

Egyetemi doktori (PhD) értekezés tézisei

Az ETO-jelzetek automatikus interpretálásának és elemzésének kérdései

Piros Attila

Témavezető: Dr. Boda István Károly



DEBRECENI EGYETEM

Matematika- és Számítástudományok Doktori Iskola

Debrecen, 2018.

Tartalom

Bevezetés	1
1 Osztályozásméleti alapfogalmak	3
2 Az Egyetemes Tizedes Osztályozás	3
3 Az ETO mint analitikus-szintetikus osztályozási rendszer	4
4 Az ETO gépesítése	4
5 Az ETO számok szerkezetének leírása	7
5.1 Az ETO-jelzetek szintaktikai struktúrája	7
5.2 Az ETO jelzetek szintaktikai struktúrájának leírása XML használatával	7
6 Az összetett ETO számok automatikus interpretálása	9
6.1 Az interpretálás alapelvei	9
6.2 Az interpreter működésének alapjai	10
6.3 További fejlesztési irányok és kutatási tervek	11
Introduction	13
1 Basic concepts of classification science	15
2 The Universal Decimal Classification	15
3 The UDC as an Analythico-Synthetic Classification Scheme	16
4 The Automatization of UDC Numbers	16
5 The representation of UDC notations	19
5.1 The syntactic structure of UDC numbers	19
5.2 Using XML to describe the structure of the UDC numbers	19
6 The automatic interpretation of pre-coordinated UDC numbers	21
6.1 The underlying principles of the interpreter	21
6.2 The basics of how the interpreter works	22
6.3 Future research directions and plans	23
Irodalomjegyzék / Bibliography	25

Bevezetés

Az *Egyetemes Tizedes Osztályozás (ETO)* egyike a világ legnagyobb és legelterjedtebb osztályozási rendszereinek. Harminckilenc nyelven jelentek meg különböző kiadásai, melyeket legalább százhuszonnégy országban használnak. Jelenleg csak Európában több, mint 140 ezer ETO-val indexelt gyűjtemény található.

Ez az elterjedtség részben annak köszönhető, hogy száz évvel ezelőtti feltalálása óta *folyamatos fejlesztés alatt áll* annak érdekében, hogy a benne foglalt tudásanyag mindig a lehető legpontosabban tükrözze a tudományok aktuális állását. Népszerűségének másik kulcsa *analitikus-szintetikus jellege*: miközben hierarchikusan szervezett táblázatai az emberi tudás jelentős részét felölelik, segéd táblázatai és jelzetkapcsolási megoldásai lehetővé teszik, hogy a felhasználó új fogalmakat alkothasson a már meglévőkből.

Ugyanakkor erősségei és előnyei ellenére a rendszer *használata és fenntartása mégis gyakori viták tárgya*, nem utolsósorban azért, mert a könyvtári és információkereső *szoftverek nem használják ki teljes mértékben a benne rejlő lehetőségeket*.

Bár vannak az ETO-t hatékonyan támogató rendszerek, ezek *általában egy adott aspektusból közelítik meg az osztályozás alkalmazását*, amely csak egy adott funkcionalitást támogat. Ha egy rendszer elsődleges célja a *zaj* elkerülése, akkor azzal általában csökken a *visszahívás* mértéke. Ha a cél a keresőkérdések könnyebb összeállítása, általában nehezebb eldönteni az eredményként kapott rekordok relevanciáját. A *kontroll* növelése segít az indexelési gyakorlat egységesítésében és a tárgyi böngészés hatékonyabbá tételében, de korlátozza azt a *flexibilitást*, ami a rendszer egyik legnagyobb erősségének számít. Másrészt éppen a flexibilitás csökkenti a *reprezentációs kiszámíthatóságot*, ami *információvesztéshez* vezethet a visszakeresés során. Ezek az ellentmondások magának az osztályozásnak a *komplex jellegéből* adódnak, ezért általában nem oldhatók fel egyetlen rendszeren belül.

Kutatásom elsődleges célja megmutatni, 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 olyan formában történő eltárolása, 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**, mely az ETO számokat a kompozicionalitás elvét szem előtt tartva, azok teljes szintaktikai struktúráját figyelembe véve elemzi. Az algoritmus alapján **megvalósított alkalmazás** könnyen alkalmazható ETO-jelzetek elemzésére, akár visszamenőlegesen, már meglévő gyűjteményekben is.

