” In *Faceted classification today: theory, technology and end users: proceedings of the International UDC Seminar 2017, London (UK), 14-15 September*, edited by Aida Slavić, and Claudio Gnoli, 203-218. Würzburg: Ergon Verlag.
- Piros Attila. 2018. „New automatic interpreter for complex UDC numbers.” *Extensions and Corrections to the UDC* 36-37 (2014-2015). [megjelenés alatt]
- „Revised UDC tables 2003.” *Extensions & Corrections to the UDC* 25.
- Ranganathan, S. R. 1967. *Prolegomena to library classification. 3rd ed.* New York: Asia Publishing House. <http://hdl.handle.net/10150/106370>. (megtekintve: 2017-11-01)

- Robinson, Geoffrey. 2002. „More about the meaning of common auxiliaries.” *Extensions and Corrections to the UDC* 24: 8. <http://www.udcc.org/notes.htm#common>. (megtekintve: 2017-11-01)
- Robinson, Geoffrey. 2003. „Citation Order in UDC.” *Extensions and Corrections to the UDC* 25: 19-27.
- Slavić, Aida. 2004. „UDC Translations: a 2004 Survey Report and Bibliography.” *Extensions and Corrections to the UDC* 26: 58-80.
- Slavić-Overfield, Aida. 2005. „Classification Management and use in a Networked Environment: The Case of the Universal Decimal Classification.” PhD diss., London: University College London. http://discovery.ucl.ac.uk/1334914/1/AidaSlavicOverfield_thesis_UCL2005.pdf. (megtekintve: 2017-11-01)
- Slavić, Aida, és Maria Inês Cordeiro. 2005. „Sharing and re-use of classification systems: the need for a common data model.” *Signum* 37 (8): 19-24. <http://arizona.openrepository.com/arizona/handle/10150/105132>. (megtekintve: 2017-11-01)
- Slavić, Aida, és Antoine Isaac. 2009. „Identifying management issues in networked KOS: examples from classification schemes.” Presentation at NKOS 2009 workshop, ECDL conference. Corfu, Greece. https://at-web1.comp.glam.ac.uk/pages/research/hypermedia/nkos/nkos2009/presentations/Slavić_isaac_NKOS2009_06.pdf. (megtekintve: 2017-11-01)
- „UDC Online.” UDCC (UDC Consortium). <http://www.udc-hub.com/>. (megtekintve: 2017-11-01)
- „UDC Summary. Multilingual Universal Decimal Classification Summary. UDCC Publication No. 088.” UDCC (UDC Consortium). <http://www.udcc.org/udcsummary/php/index.php>. (megtekintve: 2017-11-01)
- „UDC Summary Translations.” UDCC (UDC Consortium). <http://www.udcsummary.info/translation.htm>. (megtekintve: 2017-11-01)
- Vukadin, Ana. 2015. „Development of a classification-oriented authority control: the experience of the National and University Library in Zagreb.” In *Classification & Authority Control: Expanding Resource Discovery : Proceedings of the International UDC Seminar 2015, 29-30 October 2015, Lisbon, Portugal*, szerk. Aida Slavić, és Maria Inês Cordeiro, 111-120. Würzburg: Ergon Verlag.
- Wimmer, F. 2000. „The Apostrophe in UDC: A critical examination of an unwelcome development.” *Extensions and Corrections to the UDC* 22: 19-27.
- „XML Schema.” World Wide Web Consortium (W3C). Legutóbb módosítva 2004. október 28-án. <https://www.w3.org/XML/Schema>. (megtekintve: 2017-11-01)

6 Az összetett ETO számok automatikus interpretálása

„Az ETO jelei és szimbólumai arra szolgálnak, hogy megtörjék a számok sorrendjét és lehetővé tegyék a prekombinált jelzetek alkotóelemeikre bontását (azaz elemzését). Az ETO tizedes, hierarchikus és szintaktikusan kifejező jelölésmódot használ.”⁵⁰⁰ (UDC Consortium)

6.1 Bevezetés

Az előző fejezetben ismertetett formátum egyaránt alkalmas az egyszerű és szintetizált jelzetek leírására automatikusan feldolgozható formában a jelzetek szerkezetére vonatkozó információ megtartásával. Ugyanakkor *a formátum bonyolultsága miatt szükség van olyan eszközökre is, melyek lehetővé teszik a jelzetek ilyen formátumú reprezentációjának automatikus előállítását.*

Egyfelől lehetséges egy olyan *jelzetépítő szoftver* tervezése és megvalósítása, amely támogatja a jelzetek építését, miközben a háttérben felépíti azok XML reprezentációját is. Bár egy ilyen program hasznos eszköz lehet az osztályozói munka során, az ilyen jelzetépítés akkor igazán hatékony, ha a szoftver *tartalmazza az ETO táblázatokat* is. Ezen kívül ez a megoldás *nem alkalmazható a már meglévő katalógusok feldolgozásához*, formátumok, adatbázisok és szótárak közötti konverzióhoz.

Egy másik megoldás lehet egy olyan **jelzetelemző program** írása, amely a jelzeteket eredeti formájukban karaktorsorozatként fogadja és amennyire lehetséges, *automatikusan* feldolgozza őket.

A lehető legnagyobb rugalmasság érdekében legjobb, ha az algoritmus *bármilyen előzetes átalakítás nélkül* képes az ETO-jelzetek értelmezésére. Ez a megközelítés az alábbi **előnyökkel** jár:

- Bár a különböző információ visszakereső rendszerek különböző adatbázisokat és formátumokat használnak, az *eredeti ETO számnak mindegyikben szerepelnie kell.*
- Ezzel megvalósítjuk azt a célt, hogy az osztályozónak, illetve az authority fájl összeállítójának *csak a jelzetek minél pontosabb összeállításával kell foglalkoznia*, és nem kell speciális, csak egy adott programra jellemző ismeretekkel rendelkeznie.
- A feldolgozási folyamat végén a jelzetek alapján előállított XML könnyen konvertálható más, szabványos vagy nem szabványos formátumokra.
- Az ETO számok feldolgozása *további kutatások alapjául szolgálhat* például a keresés és böngészés hatékonyságának növelésére vagy a komplex számok közötti kapcsolatok feltárására vonatkozóan.

Ez a fejezet egy olyan **ETO interpreter működésének alapelveit, használatát és elérhetőségét** ismerteti, amely a fenti céloknak megfelelően képes az ETO-jelzetek átalakítására a korábban említett XML formátumba. Mind az algoritmus, mind az implementáció **önálló kutatómunka eredménye**.⁵⁰¹ Az alapjául szolgáló ETO kiadások, tankönyvek és folyóiratcikkek fel vannak sorolva, illetve hivatkozva vannak a megfelelő helyeken.

⁵⁰⁰ „UDC Notation”.

⁵⁰¹ A program először a 2014. évi ICAI konferencián (Piros 2014a), illetve a Tudományos és Műszaki Tájékoztatásban lett bemutatva (Piros 2015b), részletesebben pedig a 2015. évi ETO Szemináriumon lett ismertetve (Piros 2015a).

6.2 Az ETO számok automatikus felbontása

A *hatvanas évek óta* több kutatás foglalkozott az ETO számok *felbontásának kérdésével* és mutatott rá a *rendszer jelzetkezelésének következetlenségére*, illetve dolgozott ki javaslatokat annak átalakítására. Freeman és Atherton az *AIP projekt* során az alábbi **problémákat** sorolta fel:

- Vannak olyan *jelek, amelyek csak megjelenítési célt szolgálnak* (pl. a tizedespont)
- Egyes *jelzetelési eszközök ugyanazokat a központosítási jeleket használják*, ami problémát okozhat a számok sorba rendezése során
- Egyes *alosztások felismeréséhez két karakter figyelembe vételére van szükség* (mint például a „=” és „(=” megkülönböztetésekor)
- A „,0” és „,00” esetében a *tizedespont jelentéssel bír*, miközben minden más esetben csak a könnyebb áttekinthetőséget szolgálja.

A fenti problémák kiküszöbölése érdekében az *alosztások bevezető karaktereit betűkre cserélték*, melyek megkönnyítették az elemek beazonosítását, és biztosították a jelzetek megfelelő sorba rendezhetőségét.⁵⁰² Dahlberg számos egyéb, a rendszer egészét érintő módosítás mellett a *pont és az alosztások záró karaktereinek elhagyását* javasolta. Meglátása szerint az összes javaslata életbe lépését követően, mindössze a hely, idő és szempont szerinti alosztások igényeltek volna önálló bevezető karaktereket.⁵⁰³ Buxton szintén az osztályozási rendszer szintjén javasolta a bevezető karakterek betűkre cserélését, és az alosztások bevezető karaktereivel kapcsolatos következetlenségek felszámolását.⁵⁰⁴

Szintén a pont használatával kapcsolatos problémákra mutat rá Strachan⁵⁰⁵ és Mándy⁵⁰⁶, Wimmer pedig az *aposztrófós korlátozottan közös alosztások* egységesebb alkalmazásának szükségességét veti fel.⁵⁰⁷ Gnoli a korlátozottan közös alosztásokat elméleti megalapozottsággal, átfogóan elemezve nemcsak a jelzetelemzés, de a következetesebb facettásítás szemszögéből is bemutatja, milyen módosításokra lenne szükség az osztályozás egységesebbé tételére.⁵⁰⁸

A *jelzetelési és jelzetépítési szabályok formalizálásának szükségességét* a Konzorcium is felismerte, így a következetes, elméleti megalapozottságú facettásítás mellett *ez jelenti a rendszer fejlesztésének másik fő irányát*.⁵⁰⁹ Ez egyfelől jelentősen megkönnyíti a legújabb ETO kiadások gépesítését, másfelől komoly kihívások elé is állítja a szoftverek készítőit, amennyiben a korábbi és az újabb verziókra egyaránt fel kívánják készíteni a programokat.⁵¹⁰

Miközben számos *könyvtári rendszer tartalmaz többé-kevésbé teljes megoldásokat* a jelzetek részekre bontására, tudományos igényű *kutatások is folytak* ezzel a kérdéssel kapcsolatban. A

⁵⁰² Freeman és Atherton 1967; idézi Buxton 1990.

⁵⁰³ Dahlberg 1972.

⁵⁰⁴ Buxton 1990.

⁵⁰⁵ Strachan 2000.

⁵⁰⁶ Mándy 2011.

⁵⁰⁷ Wimmer 2000.

⁵⁰⁸ Gnoli 2011.

⁵⁰⁹ Slavić és Davies 2017.