A disszertáció első fele az ETO-val és annak gépesítésével kapcsolatos **szakirodalmat** tekinti át. Az *első fejezet* azokat az *osztályozáselméleti alapfogalmakat* és kifejezéseket vezeti be, amelyekre a dolgozat későbbi részei támaszkodnak. A *második fejezet* célja az ETO történetének és jelentőségének bemutatása. A *harmadik fejezet* magáról az osztályozásról szól, kiemelve azokat a *jellemzőit*, amelyek *leginkább befolyásolják az automatizálás kérdéseit*. Ezek a kérdések részletesebben a *negyedik fejezetben* kerülnek kifejtésre.

A disszertáció második fele **saját eredményeket** tartalmaz. Először részletesen bemutatok egy **formátumot, mely alkalmas az ETO számok géppel olvasható formában történő eltárolására, a részeikre és azok kapcsolódási módjára vonatkozó valamennyi információ megtartásával**. A következő fejezetben bemutatok az általam fejlesztett **interpreter szoftvert, amely képes a legbonyolultabb ETO-jelzetek elemzésére és a megelőző fejezetben leírt formátumra történő konvertálására is**. A szoftver képes a feldolgozás eredményének további feldolgozására is, a különböző megközelítéseket támogató formátumokra történő konverzió érdekében. Legvégül röviden áttekintem az elért **eredmények felhasználási lehetőségeit** és a további **lehetséges kutatási irányokat**.

1 Osztályozáselméleti alapfogalmak

Az **osztályozás** általános értelemben véve bizonyos objektumok adott szempontok szerinti besorolását jelenti. Az információkeresési szakirodalomban osztályozáson tartalmak valamilyen (általában mesterséges nyelvű, hierarchikus) *információs nyelven* történő kifejezését értjük.

Az osztályozási rendszernek a fogalmak leírására használható szemantikai egységek mellett tükröznie kell azok **szemantikai kapcsolatait** is. Ezek a kapcsolatok adják meg a rendszer szerkezetét és teszik lehetővé a benne történő tájékozódást. Az osztályozási rendszer harmadik összetevője a **szintaktikai kapcsolatok** leírására szolgáló szabályrendszer. Ezek azért különösen fontosak, mert az osztályozási jelzetnek nem csak a leírt szöveg ismérveit kell tartalmaznia, hanem azok *fogalmi viszonyait* is tükröznie kell annak érdekében, hogy a tartalom minél pontosabb kifejezésével segítse a megfelelő relevanciadöntések meghozatalát.

Attól függően, hogy egy adott gyűjtemény vagy téma szempontjából egy reláció állandósult, gyakran előforduló, illetve eseti kapcsolatokat jelöl szokás **paradigmatikus** és **szintagmatikus relációkról** beszélni. A paradigmatis relációkkal kifejezett fogalmak felvehetők az információkeresési szótárba, megfelelő kontrollal a számítógépes kezelésük is hatékonyan megoldható. Ezzel szemben a szintagmatikus szintaktikai kapcsolatok automatikus kezelése gyakran problémákat okoz.

A rendszer alapvető struktúrája alapján az ETO-t általában az **analitikus-szintetikus** rendszerek közé szokás sorolni. Az ilyen rendszerek jellemzője, hogy az általuk kiépített nagy és mély *fogalmi hierarchia* mellett lehetővé teszik a fogalmak további pontosítását, illetve összekapcsolását is a tartalmak minél részletesebb leírása érdekében. Emiatt a komplex jelleg miatt a szintaktikai relációk kezelése kulcskérdés az osztályozás gépesítése során.

2 Az Egyetemes Tizedes Osztályozás

Az ETO egyike a három klasszikus, legrégebb óta és legszélesebb körben használt osztályozási rendszernek. A tizenkilencedik század végén Dewey tizedes osztályozásának kibővítésével hozták létre, egy, a világ dokumentumtermését számba vevő katalógus számára - a bővítések eredményeként *a világ első facettás, analitikus-szintetikus osztályozási rendszereként*. Bár elsősorban dokumentációs céllal jött létre, már a kezdetektől használták könyvtári osztályozásként is. Emellett rugalmassága és terminológiai gazdagsága miatt gyakran szolgál alapul szakterületi osztályozási rendszereknek és tezaurszoknak is.

Az ETO-t a Nemzetközi Dokumentációs Szövetség (FID) gondozta egészen az 1990-es évekig, amikor karbantartásának feladatát és a forgalmazás jogait átadta egy nonprofit de önfenntartó konzorciumnak.

A hetvenes években az igények és az osztályozástudomány változásai, illetve a rendszer számítógépesítésével kapcsolatos tapasztalatok alapján jelentős átszervezésre került sor, amely változásokat hozott a jelzetszerkesztés szemléletében is. Az elmúlt harminc évben pedig a rendszer tudatos *facettás átszervezése* került előtérbe. Ezzel egy időben kiadták a korábbi középkiadásoknak megfelelő méretű **Master Reference File**-t, amely az ETO központi anyagát tartalmazó, a kiadások és fejlesztések alapjául szolgáló, kb. 70.000 rekordból álló adatbázis. Az MRF felhasználásához a megfelelő *licenz* szükséges, illetve hét nyelven online is elérhető - szintén térítés ellenében.

Az ETO történelme 120 éve alatt *a világ második legelterjedtebb osztályozási rendszerévé* vált. A legutóbbi felmérések alapján legalább 124 országban használják, csak Európában legalább 28 ország 144.000 könyvtára alkalmazza. Legalább 39 nyelven jelentek meg különböző ETO kiadások, online rövidített verziója pedig ötvenhét nyelvre lefordítva érhető el.

Magyarországon az ETO 120 éve ismert és használt, a huszadik század második felére pedig egyértelműen az elsődleges könyvtári osztályozási rendszerré vált. Bár használatát viták övezik, a

legutóbbi, 2016-ban készült felmérés alapján még mindig aktívan osztályoznak vele a különböző típusú könyvtárak többségében.

A legutóbbi nyomtatott magyar nyelvű kiadás 2005-ben jelent meg, jelenleg pedig előkészítés alatt van a legújabb MRF magyar nyelvű fordításának integrálása a fejlesztés alatt lévő új nemzeti névtérrendszerbe.

3 Az ETO mint analitikus-szintetikus osztályozási rendszer

Ahogy azt már korábban említettük, az *ETO analitikus-szintetikus rendszer*, azaz egyszerre tartalmazza az enumerikus és a facettás osztályozási rendszerek elemeit.

Az enumerikus jelleget elsősorban a mély tartalmi hierarchiát tükröző, egyetemes *fő táblázatok*, illetve a szintén hierarchikusan felépített *általánosan közös alosztási táblázatok* biztosítják. Emellett az ETO számos *korlátozottan közös alosztást* és egyéb lokális jelzetszerkesztési megoldást is biztosít a facetták, tehát egy adott osztályon belül jellemzően előforduló jellemzők kifejezésére. Legvégül, a szintetikus jelleget az *általánosan közös alosztások* biztosítják, amelyek lehetőséget adnak a (fő- és általánosan közös alosztási) táblázati számok, illetve a belőlük képzett tárgykörök további összekapcsolására is, a tartalmak minél pontosabb kifejezése érdekében.

Ez ETO erősen *prekoordinált* rendszer, melynek komplex jellegéből következik, hogy a jelzetek értelmezéséhez kevés az őket alkotó fogalmak figyelembe vétele: a kapcsolódási módok, a relációk típusa, a jelzetelemek sorrendje mind jelentéssel bírhat. Ezek figyelmen kívül hagyása a precízióra és a visszahívásra egyaránt negatívan hathat.

4 Az ETO gépesítése

Ahogy korábban láthattuk, az ETO szintetizálási lehetőségei lehetővé teszik az indexelő számára a bonyolultabb kifejezések hatékony leírását is a minél hatékonyabb visszakeresés érdekében. Ilyenkor a fogalmak prekoordinációja révén egy új fogalom jön létre, amely önálló, új jelentéssel rendelkezik, ahelyett, hogy egyszerűen örökölné a részek jelentését. *Ugyanazok az ETO számok más relációkkal összekapcsolva egy másik tárgykört reprezentálnak.* Ezért a szakirodalomban konszenzus uralkodik azzal kapcsolatban, hogy az adatelemek integritásához és az ETO szemantikájának az értelmezéséhez a kulcs a *jelzetek megfelelő kezelésében* rejlik.

Ez két fontos *követelményt* támaszt az információkereső rendszerekkel szemben:

- a) biztosítani kell az ETO-jelzetek szemantikai komponenseinek az integritását, például lehetővé kell tenni a jelzetelemek és azok leírásai szerinti keresést
- b) az összetett ETO számokat a szintaktikai struktúrájuk megőrzésével, egészként is elérhetővé kell tenni a leírásukkal együtt, hogy részek kapcsolódási módja által hordozott plusz jelentést is figyelembe tudjuk venni.