⁵¹⁰ Piros 2014b.

legalaposabban **Gerhard Riesthuis** járta körül a témát⁵¹¹, a komplex ETO számok felbontásának valamennyi kérdését megvizsgálva annak érdekében, hogy a számokat visszakereshetővé tegye a részeik – Master Reference File-ből származó - természetes nyelvű leírásai alapján. *Algoritmusokat* dolgozott ki az ETO számok több lépcsőben történő feldolgozására, és elkészítette az azokat megvalósító *példaprogramokat* is (Pascal nyelven).⁵¹²

Magyarországon **Mándy Gábor** végzett hasonló jellegű vizsgálatokat, illetve adott közre olyan *példaprogramokat*, melyek képesek egyes szintaktikai relációk és alosztások felismerésére. Az elképzelésének alapja egy olyan programcsomag kidolgozása, melynek algoritmusai lépésenként, egyfajta „szűrőként” viselkedve - minden eljárás az őt megelőző kimenetét kapva bemenetként - bontják részeire az ETO-jelzetet. *Célja az ETO-jelzetek-posztkoordinált használatának elősegítése volt* azzal, hogy kész algoritmusokat nyújt a szoftverek fejlesztőinek.⁵¹³

Ha a szabályokkal kapcsolatos észrevételeket és a megvalósított jelzetelemző megoldásokat összegezzük, általánosságban kijelenthetjük, hogy a jelzetek felbontására vonatkozó korábbi próbálkozások **elsősorban a jelzetek részeik alapján, Boolean-operátorokkal történő posztkoordinált visszakereshetővé tételét tűzték ki célul**. A korábbi kutatások tehát *nem fordítottak kiemelt figyelmet a jelzetelemek közötti kapcsolatok szintaktikai értelmezésére és eltávolására*.

6.2.1 Az ETO számok felbontásának néhány problémája

Ebben a fejezetben néhány példán keresztül is áttekintjük, milyen jellegű problémákkal találkozhatunk az ETO-jelzetek elemzése során.

A leggyakrabban említett következetlenséget a **.0-ás alosztások** okozzák.⁵¹⁴ A pont általában jelentés nélkül, csak a könnyebb olvashatóságot szolgálva szerepel, ezekben az esetekben azonban az alosztást bevezető karaktorsor része. Bár az osztályozás szerkesztői figyelemmel vannak erre, és a táblázati számok nem tartalmazzak .0 karaktorsort, a gyakorlatban vannak olyan jelzetszerkesztési megoldások (elsősorban a $\cong 0/9$ párhuzamos továbbosztás), melyek ilyen eredményeznek.⁵¹⁵

A *párhuzamos továbbosztások* ($\cong$) általában elsősorban a táblázatok rövidebbé és átláthatóbbá tételét, illetve a fogalmi szerkezetek párhuzamosságának a kiemelését szolgálják. Az imént említett $\cong 0/9$ típusúak azonban gyakorlatilag szintagmatikus relációk. Mándy Gábor hozzáírásos jelzetszerkesztésként hivatkozik erre a megoldásra [Mándy 2013]), és az ilyen jelzeteket is felbontandónak tartja, mindkét jelzetelem visszakereshetővé tétele érdekében. Ez viszont pusztán szintaktikai eszközökkel, az MRF használata nélkül megoldhatatlan, hiszen a hozzáférési pontok azonosításához a továbbosztás táblázaton belüli helyének és a forrásának a pontos ismerete egyaránt szükséges.⁵¹⁶ Az itt bemutatott kutatás azt a megközelítést követi, hogy a párhuzamos továbbosztás *nem művelet, hanem speciálisan jelölt enumerikus jelzet*, ezért elemeinek felismerése nem a parser vagy interpreter feladata, hanem további algoritmusoké, amelyek annak a kimenetein alapulnak.

⁵¹¹ Riesthuis 1998.

⁵¹² A kutatást összegző jelentések az Extensions and Correctionsben jelentek meg (Riesthuis 1997, 1999).

⁵¹³ Mándy 2013.

⁵¹⁴ Ide tartoznak az 1999-ben és 2001-ben kivezetésre került .00-ás szempont szerinti alosztások is.

⁵¹⁵ Az 1990. évi magyar kiadás tantárgyakat kifejező 372.8 osztályából például könnyedén képezhetünk olyan jelzeteket, mint 372.814.08 („etika alapjai tantárgy”, Mándy 2011) vagy 372.851.06 („matematikai logika tantárgy”, Piros 2015b).

⁵¹⁶ Riesthuis 1999.

Bármelyik megközelítést is kövessük, jó hír, hogy *az ilyen megoldások apránként, szisztematikusan kivezetésre kerülnek* és enumerikus facetták, illetve csillaggal bevezetett jelzetek veszik át a helyüket.⁵¹⁷

A .0-ás alosztásokéhoz hasonló probléma a **kötőjeles alosztásoké**. Az alapszabály az, hogy az általános ismérvek szerinti nem önálló alosztások -02/-05-el, a kötőjeles korlátozottan közös alosztások pedig -1/-9-el (esetleg -00-val) kezdődnek.

Ugyanakkor az utóbbi nem minden esetben teljesül. Például a patológia és kórtan (616) alatt -02, -03 és -05 kezdetű korlátozottan közös alosztásokkal is találkozhatunk. A -03 alatt található megjegyzés szerint a (-03) általánosan közös alosztások a 616 alatt nem használhatók, így (ha nem is mechanikusan) az ilyen kezdetű számok beazonosíthatók. A -05 kezdetű korlátozottan közös alosztások párhuzamos továbbosztással követik az általánosan közöseket. Ugyanakkor a -022.2 általánosan, a -022.6/.9 korlátozottan közös alosztást jelöl, a -022.1 és -022.3 egyaránt értelmezhető általános és speciális alosztásként is, tehát kétértelmű. A -021.1 pedig már az ETO hierarchikus jellegének is ellentmond, ugyanis míg ez a jelzet egy speciális alosztást jelöl, a -021.111 nem ennek a továbbosztása, hanem egy általánosan közös alosztás.⁵¹⁸

Az **alosztásokon belüli műveletek** szintén okozhatnak meglepetéseket, ha azokat mechanikusan, a kontextus figyelembe vétele nélkül értelmezzük, feltételezve, hogy azok főtblázati számokat kötnek össze.⁵¹⁹ Figyelembe kell venni azt is, hogy etnikai alosztás állhat a földrajzi alosztáson belül, annak száma után, illetve az így kifejezett helyi etnikai csoportok közötti műveleteket is.

A hármas fejezetben már említettük a **külső forrásból származó alosztások** kérdését, azt, hogy A/Z illetve * után a szeparátor karakterekről nem dönthető el, hogy a külső jelzetnek a részei-e vagy új jelzetelemet vezetnek be.⁵²⁰

Az egyik legellentmondásosabb alosztás csoport az **aposztrófos korlátozottan közös alosztásoké**, melyek a táblázatok egyes részein egyszerű enumerikus alosztásokként, máshol pedig – a harmadik fejezetben részletesen tárgyalt - aposztrófos jelzetszintézisként, azaz korlátozottan alkalmazható relációként jelennek meg. Ugyanakkor arra is van példa, hogy az aposztróf névleg szintézist jelöl, azonban az általa bevezetett jelzet a táblázatok más részéről származik, ami gyakorlatilag egy a korlátozottan közös alosztásokon belüli párhuzamos továbbosztásnak felel meg.⁵²¹

A **közbeékelés (intercalation)** elsősorban a jelzetek besorolási sorrendjének a befolyásolása miatt fenntartott lehetőség, ami több kérdést is felvet. Először is a szoftvereknek fel kell készülniük arra, hogy a külső jelzet az alosztás utáni résszel kiegészítendő. Másrészt félreértésekre adhat okot, hogy

⁵¹⁷ Gnoli 2011.

⁵¹⁸ Piros 2015, OSZK KI 2005: 1/106, 2/31-33.

⁵¹⁹ Smiraglia et. al. (2013) például leírja, hogy a jelzeteket a pluszjelek mentén mechanikusan felbontva olyan jelzeteket kaptak, amelyek a négyes, tehát betöltetlen főosztályba tartozónak tündek. Az eredményeket megvizsgálva kiderült, hogy a problémát a földrajzi alosztásokon belüli összekötések okozták.

⁵²⁰ Az itt bemutatott interpreter a másodjára említett, a szabályokhoz jobban alkalmazkodó megoldást implementálja, azaz a külső alosztásból származó jelzeteket követően szereplő terminátor karaktereket is érvényesnek tekinti.

⁵²¹ Erre példa az állati termékek 637 vagy az ötvözetek 669 alatti kifejezése. Az aposztrófok kérdését részletesen tárgyalja Wimmer (2000). A helyzetet tovább bonyolítja, hogy egyes ETO kiadásokban arra is van példa, hogy egy szintézist nem aposztróf, hanem kötőjel vezet be, mint a [355/359]-5/9 esetében (OSZK KI 2005: 315).

ha az alosztás a pont helyére van beszúrva, akkor a pont meghagyása és elhagyása egyaránt elfogadott.⁵²²

A formai alosztások, illetve a -029 kezdetű tulajdonságok szerinti **alosztások tartalmazhatnak főtáblázati számokat** (kettősponttal bevezetve). Ilyenkor figyelembe kell venni, hogy ezek a számok az általánosan közös alosztáshoz tartoznak.

Az **idő alosztásokat** leíró egyszerű XML típus kapcsán láttuk, hogy ezek egy része formailag teljesen eltér a többi ETO-számtól, köszönhetően annak, hogy a dátumok és időpontok, továbbá az időmértékek leírása is ezekkel történik. A **földrajzi alosztások** is hasonlóan sokrétűek, miután a természetföldrajzi és geopolitikai szempontból meghatározható helyek mellett lehetőséget adnak a hosszúság, terület és térfogatmértékek (a (18) alatt), illetve a földrajzi szélesség és hosszúság [(161/164)] kifejezésére is. Ezek a régebbi kiadásokban gyakorlatilag számvégződéses alosztásokként jelentek meg.⁵²³ A *revíziók* óta már csillaggal kell őket jelölni, ebből azonban az is következik, hogy meg kell feleltetni egymásnak a régi számvégződéses, és az új, külső forrásból származóként felismert jelzeteket.⁵²⁴

A fejezetben szereplő példák a legszembetűnőbb speciális esetek, melyek megnehezítik egy ETO elemző írását. Ezek mellett *a különböző ETO kiadások számos további speciális esetet tartalmaznak*, amelyekre a szoftvereknek mind fel kell készülniük.