A korábbi rendszerekben a *jelzetek tárolására* az alábbi megoldások terjedtek el:

- 1) A jelzetek „ahogy vannak”
 - a) egyszerű megközelítés – a jelzetek **karaktorsorozatként** való kezelése
 - b) **KWOC-index** – az ETO-jelzetek részeinek önálló visszakereshetővé tétele
 - c) **KWIC-index** – a jelzetelemek a saját kontextusukon belül kereshetők vissza
- 2) A kombinált jelzetek strukturált adatként való eltárolása
 - a) **authority control** – a jelzetek osztályozási authority állományokban kerülnek eltárolásra, amelynek rekordjai a jelzetek elemei és azok leírásai, illetve a köztük fennálló szemantikai kapcsolatok alapján is hozzáférhetők.

Az első csoport megoldásainak közös jellemzője, hogy a keresés általában nagy *zajjal* jár a szintaktikai relációk hiányos kezelése miatt. Emellett gyakran nagy az *információs veszteség* is a jelzetelemek nem kellő mértékű feltárása, illetve a reprezentációs kiszámíthatatlanság miatt.

Az *authority rekordok* nemcsak az osztályozási jelzetet és annak metaadatait tárolhatják, hanem járulékos szemantikai információt is. Ilyenek lehetnek már említett szemantikai kapcsolatok az osztályok között (például az osztály hierarchikus felső és alsó kapcsolatainak, illetve az egyéb kapcsolódó osztályok jelölésével), teauruszok vagy tárgyszólisták kapcsolódó deskriptorainak a felsorolása, vagy akár másik osztályozási rendszerből származó jelzetek és azok leírásai. A szemantikai információk alapos feltárásának köszönhetően az authority control segítségével a tárgyi keresés és böngészés egyaránt hatékonyan támogatható.

Másrészről az authority control fenntartása meglehetősen *idő- és költségigényes* folyamat és olyan tapasztalatot és szakértelmet igényel, amelynek legtöbb könyvtár nincs birtokában. Ezért az authority control általában csak a jól menedzselt nemzeti és egyetemi könyvtárak gyakorlatára jellemző. A könyvtári hálózatban található bibliográfiai adatok nagy része tehát nincs authority állományokban karbantartva, és messzemenően indokolt lenne a módszer automatizált támogatása. Az authority control további hátránya, hogy *nem alkalmazható a szintagmatikus fogalmak esetében*, melyek kifejezhetősége az analitikus-szintetikus rendszerek legnagyobb előnyei közé tartozik.

5 Az ETO számok szerkezetének leírása

Az ETO jelzetkapcsolási lehetőségeinek elemzése alapján beláthatjuk, hogy szinte *valamennyi ETO szám reprezentálható egy hierarchikus struktúrával*, amelyet a műveletek, illetve az általánosan és korlátozottan közös alosztások precedenciája határoz meg (a kivételek többnyire a jelzetalkotási szabályok megszegéséből adódnak). Az így felépített *fa leírja a jelzet szintaktikai struktúráját*, megőrizve minden fontos információt annak elemeivel, azok kapcsolódási módjával, jelzeten belüli szerepével, sorrendjével kapcsolatban.

A jelzetek ilyen formában történő reprezentációjának nagy előnye, hogy *független a jelzetelemek sorrendjétől*. Ha két szám logikai szerkezete ugyanaz és csak az elemek sorrendjében különböznek, akkor reprezentációjuk is meg fog egyezni. Ugyanakkor a sorrend az elemek tulajdonságaként meg fog jelenni a reprezentációban, hogy lehetővé tegye a jelzetek megkülönböztetését azokban a ritka esetekben, amikor a sorrendnek szemantikai jelentősége van.

5.1 Az ETO-jelzetek szintaktikai struktúrája

Az ETO műveletei között a következő **precedencia sorrendet** állíthatjuk fel:

1. + Összekötés
2. : Mellérendelő viszonyítás
3. :: Sorrendrögzítés
4. [] Csoportképzés
5. ' Jelzetszintézis (a korlátozottan közös alosztásokban)

A fenti műveletek *operandusai* ETO-számok (a prioritási sorrend alapján), kiterjesztéssel alkotott intervallumok lehetnek. A felső három művelet *csoportképzéseket* is összeköthet, ugyanakkor a csoportképzés is tartalmazhatja ezen műveletek bármelyikét, ugyanezen precedencia sorrend alapján. Az általánosan használható alosztási jelek mellett a táblázatok egyes helyein, korlátozottan közös alosztásként alkalmazható az *aposztróf* is jelzetek szintézisére (pl. kémiai elemek vagy politikai pártok és mozgalmak leírása esetén). Arra is van példa, hogy egyes műveletek általánosan bizonyos típusú közös alosztásokon belül is alkalmazhatók.

Az operátorokkal összekapcsolt fogalmak legfeljebb egy fő táblázati számot (vagy intervallumot) tartalmazhatnak a hozzá kapcsolódó nem önálló általánosan vagy korlátozottan közös alosztásokkal, valamint tetszőleges számú önálló általánosan közös alosztást (amelyek tartalmazhatnak táblázati számokat, korlátozottan közös alosztásokat, műveleteket stb. is).

Az ETO *korlátozottan közös alosztásai* számos eltérő megoldást tartalmaznak annak érdekében, hogy a legspeciálisabb jellemzők is leírhatók legyenek velük (pl. tulajdonságok, eljárások, műveletek, technikák, eszközök stb.) amelyeket eltérő módokon is kell kezelni.

5.2 Az ETO jelzetek szintaktikai struktúrájának leírása XML használatával

Mint minden fa, az egyszerű és kombinált ETO számok struktúrája is leírható egy XML dokumentum formájában.

Disszertációmban részletesen bemutatok egy határozottan az ETO számok leírására tervezett **XML séma definíciót**, amely *alkalmas a legösszetettebb jelzetek leírására is*, beleértve nem csupán az általánosan közös, de a legspeciálisabb korlátozottan közös alosztásokat is. A bemutatott séma definíció letölthető a <http://piros.udc-interpreter.hu#xsd> oldalon.

A séma definíció alapján generált XML file *felhasználható bármely ETO szám leírására*, melynek összeállítója szem előtt tartotta az ETO jelzetépítési szabályait, és teljessége miatt tetszőleges további formátumra konvertálható.

Az XSD **komplex típusai** írják le az ETO számok lehetséges összetevőit: a *táblázati számokat*, *alosztási jeleket* vagy *magát a jelzetet*. Például az alábbi típus alkalmazható a formai általánosan közös alosztások leírására:

```
<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>
```

A műveletek olyan komplex típusokként írhatók le, melyeknek elemei a művelet operandusai. Az alábbi típus például az összekötés leírására használható:

```
<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>
```

A séma **egyszerű típusai**, melyek általában a különböző táblázatok számait leíró reguláris kifejezéseket tartalmaznak, elsősorban *validációs célokat* szolgálnak. Az alábbi példa az idő szerinti általánosan közös alosztásokhoz tartozó egyszerű típust mutatja be:

```
<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})?(?)?)?">
    </xsd:pattern>
    <xsd:pattern value="[3-9]((\.)?\d{1,4})(\.\d{1,4})*">
    </xsd:restriction>
  </xsd:simpleType>
```