6.3 Az interpretálás alapelvei

A **legfontosabb elvárás** az újonnan kifejlesztett interpreterrel szemben, hogy a feldolgozás során *vegye figyelembe az ETO összetett számok képzésére vonatkozó szintaktikai szabályait* úgy, hogy amennyire lehetséges, a jelzetekben található valamennyi információ megmaradjon.⁵²⁵ Részletesebben kifejtve ez az alábbiakat jelenti:

- Az algoritmusnak pontosan **azokat a számokat kell felismernie, amelyek betartják az ETO jelzetalkotási szabályait**, még akkor is, ha azok ritkán vagy szinte soha nem jelennek meg egy adott állomány indexelési gyakorlatában.⁵²⁶
- Az algoritmusnak **a jelzetben tárolt valamennyi információt ki kell nyernie**. Mindössze néhány jelzetelem vagy reláció felismerése nem elég, ahogy az sem, ha *a jelzetet alkotó számok* mind felismerésre kerülnek ugyan, de *a kontextusukra és szerepükre vonatkozó információ* nem.
- A feldolgozó folyamatnak, amennyire csak lehetséges, **szintaktikai alapúnak kell maradnia**: a számok feldolgozásának a jelzetszerkesztési szabályokon kell alapulnia, a számok jelentésére vonatkozó információt csak ott vehet igénybe, ahol ez elkerülhetetlen.
- A feldolgozásnak teljesen **automatikusnak kell lennie**: a programnak képesnek kell lennie a számok elemzésére felhasználói beavatkozás nélkül.

⁵²² „Punctuation and intercalation”.

⁵²³ OMIKK és OSZK-KMK 1990: 21-22, OSZK KI 2005: 1/38-39, BSI 2005: 26.

⁵²⁴ Hasonló módosítások a táblázatok számos pontján következtek be, mindenütt felvetve ugyanezt a kérdést.

⁵²⁵ A fejezetben leírt alapelvek a 2015. évi ETO Szemináriumon bemutatásra kerültek (Piros 2015a).

⁵²⁶ Ezzel nagyobb stabilitás és időbeli konzisztencia érhető el, mintha a gyakorlatban általában használt jelzetekből indulnánk ki. Ugyanakkor előfordulhat, hogy egyes, akár elterjedt, de a szabályoknak nem megfelelő jelzetelési gyakorlatok feldolgozhatatlan számokat fognak eredményezni.

A szintaktikai alapú feldolgozást alapvetően három dolog indokolja:

- A jelzetváltozások által okozott problémák elkerülése. Ha egy szám megszűnik, utána újra felvehető, megváltozott jelentéssel. A korábbi szabályzat szerint volt egy tíz éves kötelező várakozási idő az osztály megszűnését követően⁵²⁷, melytől később a módosítások tervezett tempója miatt el kellett tekinteni.⁵²⁸
- Ha általános, netán nyílt forráskódú megoldást szeretnénk kidolgozni, le kell mondanunk arról, hogy a szoftver tartalmazza a Master Reference File-t, mivel annak felhasználása licenszhez kötött.⁵²⁹
- A nyelvi elemzés során nagyobb stabilitást eredményez az, ha elsősorban a szintaktikai szabályokra építünk, és csak a kivételek esetén használunk szemantikai információt, másfelől ez az időbeli konzisztenciát is növeli, a program működését ugyanis csak a jelzetszerkesztési szabályokat is érintő változtatások befolyásolják, a táblázatok átszervezései nem.

Mivel a távlati cél az interpreter **Software-as-a-Service (SaaS)** modellt⁵³⁰ követő elérhetővé tétele, az online elérhetőség szintén kiemelt szempont volt a fejlesztés során. Ez a gyakorlatban azt jelentené, hogy *az ETO számokat használó programok fejlesztőinek a jövőben nem lenne szükségük arra, hogy - akár saját program írásával, akár külső modulok beépítésével – a programon belül oldják meg az ETO számok felbontását.* A programoknak csak egy vékony középrétegre lenne szükségük, mely képes a külső szolgáltatás megszólítására, illetve az általa küldött válasz feldolgozására.

Ehhez, a minél szélesebb körű felhasználhatóság érdekében természetesen a jelzetek feldolgozása mellett a szolgáltatásnak képesnek kell lennie az eredmény **minél több formátumba történő konvertálására** is.

A fenti követelmények alapján a megvalósításra a **Java programozási nyelv** lett kiválasztva, a kapcsolódó technológiák – elsősorban az XML dokumentumok összeállításának, elemzésének és validálásának, illetve a webes szolgáltatások fejlesztésének – nagymértékű és rugalmas támogatása miatt.

6.4 Az interpreter működésének alapjai

Az ETO jelzetalkotási szabályai egy **formális nyelvet** alkotnak a decimális számok, a pont és az ETO metanyelvi jelei által alkotott **ábécé** fölött.⁵³¹ Az interpreter egy LL(2) elemző **automata**⁵³², amely ezt a

⁵²⁷ FID 1983.

⁵²⁸ FID 1990; hiv. Slavić-Overfield 2005: 19.

⁵²⁹ „UDC Licenses”

⁵³⁰ Levinson 2007.

⁵³¹ Az ábécé elméletileg a decimális számjegyeket, az alosztási jeleket és az alosztások bevezető karaktereit tartalmazza (Piros 2014a). Bár a külső forrásból származó, illetve névalosztások tetszőleges írásjelet tartalmazhatnak, ennek nincs jelentősége, mivel ezeket az alosztásokat a parser fel tudja ismerni, és a következő terminátor jelig el tudja fogadni a karaktereiket.

⁵³² Bach 2000: 167-182. LL(1) elemzésre azok miatt az alosztások miatt van szükség, amelyeket két karakter vezet be, mivel az első karakter önmagában általában bevezethet más típusú alosztást is (például formai, illetve népi alosztások). Az LL(2) elemzést kizárólag a (már törölt) szempont szerinti alosztások indokolják. A publikált parserek közül szintén balelemzést valósítanak meg Mándy Gábor algoritmusai (2013).

nyelvet ismeri fel. A program elkészítésekor ugyanazok a kiadások⁵³³, tankönyvek⁵³⁴ és egyéb források⁵³⁵ lettek felhasználva, melyek az előző részben leírt formátum tervezésekor.

A bemenő szavak az ETO-jelzetek, míg az automata állapotai annak a számnak a típusát és a jelzeten belüli szerepét tükrözik, amelyikhez az aktuális bemenő karakter tartozik. A feldolgozás során az automata egy fát generál, mely tartalmazza a jelzet részeit az alosztási jelek és táblázatok korábbi fejezetekben már leírt precedenciájának megfelelően. A generált fa a jelzetrészeken és kapcsolataikon kívül még tartalmaz egyéb járulékos információkat is, melyek fontosak lehetnek a keresés során.

Ha az automata **nem ismeri fel** a bemenő szót, az azt jelenti, hogy *a megadott jelzet nem helyes, vagy legalábbis nem interpretálható a megadott ETO verzióban*. Ez azt jelenti, hogy az automata által felismert nyelv éppen a megadott évi ETO verzióban összeállítható jelzetek halmaza.

Az **algoritmus bemenete** az ETO-jelzet és az összeállításához használt ETO kiadást azonosító évszám, illetve a jelzet *természetes nyelvű leírásai* (tetszőleges nyelveken). A **kimenete** a jelzet XML reprezentációja, vagy amennyiben azt nem lehetett feldolgozni a megadott verzióban, *a hibát leíró üzenet*.⁵³⁶

Ahogy már korábban említettük, bár az évszám használata – amennyiben a szabályok változásának a bevezetését vesszük alapul – megfelel a szintaktikai analízis céljaira, a táblázati számok azonosítására és a Master Reference File rekordjaiknak való megfeleltetésére *az MRF adatbázis azonosítójának használata célszerűbb lenne*⁵³⁷. Ezért a további fejlesztési tervek között szerepel a bemenő adatok ennek megfelelő módosítása.

A cél tehát egy olyan algoritmus kidolgozása, amely képes az ETO-jelzetek feldolgozására, amennyire csak lehet, szintaktikai úton. Ez azt jelenti, hogy *ha két jelzet struktúrája azonos, a feldolgozásuk eredményeként kapott struktúrák is hasonlóak* lesznek, mindössze az őket alkotó ETO számokban térve el.

A szintaktikai úton történő feldolgozás igényét úgy is megfogalmazhatjuk, hogy *a feldolgozás során az interpreternek nem kell azonosítania magukat a számokat, csak a jelzetben szereplő központosítási jelek, illetve a jelzetalkotási szabályok alapján kell meghoznia a döntéseket*. Ugyanakkor vannak olyan korlátozottan közös alosztások és **speciális jelzetalkotási módok**, amelyek nem értelmezhetők egyértelműen a táblázatok béli helyük pontos azonosítása nélkül. Korábban már láttuk, hogy az aposztrófos korlátozottan közös alosztások a táblázatok egyes helyein (pl. a 329, illetve az 547 osztályokban) az eredeti funkció szerint jelzetek szintetizálására szolgálnak, másutt viszont enumerikus korlátozottan közös alosztásként használandó, ahogyan a párhuzamos alosztások

⁵³³ Tehát a legutóbbi magyar nyelvű nyomtatott kiadások (OMIKK és OSZK-KMK 1990 és OSZK KI 2005), a British Standard Institute nyomtatott angol nyelvű szabvány kiadása (BSI 2005), a UDC Summary („UDC Summary”) és az angol nyelvű UDC Online („UDC Online”).

⁵³⁴ Babiczky Béla: Osztályozás és szakkatalógus szerkesztés (1975), Babiczky B. és Barátné Hajdu Ágnes: Bevezetés az információkereső nyelvek elméletébe és gyakorlatába (1998), Eszenyiné Borbély Mária, és Salgáné Medveczki Marianna. 2010. Komplex könyvtári feldolgozó munka gyakorlata (2010) és Pálvölgyi Mihály: Információfeldolgozás, információkereső nyelvek (1999).

⁵³⁵ „Extensions and Corrections to the UDC.”, „Cancellations to the UDC” és „Major Changes to the UDC 1993-2013”.

⁵³⁶ Piros 2014a.

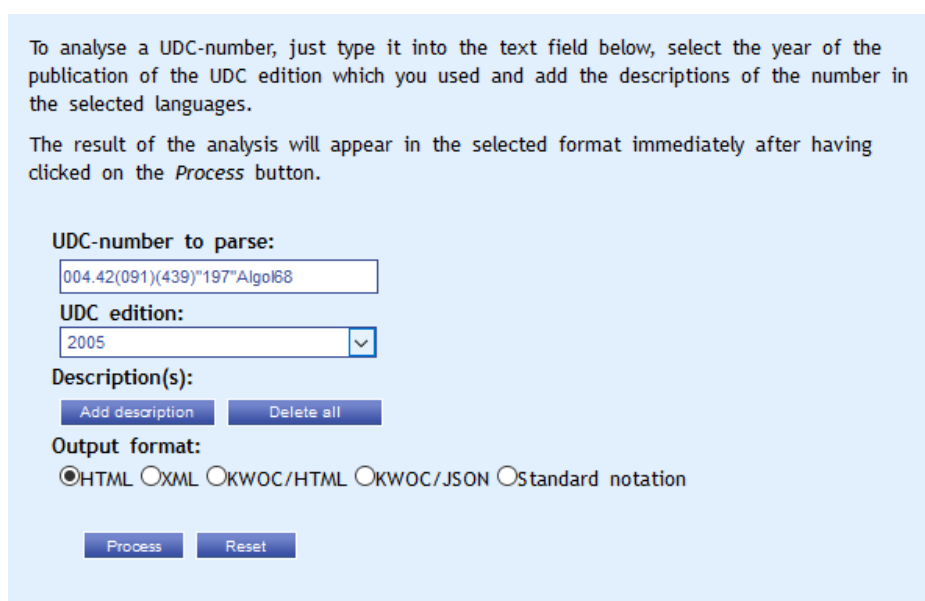
⁵³⁷ Slavić és Isaac 2009

kérdéséről is volt szó korábban. A **verziók közötti eltérések** – pl. bizonyos alosztások be- és kivezetése – szintén indokolhatja a részjelzetek azonosítását.

Az ilyen jellegű döntésekhez szükséges ismeretek egyelőre az interpreter egy lazán csatolt moduljában, *forráskód szinten* vannak tárolva. A későbbiekben, éppen a lazán csatoltság miatt ez kiváltható egy olyan modullal, amely az MRF adatbázist használja.

Az interpreter – az ETO változásait, illetve a kutatás újabb eredményeit követve – folyamatos fejlesztés alatt áll. Mindenkori legfrissebb verziója elérhető online az alábbi címen: **<http://piros.udc-interpreter.hu#process>**.⁵³⁸

A „ETO számok feldolgozása” menüpontra kattintva begépelhetjük az ETO számot, megadhatjuk az ETO kiadás évszámát és megadhatjuk a szám leírásait különböző nyelveken. A „Feldolgozás” gombra kattintva a szám feldolgozásra kerül (9. ábra).



9. ábra Az interpreter beviteli űrlapja

Az eredmény elérhető közvetlen **XML formátumban** is, de a szoftver - a kiválasztott formátumnak megfelelően – képes a **HTML** formátumú megjelenítésre, illetve a jelzetelemekből alkotott **KWOC index** összeállítására is.

6.5 A program kimenete

6.5.1 Példák az ETO számok elemzésére és interpretálására XML formátumban

A program elsődleges és legfontosabb kimeneti formátuma az XML. Ebben a fejezetben erre látunk példákat.

Az első az **irodalmi művek** kifejezésére szolgáló speciális jelzetszerkesztési mód leírását mutatja be:

```
<?xml version="1.0" encoding="UTF-8"?>
```

⁵³⁸ A jelenleg elérhető szoftver az XML séma definíció 1.1-es verziója alapján működik, de folyamatban van az itt ismertetett 1.2-es verzióra történő továbbfejlesztés, illetve az új verzió publikálása.

```

<ns:udc_concept xmlns:ns="http://library.inf.unideb.edu/udc/xml"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" udc_edition="2005"
notation="821.111SHAK7ROM.03=112.2">
  <ns:description xml:lang="EN">
    Shakespeare: Romeo and Juliet (translated to German)
  </ns:description>
  <ns:main_table_number number1="821.111">
    <ns:special_auxiliary
      xsi:type="ns:special_auxiliary_number_finaldigits"
      number1="7"/>
    <ns:special_auxiliary
      xsi:type="ns:special_auxiliary_number_pointnought"
      number1=".03"/>
    <ns:alphabetical_specification order="1" text="SHAK" standard=""/>
    <ns:alphabetical_specification order="2" text="ROM" standard=""/>
    <ns:common_auxiliary_independent
      xsi:type="ns:common_auxiliary_of_language" order="1">
        <ns:common_auxiliary_of_language_number number1="=112.2"/>
      </ns:common_auxiliary_independent>
    </ns:main_table_number>
  </ns:udc_concept>

```

A fenti XML az 821.111SHAK7ROM.03=112.2 jelzet (Shakespeare: Rómeó és Júlia német nyelven) feldolgozásának eredménye.

A fogalom objektum tartalmazza magát a jelzetet az eredeti formájában, a leírását különböző nyelveken és az ETO kiadást azonosító évszámot. Az elemei a fő táblázati számot és a hozzá kapcsolódó alosztásokat írják le. Az interpreter képes azonosítani mind a szerzőt és a címet leíró betűjeles, mind a „művet” és a „fordítást” jelentő korlátozottan közös alosztásokat.

Az alábbi, összetettebb példa⁵³⁹ a kiterjesztést, azaz az ETO számok intervallumának képzését mutatja be:

```

<?xml version="1.0" encoding="UTF-8"?>
<ns:udc_concept xmlns:ns="http://library.inf.unideb.edu/udc/xml"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" udc_edition="2005"
notation="323 (47+57) 1/.2:81 (47+57) '01">
  <ns:description xml:lang="EN">
    National and ethnic minorities. Internal political activities. Origins (of
    languages). Soviet Union.
  </ns:description>
  <ns:main_table_relation>
    <ns:main_table_number number1="323.1" number2="323.2">
      <ns:common_auxiliary_independent
        xsi:type="ns:common_auxiliary_of_place">
          <ns:common_auxiliary_of_geography_addition>
            <ns:common_auxiliary_of_geography_number number1="(47)"/>
            <ns:common_auxiliary_of_geography_number number1="(57)"/>
          </ns:common_auxiliary_of_geography_addition>
        </ns:common_auxiliary_independent>
      </ns:main_table_number>
    <ns:main_table_number number1="81">
      <ns:special_auxiliary
        xsi:type="ns:special_auxiliary_number_apostrophe" number1="'01"/>
      <ns:common_auxiliary_independent
        xsi:type="ns:common_auxiliary_of_place">
          <ns:common_auxiliary_of_geography_addition>
            <ns:common_auxiliary_of_geography_number number1="(47)"/>

```

⁵³⁹ A példa forrása Riesthuis (1997).

```

        <ns:common_auxiliary_of_geography_number number1="(57)"/>
    </ns:common_auxiliary_of_geography_addition>
</ns:common_auxiliary_independent>
</ns:main_table_number>
</ns:main_table_relation>
</ns:udc_concept>

```

A példa tartalmazza a 323.1/2 intervallumot, a végpontjaival kifejezve. Ez a módszer a paradigmatis, a táblázatok által tartalmazott kiterjesztések és a szintagmatikus, az indexelő által képzett kiterjesztések esetén egyaránt hasznos lehet. Így ugyanis *a jelzet ugyanúgy visszakereshető az intervallum részei, mint maga a kiterjesztés alapján*. Ez ilyen jelzetrepresentáció alapján (az MRF adatbázis használatával) *lehetséges az intervallum felbontása is*.⁵⁴⁰ A földrajzi alosztáson belüli összekötést szintén meg kell tartani ebben az esetben, hiszen ez szintén egy a táblázatok által tartalmazott, tehát paradigmatis jelzetkombináció, ezért a teljes összekötés és annak elemei alapján egyaránt visszakereshetővé kell tenni.

A példa azt is bemutatja, hogy az interpreter a közbeszúrás (*intercalation*) és az általánosan közös alosztásokat követő korlátozottan közös alosztások kezelésére is képes.

6.5.2 KWOC index

Bár az XML szabványos, automatikusan feldolgozható formátum, közvetlenül nem támogatja az ETO használatának minden formáját. Ezért logikusnak és szükségesnek tűnik olyan konverziós módszerek elérhetővé tétele, amelyek *a szoftverek számára könnyebben felhasználható* formátumokat produkálnak.

Az XML formátum mellett az interpreter HTML formában is képes megjeleníteni a legenerált fákat, illetve, ha szükséges, összeállítja a jelzetelemek listáját a kontextuális információ nélkül, **JSON formátumban**.⁵⁴¹

Az alábbi példa egy ETO-jelzet elemeinek listáját mutatja, a szoftver által összeállított JSON sztring formájában:

```

{
  "concept": "378.007.1",
  "udc_edition": "1990",
  "pref_labels": {
    "pref_label_1": {"pref_label": "Főiskolák vezetése", "language": "HU"}
  },
  "udc_numbers": {
    "number_1": {
      "notation": "378", "filing": "3T7T8C", "uri": "http://udcdata.info/025169",
      "pref_labels": {
        "pref_label_1": {"language": "HU", "pref_label": "Felsőoktatás. Egyetemek. Főiskolák"}
      }
    },
    "number_2": {
      "notation": ".007.1", "filing": "P0T0T7T1C",
      "pref_labels": {
        "pref_label_1": {"language": "HU", "pref_label": ""}
      }
    }
  }
}

```

⁵⁴⁰ Az intervallumok felbontásának módszerét részletesen tárgyalja Buxton (1990) és Riesthuis (1999)

⁵⁴¹ A JSON (JavaScript Object Notation) egy önleíró, könnyen érthető, nyelvfüggetlen adatsere formátum („JSON Data Interchange Format”). A JSON szerializáció során a program az ETO szám által meghatározott objektumhierarchiát alakítja át a megfelelő sztringgé az alkalmazások közötti adatsere számára.

```

}
}
}

```

A 10. ábrán látható példa pedig egy **HTML-formátumú KWOC-indexet** mutat be.

Concept:		
004.42(091)(439)"197"Algol68		
UDC edition:		
2005		
Description(s):		
KWOC-index*:		
<u>Notation</u> ▲	<u>URI</u>	<u>English</u>
004.42	http://udcdata.info/013945	Computer programming. Computer programs
Algol68		
(091)	http://udcdata.info/001930	Presentation in chronological, historical form. Historical presentation in the strict sense
(439)	http://udcdata.info/003965	Hungary. Hungarian Republic. Magyarország. Magyar Köztársaság
"197"		
*The descriptions are taken from UDC Summary page , which has been released under a Creative Commons Share Alike 3.0 license .		

10. ábra KWOC index, mint a feldolgozás eredménye

6.5.3 A jelzetelési hibák felismerése

Az ETO számok elemzése az esetleges **indexelési hibák** felismerésére is alkalmat ad. A 11. ábra néhány gépelési hiba felismerésére mutat példát.⁵⁴²

⁵⁴² A fejezet példái bemutatásra kerültek a 2015. évi ETO Szemináriumon (Piros 2015a).

The parsing process failed because of an unexpected error.
There must be a point ('.') after every three digits!
796.332092.2

The parsing process failed because of an unexpected error.
The point ('.') stands in an invalid place.
796.332.092.2

The parsing process failed because of an unexpected error.
The character is invalid in this place.
796.332.092.2.097(100"92"

11. ábra Gépelési hibák felismerése

Az ETO kiadás hibás beállítása szintén hibákat eredményezhet, ahogy az a 12. ábrán látható.

UDC-number to parse:
511-027.22-37
UDC edition:
1996
Description:
English, American
Description
Add description

The parsing process failed because of an unexpected error.
The character is invalid in the given UDC version.
511-027.22-37

UDC-number to parse:
511-027.22-37
UDC edition:
1999
Description:
English, American
Description
Add description

Processing the given UDC-number resulted in
- 511-027.22-37
(UDC edition: 1999)
- main table number
 number : 511
- common auxiliary of properties
 number : -027.22
- special auxiliary subdivision
 number : -37

12. ábra Hibás verzió felismerése

6.6 A szoftver tesztelése

A kutatásnak már a kezdeti szakaszában szükségessé vált egy **tesztkészlet** felépítése a szoftver integritásának megőrzése, továbbá az ETO szabályainak kellően alapos feltérképezése és analízise érdekében.⁵⁴³

A tesztkészlet több, mint 700 *tesztesetet* tartalmaz, a tesztek céljai szerint csoportosítva. Vannak tesztek a különböző szabályok, illetve az alosztási jelek, kiterjesztés, csoportképzés, külső forrásból származó jelzetek és névalosztások stb. alkalmazásakor felmerülő speciális esetek megfelelően kezelésének ellenőrzésére.

Minden teszteset egy ETO-számot tartalmaz, az összeállításához használt ETO-verzió évszámával, illetve az XML-t, amit a szám feldolgozása után az interpreternek produkálnia kell.

A tesztek *manuálisan vagy automatikusan is felhasználhatók* annak ellenőrzésére, hogy a feldolgozás eredménye megfelel-e az elvárásoknak. Mivel az interpreter online elérhető, az automatikus ellenőrzéshez elegendő egy alkalmazás, amely a megfelelő paraméterekkel elküld egy http kérést, majd a választ összeveti az ETO szám mellé eltárolt XML-lel.

Általában a teszteseteknek nem szükséges valós, jelentéssel bíró ETO-számokat tartalmazniuk, erre csak akkor van szükség, ha a központosítási jelek önmagukban nem határozzák meg egyértelműen a jelzetelemek típusát és feladatát.

Legtöbb jelzet könyvtári katalógusokból, osztályozási tankönyvekből és cikkekből lett összegyűjtve, illetve, amennyiben nem sikerült olyan példát találni, amelyekben a tesztelendő megoldás előfordult volna, határozottan a teszt céljára lett létrehozva.

A szoftver karbantartása mellett egy ilyen tesztkészlet hasznos példákat szolgáltat az XML-formátum használatára, illetve *lehetőséget ad az ETO-számok prekoordinációjára vonatkozó szabályok felülvizsgálatára és jelentésük megértésére.*⁵⁴⁴

6.6.1 Esettanulmányok

Egy teszter alkalmazás nemcsak a szoftver ellenőrzésére alkalmas, hanem arra is, hogy teljes gyűjteményeket egyszerre feldolgozzunk vele, elmentve a kimeneteket. Ez lehetőséget ad esettanulmányok készítésére.

Az első ilyen esettanulmányra 2015 végén került sor a *The European Library (TEL)* szabad hozzáférésű adatbázisa alapján.⁵⁴⁵ A több mint 100 elérhető gyűjtemény közül a **Portugál Digitális Nemzeti Könyvtár** került kiválasztásra, elsősorban a közepes méretű állomány, és az ETO számokkal való nagy mértékű lefedettsége miatt.⁵⁴⁶

Az adatbázis RDF/XML formában való letöltését és az ETO-számok kinyerését követően a duplikációk törlésre kerültek. Az így nyert lista 13 741 különböző ETO-számot tartalmazott, melyek egy teszter

⁵⁴³ A tesztkészlet is bemutatásra került a 2015. és 2017. évi ETO Szemináriumokon (Piros 2015a, 2017).

⁵⁴⁴ A tesztkészlet szintén letölthető és felhasználható Creative Commons Nevezd meg! - Ne add el! - Így add tovább! 4.0 Nemzetközi Licenz („CC BY-NC-ND 4.0”) szabályai szerint az alábbi címről: <http://piros.udc-interpreter.hu/#tests>.

⁵⁴⁵ Az esettanulmány eredményei be lettek mutatva a 2017. évi ETO Szemináriumon (Piros 2017).

⁵⁴⁶ A The European Library (TEL), és a Portugál Digitális Nemzeti Könyvtár tesztelésre való felhasználásának ötletét Dr Nuno Freire vetette fel a 2015. évi ETO Szemináriumon. Bár a TEL portál karbantartása és frissítése 2016. december 31-én véget ért, az adatbázisok továbbra is elérhetők („The European Library Open Dataset”).

alkalmazáson keresztül egyesével történtek feldolgozásra. Ezen a módon a szolgáltatás néhány perc alatt a teljes listát feldolgozta.

A 13 741 szám közül 13 604-et sikerült hiba nélkül feldolgozni. A maradék 137 rekord megvizsgálása során két *programhiba*ra és öt még olyan *speciális osztályozási megoldásra* derült fény, amelyeket a program még nem támogatott. A többi problémát *gépelési hibák* és *szabálytalan indexelési megoldások* okozták, vagy a jelzetalkotáskor használt és a feldolgozáskor megadott ETO verziók különbsége. A feldolgozási hibák mellett az *XML validáció* további gépelési hibákat és szabálytalan megoldásokat is felszínre hozott.

A teljesítmény tesztelése mellett az esettanulmány tapasztalatai kiváló visszajelzést jelentettek az adatformátum és a program további korrekciója és továbbfejlesztése számára. Az így szerzett tapasztalatok már be vannak építve az XML Séma definíció itt ismertetett 2.1-es verziójába.

Jelenleg hasonló esettanulmányok vannak tervben a *Spanyol Nemzeti Könyvtár* és *magyar könyvtárak* OPAC-jaiból származó jelzetek, illetve a *Köztaurusz* bevonásával.

6.7 További fejlesztési irányok és kutatási tervek

6.7.1 A szoftver fejlesztése

A bemutatott interpreter jelenlegi verziója az XML séma definíció egy korábbi (1.1) verziója alapján működik. Ezért a korábbiakhoz képest számos fejlesztést tartalmazó 2.1-es verziójú **formátumhoz való igazítása** minden egyéb fejlesztéshez képest előnyt élvez.

Hasonlóan fontos az esetlegesen talált hibák, nem megfelelően kezelt jelzetelési megoldások és szabályok *javítása* is. Ebben nagy segítséget nyújtanak a már említett esettanulmányok.

Nagyon fontos **további kimeneti formátumok** megvalósítása is, amely lehetővé teszi az ETO-jelzetek további szabványos formátumokban történő adatszerezését, feldolgozását más programok által. Jelenleg a korábban már említett *osztályozási MARC formátumokra* való konverzió (beleértve az UNIMARC javaslatok szerint továbbfejlesztett, de nem szabványosított formáját) van tervben és folyamatban. Nagyon fontos az *ontológia leíró nyelvekre* (pl. OWL 2) és más *RDF alapú formátumokra* való konverzió is. Ehhez azonban először az ETO jelzetek ilyen formában való reprezentációjának elméleti hátterét kell megvizsgálni.

Mivel a szoftver elkészítésének a fő célkitűzése, hogy az ETO számokat használó programok közvetlenül is felhasználhassák a kimeneteit, fontos, hogy a megfelelő interfészekon keresztül elérhető legyen. Ezért tervben van, illetve elkezdődött *a szolgáltatás RESTFul formában való újratevezése*.⁵⁴⁷

6.7.2 A szoftver és a formátum felhasználása egyéb területeken

Az osztályozástudomány célja az osztályozási rendszerek formájának, struktúrájának, jelentésének és fejlődésének a kérdéseit és az információ visszakeresésre tett hatásait vizsgálja. A **kvantitatív módszerek** nem tartoznak a tudományterület hagyományos, standard eszközei közé, pedig a

⁵⁴⁷ A REST („Representational State Transfer”, Reprezentációs Adatátvitel) egy osztott hipermédia rendszerek számára létrehozott tervezési stílus. A REST olyan architektúráis megszorításokat definiál, melyek maradéktalan megvalósítása biztosítja a komponensek interakciójának gyorsaságát és skálázhatóságát, az interfészek generikusságát, a komponensek független telepíthetőségét (Fielding 2000).

statisztika és a komplexitáselmélet módszerein keresztül hasznos következtetéseket lehet levonni egy adott rendszer szerkezetére és használatára vonatkozóan.⁵⁴⁸ A jelzetek részletes szintaktikai elemzése nyomán sokkal több információnk áll rendelkezésre automatikusan értelmezhető formában, ezért kifinomult statisztikai vizsgálatokat végezhetünk az ETO-val feltárt gyűjtemények feldolgozása alapján.

A megfelelő elméleti kutatások után a szoftver kimenete felhasználható az **ETO-jelzetel hasonlósági vizsgálatainak és a jelzetek klaszterezésének**⁵⁴⁹, illetve kifinomult algoritmusokkal történő **visszakeresésének** alapjaként. A jelzetek XML reprezentációja elősegítheti az **ETO számok vizualizációját**, vagy felhasználható olyan speciális kutatásokban, mint az ETO-jelzetek alapján történő **címadás**.⁵⁵⁰

A kutatás során felgyülemlett tapasztalatok, az interpreter algoritmus és a tesztkészlet alapján egy új szemszögből tekinthetjük át az ETO jelenlegi és korábbi szintaktikai szabályait. Megvizsgálhatjuk, hogy az utóbbi idők módosításai valóban minden esetben elősegítették-e a következetesebb jelzetkezelést, vagy hogy a szintaktikai relációkból milyen következtetések vonhatók le a szemantikaiakra vonatkozóan, ez pedig elvezethet a prekoordinált ETO-jelzetek **Linked Data** formában történő alkalmazásához.⁵⁵¹ Végül, de nem utolsósorban leírhatjuk az **ETO formális nyelvtanát**.⁵⁵²

6.8 Irodalomjegyzék

- Babiczy Béla. 1975. *Könyvtári feldolgozó munka: A tanítóképző intézet II. évfolyama számára (népművelő-könyvtáros szak). 2. rész. Osztályozás és szakkatalógus szerkesztés.* Budapest: Tankönyvkiadó.
- Babiczy Béla, és Barátné Hajdu Ágnes. 1998. *Bevezetés az információkereső nyelvek elméletébe és gyakorlatába.* Budapest: Universitas.
- Bach Iván. 2002. *Formális nyelvek.* Budapest: TYPOTEX. <http://mek.oszk.hu/05000/05099/05099.pdf>. (megtekintve: 2017-01-01)
- Buxton, Andrew. 1990. „Computer Searching of UDC Numbers.” *Journal of Documentation* 46 (3): 193-217.
- British Standards Institution (BSI). 2005. *Universal Decimal Classification: standard edition: volume 1: systematic tables.* London: British Standards Institution.
- „Cancellations to the UDC.” UDCC (UDC Consortium). <http://udcc.org/index.php/site/page?view=cancellations>. (megtekintve: 2017-11-01)
- Csabay Károly. 1999. Cédulák, ETO-szám, szócikkek. *Tudományos és műszaki tájékoztatás* 46 (1): 12-32. http://tmt-archive.omikk.bme.hu/show_news.html?id=1933&issue_id=4. (megtekintve: 2017-11-01)
- Dahlberg, Ingetraut. 1972. „Az ETO újjáalakításának lehetőségeiről” [Fordította Horváth Magda]. *Könyvtári Figyelő* 18 (4): 412-430.

⁵⁴⁸ Salah et. al. 2011 és Smiraglia et. al. 2013.

⁵⁴⁹ Lula és Cieraszewska 2017.

⁵⁵⁰ Šauperl 2015.

⁵⁵¹ Piros 2015a.

⁵⁵² Az ETO formális nyelvtanának BNF formában történő leírásával kísérletezett Csabay Károly is (1999). Az általa közreadott grammatika az ETO legtöbb általános szabályát megfelelően kezeli, de nem veszi figyelembe a speciális szabályokat.

- Eszenyiné Borbély Mária, és Salgáné Medveczki Marianna. 2010. *Komplex könyvtári feldolgozó munka gyakorlata*. Budapest: Kempelen Farkas Hallgatói Információs Központ.
http://www.tankonyvtar.hu/hu/tartalom/tamop425/0046_komplex_konyvtari_feldolgozo_munka_gyakorlata/index.html. (megtekintve: 2017-11-01)
- „Extensions and Corrections to the UDC.” UDCC (UDC Consortium).
<http://udcc.org/index.php/site/page?view=ec>. (megtekintve: 2017-11-01)
- FID (International Federation de Documentation). 1983. *Az Egyetemes Tizedes Osztályozás (ETO) alapelvei, revíziójának és kiadásának szabályai (FID 603)*. Budapest: OMIKK.
- FID (International Federation de Documentation). 1990. *Task Force For System Development: final report*. Hága: International Federation de Documentation.
- Fielding, R. T. 2000. „Architectural Styles and the Design of Network-based Software Architectures.” PhD dissz., Irvine: University of California.
http://www.ics.uci.edu/~fielding/pubs/dissertation/fielding_dissertation.pdf. (megtekintve: 2017-11-01)
- Freeman, Robert R., és Pauline Atherton. 1967. *File organization and search strategy using the universal decimal classification in mechanized reference retrieval systems. Report no.AIP/UDC-5*. New York: American Institute of Physics.
- Gnoli, Claudio. 2011. „Facets in UDC: a review of current situation.” *Extensions and Corrections to the UDC* 33: 19-36. <http://arizona.openrepository.com/arizona/handle/10150/236511>. (megtekintve: 2017-11-01)
- „The JavaScript Object Notation (JSON) Data Interchange Format. RFC 7159.” Internet Engineering Task Force (IETF). <https://tools.ietf.org/html/rfc7159> (megtekintve: 2017-11-01)
- Levinson, Meridith. 2007. „Software as a Service (SaaS) Definition and Solutions.” *CIO Digital Magazine*. <https://www.cio.com/article/2439006/web-services/software-as-a-service--saas--definition-and-solutions.html>
- Lula, Paweł, és Urszula Cieraszewska. 2017. „Similarity measurement between UDC classmarks and its application.” In *Faceted classification today: theory, technology and end users: proceedings of the International UDC Seminar 2017, London (UK), 14-15 September*, edited by Aida Slavić, and Claudio Gnoli, 219-239. Würzburg: Ergon Verlag.
- „Major Changes to the UDC 1993-2013.” UDCC (UDC Consortium).
http://www.udcc.org/index.php/site/page?view=major_revisions. (megtekintve: 2017-11-01)
- Mándy Gábor. 2011. „Régi ETO, új ETO.” *Könyvtári Figyelő* 57 (2): 317-332.
<http://ki.oszk.hu/kf/2011/07/regi-eto-uj-eto/>. (megtekintve: 2017-11-01)
- Mándy Gábor. 2013. „A posztkoordináció esélyei az ETO-ban.” *Könyvtári figyelő* 59: 65-84.
http://epa.oszk.hu/00100/00143/00086/pdf/EPA00143_konyvtari_figyelo_2013_1_065-083.pdf. (megtekintve: 2017-11-01)
- Országos Műszaki Információs Központ és Könyvtár (OMIKK), és Országos Széchényi Könyvtár Könyvtártudományi és Módszertani Központ (OSZK-KMK). 1990. *Egyetemes tizedes osztályozás. Rövidített kiadás. 1. kötet Táblázatok. FID Publication 691*. Budapest: OMIKK, OSZK-KMK.
- Országos Széchényi Könyvtár Könyvtári Intézet (OSZK KI). 2005. *Egyetemes tizedes osztályozás. 1. kötet Táblázatok 1-2. Részt. UDCC Publication P057*. Budapest: Országos Széchényi Könyvtár Könyvtári Intézet.

- Pálvölgyi Mihály. 1998. *Információfeldolgozás, információkereső nyelvek*. Szombathely: Berzsényi Dániel Tanárképző Főiskola.
- Pethes Iván, Tolnai György, Turcsányi Tamás, és Varga Jánosné. 1976. „ETO-jelzetek számítógépes rendezése.” *Könyvtári figyelő* 22 (1-2): 66-87.
http://epa.oszk.hu/00100/00143/00228/pdf/EPA00143_konyvtari_figyelo_1976_01-02_066-087.pdf. (megtekintve: 2017-11-01)
- Piros Attila. 2014a. „A different approach to Universal Decimal Classification in a Mechanized Retrieval System.” In *Proceedings of the 9th International Conference on Applied Informatics. Eger, Hungary, January 29–February 1, 2014.*, 29-39. Eger: Debreceni Egyetem, Eszterházy Károly Főiskola.doi:10.14794/ICAI.9.2014.2.29
- Piros Attila. 2014b. „Az ETO alapú számítógépes információkeresés egy új megközelítése.” *Könyvtári Figyelő* 60 (2): 156-166.
http://epa.oszk.hu/00100/00143/00091/pdf/EPA00143_konyvtari_figyelo_2014_2_156-166.pdf. (megtekintve: 2017-11-01)
- Piros Attila. 2015a. „Automatic interpretation of complex UDC numbers: towards support for library systems.” In *Classification & authority control: expanding resource discovery: proceedings of the International UDC Seminar, 29-30 October 2015, Lisbon, Portugal*, szerk. Aida Slavić, és Maria Inês Cordeiro, 177-193. Würzburg: Ergon Verlag. „Revised UDC tables 2000”. *Extensions and Corrections to the UDC 22*.
- Piros Attila. 2015b. „Az ETO-jelzetek automatikus elemzésének kérdései.” *Tudományos és műszaki tájékoztatás* 62 (4): 127-142. http://tmt-archive.omikk.bme.hu/show_news.html?id=6025&issue_id=569. (megtekintve: 2017-11-01)
- Piros Attila. 2017. „The thought behind the symbol: about the automatic interpretation and representation of UDC numbers.” In *Faceted classification today: theory, technology and end users: proceedings of the International UDC Seminar 2017, London (UK), 14-15 September*, edited by Aida Slavić, and Claudio Gnoli, 203-218. Würzburg: Ergon Verlag.
- „Punctuation and intercalation.” Notes & Queries. UDC Consortium (UDCC).
<http://www.udcc.org/notes.htm#bitsymbol>. (megtekintve: 2017-11-01)
- Riesthuis, Gerhard. 1997. „Decomposition of complex UDC notations.” *Extensions and Corrections to the UDC 19*: 13-19.
- Riesthuis, Gerhard. 1998. „Zoeken met woorden: hergebruik van onderwerpsontsluiting.” PhD diss., Amsterdam: Leerstoelgroep Boek-, Archief- en Informatiewetenschap.
- Riesthuis, Gerhard. 1999. „Searching with words: re-use of subject indexing.” *Extensions & Corrections to the UDC 21*: 24-32.
- Salah, Almila Akdag, Cheng Gao, Krzysztof Suchecki, és Andrea Scharnhorst. 2011. „Need to categorize: A comparative look at the categories of the Universal Decimal Classification system (UDC) and Wikipedia.” In *Paper for High Throughput Humanities - a satellite meeting at the European Conference on Complex Systems 2010; Lisbon, Portugal*.
<https://arxiv.org/abs/1105.5912>. (megtekintve: 2017-11-01)
- Alenka Šauperl. 2015. „UDC as a standardisation method for providing titles of documents.” In *Classification & Authority Control: Expanding Resource Discovery : Proceedings of the International UDC Seminar 2015, 29-30 October 2015, Lisbon, Portugal*, szerk. Aida Slavić, és Maria Inês Cordeiro, 235. Würzburg: Ergon Verlag.

- Slavić-Overfield, Aida. 2005. „Classification Management and use in a Networked Environment: The Case of the Universal Decimal Classification.” PhD diss., London: University College London. http://discovery.ucl.ac.uk/1334914/1/AidaSlavicOverfield_thesis_UCL2005.pdf. (megtekintve: 2017-11-01)
- Slavić, Aida, és Antoine Isaac. 2009. „Identifying management issues in networked KOS: examples from classification schemes.” NKOS 2009 workshop, ECDL conference. Corfu, Greece. https://at-web1.comp.glam.ac.uk/pages/research/hypermedia/nkos/nkos2009/presentations/Slavić_isaac_NKOS2009_06.pdf. (megtekintve: 2017-11-01)
- Slavić, Aida, és Sylvie Davies. 2017. „Facet Analysis in UDC: questions of structure, functionality and data formality.” In *Faceted classification today: theory, technology and end users: proceedings of the International UDC Seminar 2017, London (UK), 14-15 September*, szerk. Aida Slavić, és Claudio Gnoli, 181-200. Würzburg: Ergon Verlag.
- Smiraglia, Richard, Andrea Scharnhorst, Almila Akdag Salah, és Cheng Gao. 2013. „UDC in action.” In *Classification and visualization: interfaces to knowledge: proceedings of the International UDC Seminar, 24-25 October 2013, The Hague, The Netherlands*, szerk. Aida Slavic, Almila Akdag Salah, és Sylvie Davies. Würzburg: Ergon Verlag.
- Strachan, David P. 2000. „A Chance to make UDC Notation More Computer-Friendly.” *Extensions and Corrections to the UDC 22*: 43-44.
- „The European Library Open Dataset.” The European Library. <http://www.theeuropeanlibrary.org/tel4/access/data/opendata>. (megtekintve: 2017-11-01)
- „UDC Licences.” UDCC (UDC Consortium). <http://udcc.org/index.php/site/page?view=licences>. (megtekintve: 2017-01-01)
- „UDC Notation.” UDCC (UDC Consortium). http://www.udcc.org/index.php/site/page?view=about_notation. (megtekintve: 2017-01-01)
- „UDC Online.” UDCC (UDC Consortium). <http://www.udc-hub.com/>. (megtekintve: 2017-11-01)
- „UDC Summary. Multilingual Universal Decimal Classification Summary. UDCC Publication No. 088.” UDCC (UDC Consortium). <http://www.udcc.org/udcsummary/php/index.php>. (megtekintve: 2017-11-01)
- Wimmer, F. 2000. „The Apostrophe in UDC: A critical examination of an unwelcome development.” *Extensions and Corrections to the UDC 22*: 19-27.

Összefoglalás

Az *Egyetemes Tizedes Osztályozás (ETO)*, amely a világ első facettás és analitikus-szintetikus osztályozási sémája, az egyik legrégebben használt tudásszervező rendszer a világon. A rendszer analitikus-szintetikus jellege egyfelől *számos lehetőséget biztosít a tartalmak pontosabb kifejezésére* új fogalmak képzésével, másrészt ez a gazdag kifejezőképesség azt is jelenti, hogy *nincs mód a fogalmak kapcsolatainak meghatározására az indexelési fázist megelőzően*. Ez a reprezentációs kiszámíthatatlanság *nehézségeket jelent* a meglévő bibliográfiai adatbázisok tárgyi metaadatainak hatékony használatában is.

Az elmúlt hetven évben számos kutatás foglalkozott az ETO számítógépes rendszerekben történő alkalmazási lehetőségeivel. Az 1960-as években a kutatók ennek *a gépesítésnek több aspektusát vizsgálták* az ETO-jelzetek automatikus kezelésétől a tárgyi authority fájlok alkalmazásáig. Az 1990-es években a *UDC Master Reference File (MRF)* létrehozása újabb lökést adott a kutatásoknak, melyek újra megvizsgálták a részekre bontás és a részek természetes nyelvű visszakereshetőségének kérdéseit (ezúttal már az MRF alapján), az ETO felhasználási lehetőségeit többnyelvű teauruszokban és más kontrollált szótárakban.

Bár vannak az *ETO-t támogató rendszerek*, ezek általában *egy adott szemszögből* közelítik meg az osztályozás használatát, amely csak egy adott funkcionalitást támogat. Ha egy rendszer elsődleges célja *a zaj elkerülése*, akkor általában csökken a *visszahívás* mértéke. Ha a cél a keresőkérdések könnyebb összeállítása, általában nehezebb eldönteni az eredményként kapott rekordok relevanciáját. *A kontroll* növelése segíti indexelési gyakorlat egységesítésében és a tárgyi böngészés *hatékonyabbá tételében*, de *korlátozza azt a flexibilitást*, ami a rendszer egyik legnagyobb erősségének számít. Másrészt a flexibilitás *csökkenti a reprezentációs megbízhatóságot*, ez pedig információvesztéshez vezethet a visszakeresés során. Ezek az ellentmondások magának *az osztályozásnak a komplex jellegéből adódnak*, ezért általában nem oldhatók fel egyetlen rendszeren belül.

Kutatásom során *arra törekedtem, hogy a különböző megközelítéseket közelebb hozzam egymáshoz annak érdekében, hogy az ETO-ban rejlő lehetőségeket kompromisszumok és korlátozások nélkül kiaknázzhatóvá tegyem*. Pontosabban meghatározva, **a kutatás elsődleges célja** annak megmutatása volt, *hogy lehetséges olyan módszerek kidolgozása, melyek egyaránt támogatják az ETO-jelzetek összeállítását, visszakeresését és elemzését a rendszer teljes szintaktikai gazdagságát figyelembe véve*. A megoldás alapja **az ETO számok eltárolása olyan formában, amely megtartja és feldolgozhatóvá teszi az azok szintaktikai struktúrája által tárolt információt** is. A kutatás céljai közé tartozott annak megmutatása is, hogy lehetséges olyan **jelzetelemző algoritmus kidolgozása és megvalósítása**, mely *az ETO számokat a kompozícionalitás elvét szem előtt tartva, azok teljes szintaktikai struktúráját figyelembe véve elemzi*.

A disszertációmban bemutatott megoldás alapját **egy új XML alapú, platformfüggetlen, géppel olvasható formátum** képezi, amely *alkalmas akár a legkülönlegesebb, komplikált ETO számok ábrázolására is, azok teljes szintaktikai struktúrájára – a részjelzeteikre és azok kapcsolódási módjára – vonatkozó információ megőrzésével*, illetve egy olyan **alkalmazás, amely képes az ETO-jelzeteket közvetlenül a kidolgozott formátumra konvertálni**.

Az irodalmi áttekintést követően a *formátumot* a disszertáció *ötödik fejezetében*, míg az *interpreter működését, illetve továbbfejlesztési és alkalmazási lehetőségeit* a *hatodik fejezetben* mutatom be részletesen. A *séma definíció* és a *program* egyaránt a **<http://piros.udc-interpreter.hu/>** címen érhető el.

A **program legfontosabb újdonsága** a *teljességre való törekvés* – kimondott célja, hogy kezelje a legspecifikusabb jelzetszerkesztési megoldásokat is és megőrizze a jelzetek teljes szintaktikai struktúráját – és *hogyan kezeli az osztályozási rendszer változásait* is.

Az ETO számok XML reprezentációja tovább konvertálható más formátumokra is, a szabványos leíró, adatcsere és forrásleíró formátumoktól az ETO-jelzetek javasolt hozzáférési pontjainak listájáig. A kutatás során ilyen **konverziós metódusok is kidolgozásra kerültek**. A megvalósított program **online szolgáltatásként** is elérhető, így különböző kimenetei más alkalmazások számára is elérhetők, így az algoritmus akár meglévő akár a jövőben kifejlesztendő rendszerek számára is könnyen integrálható.

Az ETO számok *XML formátumban történő ábrázolásának* az egyik legnagyobb **előnye**, hogy szemben a jelzetek eredeti formában történő eltárolásával, egy ilyen formátum átlátható és a programozási nyelvek által jól támogatott. Következésképpen *a további gépi feldolgozás és konverzió nem igényel speciális algoritmusokat és nagyobb programozási munkát*. Ez által pedig növelhető a hatékonysága minden olyan eljárásnak, amely használja az ETO számok szintaktikai struktúrájára és elemeire vonatkozó információt. Ilyen eljárások lehetnek például a *kvantitatív kutatások módszerei*, az *ETO-jelzetek hasonlóságának megállapítása*, vagy a hatékonyabb inkluzív keresést és böngészést célzó algoritmusok.

A generált XML olyan formátumok és eljárások alapjaként is szolgálhat, amelyek elősegítik intelligens osztályozási interfészek fejlesztését. A jelzet olyan formába való konvertálása, amely tartalmazza az ETO számok szabványos rendezéséhez szükséges jelölőket, megvalósítás alatt van. Ezen jelzetek és az XML alapján lehetőség nyílik olyan KWIC-indexek építésére, amelyek lehetővé teszik a jelzetelemek szerinti böngészést azok környezetét is figyelembe véve. A jelzetek elemeinek permutálása az ETO szabályait szem előtt tartva megsokszorozhatja a jelzetek hozzáférési pontjainak számát, és elvezethet olyan algoritmusok kidolgozásához, amelyek a jelzeteket automatikusan változtatják a böngészés során a kereső feltételezett kognitív státusza alapján.

A feldolgozás eredménye nem csak könyvtári katalógusokban használható, hanem - a prekoordinált jelzetek kapcsolatainak feltérképezésén keresztül az egyszerű ETO-számok ontológiájának ilyen formában történő kiterjesztésén keresztül - a *szemantikus webben*, illetve *internetes keresőkben* is.

A program és a formátum közvetlen alkalmazása mellett a kutatás első fázisa során felhalmozott tapasztalatok kiértékelése olyan elméleti vizsgálatokra ad lehetőséget, mint *az ETO formális nyelvtanának felírása*, vagy annak áttekintése, hogy az utóbbi évek revíziói mennyiben segítették elő az ETO gépesítését.

Summary

As the first analytico-synthetic classification, *Universal Decimal Classification (UDC)* is one of the oldest knowledge organization systems of the world. Its analytico-synthetic approach provides the possibility for indexing documents more precisely and specifically by *composing new subjects from the existing ones*. However, that also means, that *one cannot anticipate the possible combinations of the subjects*, and the relationships of the newly created subjects cannot be explained before they would be created, during the indexing phase. The un-determinism of the classification codes may also *bottleneck* the usage of the already generated subject keywords effectively after having converted existing collections.

Over the last seven decades there *have been many research projects conducted* on the efficient accomplishment of the hierarchical, faceted, analitico-synthetic nature of Universal Decimal Classification in computer systems. In the Sixties many researchers investigated the possibilities of building UDC based indexes or the automatic handling of UDC notations as well as utilizing classification authority files. In the Nineties the publication of the *UDC Master Reference File* encouraged research projects, for example: decomposition algorithms to parse synthesized numbers and join their parts to the descriptions contained by MRF records; investigation of using multilingual thesauri based on UDC; utilizing UDC in comprehensive controlled vocabularies.

There are software systems which are able to apply UDC on an improved level, however, *they approach the employment of the system from one specific aspect*, that usually supports only one given functionality. If a program is tuned to avoid *noise*, the level of *recall* will decrease. If the goal is to compile queries easier, then usually it is harder to make decisions about the relevance of the results. The *improved control* makes unifying indexing process and efficient subject browsing possible, however it *restricts the flexibility*, which is one of the biggest advantages of the classification. On the other hand, the *flexibility may reduce the consistency in the representations*, which will result in *loosing information* during the retrieval process. The above conflicts usually cannot be resolved inside a system, although the language carries the potential necessary to that, because these *problems come from the complex nature of the system*.

In my research *I would like to bring the different approaches closer* in order to make full use of every possibility of UDC without any restrictions as well as requiring the indexer or the searcher to utilize any special, system related knowledge in addition to information retrieval competencies and knowledge of the rules and schedules (either the numbers or the descriptions in a natural language) of UDC. **The primary goal of my research is to prove that it is possible to support building, retrieving, and analyzing UDC numbers without compromises, by taking the whole syntactic richness of the scheme by storing the UDC numbers reserving the meaning of pre-coordination.** The research has also included **the implementation of a software that parses UDC classmarks.**

The approach presented in my dissertation provides **a new XML-based platform-independent, machine processable data format**, *which can be employed to describe even the most specific and complicated UDC numbers without losing any information regarding to their whole syntactical structure* - the parts of those and the way how those have been constructed -, together with **the application which is able to convert the existing notations into such format.**

After having reviewed the literature I explained the newly *designed format* in detail in the *fifth chapter* of my dissertation, while *the basic principles how the interpreter works, its usage and a few possible applications* of that are outlined in the *sixth chapter*. Both the XML Schema Definition that defines the format and the software service are available at: <http://piros.udc-interpreter.hu/>.

The **novelty of such a program** is due to *its ability to identify even the most specific solutions* for building UDC classmarks *by retaining all of the syntactical information* stored by the number, and to *react to the changes of the schema*.

That format is applicable as a base of further conversions into different formats, such as standard mark-up, data-exchange and resource description formats or simple lists of recommended entry points of a UDC number. The research also involves **implementing conversion methods** to provide outputs that can be employed by other software directly and, **as a service**, make them available for other software. This means, that the algorithm, in principle, can be utilized both in the existing and the future systems to analyze UDC numbers.

One of the biggest advantages of storing a UDC number in an XML format is that this format, in contrast to the UDC notations themselves, is transparent and well supported by programming languages. Thus, *its further automatic analysis and conversion would not require special algorithms and programing effort*. This may improve the efficiency of methods applying information regarding the syntactic structure and the elements of UDC numbers, including *quantitative studies, similarity measurements* between composite subjects or to develop *improved algorithms to search and browse* composite subject descriptions inclusively.

The generated XML may also serve as the bases for the development of outputs and methods that can support *implementing intelligent classification interfaces*. Generating the notation that contains proper tags to arrange the numbers in the correct filing order is already implemented. Building, for instance, a KWIC index based on these notations may render it possible to browse the composite numbers by their components, by taking their context into account. Permuting the elements by abiding by the rules of UDC is also possible, based on the representation that reveals the whole inner structure of the numbers: this may multiply the number of the access points to the concepts. The final goal would be to change the notations of the numbers automatically and to display them in a form that fits to the supposed target of the searching process and the cognitive status of the searcher during subject browsing.

The outputs of the service may support not only retrieving information more effectively in library catalogs, but also to discover the relations between pre-coordinated UDC numbers and to expand the ontology of simple UDC numbers with the compound concepts, in order to find documents having been indexed by using UDC on the *semantic web* and in *Internet based search engines*.

In addition to applying the format and the program, investigating into the experience accumulated while automatically representing and interpreting the concepts gives us a high chance of defining the formal grammar of UDC or examining how the revisions and appropriate changes performed in recent years have helped in handling UDC notations.

Köszönetnyilvánítás

Mindenekelőtt szeretnék köszönetet mondani a feleségemnek, Krisztinának és kislányomnak, Eszternek a támogatásért, melyet az egész kutatómunkám során nyújtottak, illetve családom többi tagjának és valamennyi barátomnak a folyamatos bátorításért.

Külön szeretném megköszönni témavezetőm Dr. Boda István értékes és konstruktív javaslatait és támogatását, valamint Dr. Aida Slavić szívből jövő segítségét, a rengeteg értékes és hasznos információt, melyet a rendelkezésemre bocsátott.

Végül köszönettel tartozom Daniel Benediktssonnak az információkereséssel kapcsolatos kurzusokon átadott ismeretekért és szemléletért, illetve Jonathan Wildnak a publikációim angol nyelvi lektorálásáért és a konferenciákra való felkészüléshez nyújtott segítségért.

Függelék

A szerző publikációi

Referált publikációk

Piros Attila. 2018. „A gondolat a szimbólum mögött: Az ETO-jelzetek automatikus interpretálásáról és elemzéséről.” *Tudományos és Műszaki Tájékoztatás (TMT)* 65 (2). [nyelvi változat]

Piros Attila. 2017. „The thought behind the symbol: about the automatic interpretation and representation of UDC numbers.” *Knowledge Organization (KO)* 44 (6): 416-424.
doi:10.5771/0943-7444-2017-6-416

Indexelve: Social Sciences Citation Index, Web of Science, Information Science Abstracts, INSPEC, Library and Information Science Abstracts (LISA), Library, Information Science & Technology Abstracts (EBSCO), Library Literature and Information Science (Wilson), PASCAL, Referativnyi Zhurnal Informatika, és Sociological Abstracts

Impact Factor (2016): 0.831.

Piros Attila. 2015. „Automatic interpretation of complex UDC numbers: towards support for library systems.” In *Classification & authority control: expanding resource discovery: proceedings of the International UDC Seminar, 29-30 October 2015, Lisbon, Portugal*, szerk. Aida Slavić, és Maria Inês Cordeiro, 177-193. Würzburg: Ergon Verlag.

Indexelve: Literatur zur Informationserschließung (TH Köln, Institut für Informationswissenschaft)

Piros Attila. 2015. „Az ETO-jelzetek automatikus elemzésének kérdései.” *Tudományos és Műszaki Tájékoztatás (TMT)* 62 (4): 127-142. http://epa.oszk.hu/03000/03071/00085/pdf/EPA03071_tmt_2015_04_127-142.pdf.

Piros Attila. 2014. „Az ETO alapú számítógépes információkeresés egy új megközelítése.” *Könyvtári Figyelő* 60 (2): 156-166. http://epa.oszk.hu/00100/00143/00091/pdf/EPA00143_konyvtari_figyelo_2014_2_156-166.pdf.

Piros Attila. 2014. „A different approach to Universal Decimal Classification in a Mechanized Retrieval System.” In *Proceedings of the 9th International Conference on Applied Informatics. Eger, Hungary, January 29–February 1, 2014.*, 29-39. Eger: Debreceni Egyetem, Eszterházy Károly Főiskola. doi:10.14794/ICA1.9.2014.2.29

Indexelve: Zentral Blatt - IOPort

Lektorált publikációk

Piros Attila. 2018. „New automatic interpreter for complex UDC numbers.” *Extensions and Corrections to the UDC* 36-37 (2014-2015). [Megjelenés alatt]

Piros Attila. 2017. „The thought behind the symbol: about the automatic interpretation and representation of UDC numbers.” In *Faceted classification today: theory, technology and end users: proceedings of the International UDC Seminar 2017, London (UK), 14-15 September*, szerk. Aida Slavić, and Claudio Gnoli, 203-218. Würzburg: Ergon Verlag. [utánközlés]

Piros Attila. 2014. „Aspects of Automatic Information Retrieval Using Universal Decimal Classification.” In *Tavaszi Szél 2014: Konferenciakötet*, szerk. Csiszár Imre, és dr. Kőmíves Péter

Miklós, 223-231. Debrecen: Doktoranduszok országos szövetsége. <http://www.dosz.hu/dokumentumfile/TSZ-VI-kotet.pdf>.

Egyéb publikációk

Piros Attila. 2017. „A facettás osztályozás napjainkban: elmélet, technológia és a végfelhasználók.” *Könyv, Könyvtár, Könyvtáros* 26 (11): 22-32.

Piros Attila. 2015. „A New Approach to Universal Decimal Classification as the Indexing Language for a Mechanized Information Retrieval System.” In *Agría Media 2014: XI. Információtechnikai és Oktatóstechnológiai Konferencia és Kiállítás. Eszterházy Károly Főiskola - 2014. október 8-10.*, szerk. Dr. Nádasi András, 228-235. Eger: Eszterházy Károly Főiskola Médiainformatikai Intézet.

Konferencia előadások

Piros Attila. „The thought behind the symbol: about the automatic interpretation and representation of UDC numbers.” *Faceted classification today: theory, technology and end users (Nemzetközi ETO Szeminárium 2017), London (UK), 2017. szeptember 14-15.*
<http://seminar.udcc.org/2017/programme.php>. (megtekintve: 2018-01-05)

Piros Attila. „Automatic interpretation of complex UDC numbers: towards support for library systems.” *Classification & Authority Control: Expanding Resource Discovery (Nemzetközi ETO Szeminárium 2015), Lisszabon, 2015. október 29-30.*
<http://seminar.udcc.org/2015/programme.php>. (megtekintve: 2018-01-05)

Piros Attila. „A New Approach to Universal Decimal Classification as the Indexing Language for a Mechanized Information Retrieval System.” *Agría Media 2014, Eger, 2014. október 8-10.*

Piros Attila. „Az ETO jelzetek automatikus elemzésének kérdései.” *Library 2.014, 2014. október 8-9.*
<https://sas.illuminate.com/site/external/recording/playback/link/table/dropin?sid=2008350&uid=D.1CC49A03B626697C83E33699C00236>. (megtekintve: 2018-01-05)

Piros Attila. „Aspects of Automatic Information Retrieval Using Universal Decimal Classification.” *Tavaszi Szél Konferencia 2014, Debrecen, 2014. március 21-23.*

Piros Attila. „A Different Approach to UDC in a Mechanized Retrieval System.” *ICAI 2014, Eger, 2014. január 29-február 1.* <http://icaiektf.hu/?p=abstracts>. (megtekintve: 2018-01-05)