A séma szerint minden ETO-val kifejezett fogalmat egy komplex típus ír le, és minden XML-nek egy ilyen típusú elemet kell tartalmazni. Ez az elem tartalmazza a további elemeket, a fenti hierarchia szerint, illetve a kiinduló jelzetet és annak *különböző nyelvű leírásait*. Az alábbi példa egy 1990. évi Németország-Anglia labdarúgó mérkőzést leíró ETO-jelzet XML reprezentációját 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="2005"
notation="796.332.093.434 (430:410.1) &quot;1990\&quot;">
  <ns:description xml:lang="HU">
    Németország-Anglia labdarugó elődöntő mérkőzés (1990).
  </ns:description>
  <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:udc_concept>

```

6 Az összetett ETO számok automatikus interpretálása

Az ETO számoknak egy ennyire bonyolult és összetett formára való manuális konvertálása irreális elvárás lenne az osztályozókkal és az adatbázisok karbantartóival szemben. Ezért indokolt elvárás ennek a folyamatnak az automatikus úton történő támogatása. Egy ilyen megoldás további előnyei közé tartozna, hogy lehetővé tenné a már meglévő bibliográfiai rekordok konverzióját is. Disszertációm ezen fejezetében azt szeretném megmutatni, hogy lehetséges olyan algoritmus tervezése és implementálása, amely a prekoordinált ETO-jelzeteket elemezve, azok valamennyi elemét felismerve és szintaktikai struktúráját megtartva konvertálja azokat az előzőekben bemutatott XML formátumba.

Az ETO-jelzetek közvetlen feldolgozása a következő *előnyökkel* jár:

- Bár az adatbázisok típusa és az általuk használt formátumok változhatnak, a kiinduló ETO-jelzet valamennyiben fellelhető
- Az osztályozóknak és az authority fájl karbantartóinak nem szükséges speciális, rendszerfüggő ismeretekkel rendelkezniük, elég az osztályozási jelzetek lehető legpontosabb összeállítására koncentrálniuk
- A jelzetek automatikus feldolgozása további kutatások alapjául szolgálhat, például a hatékonyabb kereső és böngésző algoritmusok kifejlesztése terén.

Bár az ETO-jelzetek elemzését és felbontását az 1960-as évek óta sokan kutatták és kereskedelmi szoftverek is tartalmaznak ilyen jellegű megoldásokat, *a korábbi próbálkozások elsősorban a jelzetelemek azonosítására és visszakereshetővé tételére fókuszáltak és nem fordítottak kiemelt figyelmet azok szintaktikai kontextusára.*

6.1 Az interpretálás alapelvei

Az algoritmussal szemben elvárás, hogy kövesse az ETO jelzetépítési szabályait, amennyire lehetséges, információvesztés nélkül. Részletesebben kifejtve ez a következőket jelenti:

- Az algoritmusnak pontosan azokat a jelzeteket kell felismernie, amelyeket az ETO szabályainak betartásával alkottak meg, függetlenül a felhasznált ETO kiadástól.

- Az algoritmusnak meg kell őriznie a jelzetekben tárolt szintaktikai információ egészét.
- A jelzetelemzőnek amennyire csak lehetséges, szintaktikai alapon kell működnie.
- A folyamatnak teljesen automatikusnak kell lennie.

Egy a fenti elvárásokat kielégítő program legfontosabb **újdomsága**, hogy az ETO legkülönlegesebb jelzetépítési megoldásait is képes felismerni, megtartja a jelzetek teljes szintaktikai struktúráját, miközben kezeli az osztályozási rendszer változásait is.

Az Internet és a hozzá kapcsolódó infrastruktúra fejlődésével a **Software-as-a-Service** (SaaS, szolgáltatásként működő szoftver) modellje is egyre jelentősebbé vált. Az SaaS modell lehetővé teszi az ETO számok egy távoli szolgáltatás által történő analízisét, így a meglévő programoknak csak egy vékony középréteget kell implementálniuk, amely kapcsolódik ehhez a szolgáltatáshoz. Ezzel a megoldással támogatható az authority kontroll vagy bármilyen más, az ETO alkalmazására irányuló módszer.

A fenti követelmények alapján a megvalósításra a **Java** programozási nyelv lett kiválasztva, elsősorban a szükséges technológiák kiemelt támogatása miatt.

6.2 Az interpreter működésének alapjai

Az ETO táblázatai és jelzetépítési szabályai egy *formális nyelvet* határoznak meg az arab számjegyek és az ETO metanyelvi jelei által alkotott ábécé fölött. Az interpreter egy olyan *automata*, amely ezt a formális nyelvet ismeri fel.

A bemenő szavak az ETO-jelzetek és az automata állapotai az éppen feldolgozott jelet tartalmazó jelzetelem típusát és szerepét írják le. A szám feldolgozása során az automata legenerálja az annak elemeit az előző fejezetekben leírt precedencia sorrend szerint tartalmazó fát. A feldolgozás eredménye a jelzet elemei mellett egyéb olyan információkat is tartalmazhat, amelyekre szükség lehet a visszakeresés során.

Ha az automata nem ismer fel egy számot, az azt jelenti, hogy az nem felel meg pontosan az ETO szabályainak, vagy nem értelmezhető a megadott ETO-verzió alapján.

Az algoritmus *bemenete* egy ETO-jelzet, a jelzet összeállításához felhasznált ETO-kiadás évszám és a jelzet természetes nyelvű leírásai (tetszőleges nyelveken). A *kimenete* a szám XML reprezentációja vagy egy a problémát leíró hibaüzenet, amennyiben a feldolgozás sikertelen volt a megadott ETO verzió alapján. Az interpreter nemcsak a gépelési *hibák felismerésére* képes, de a helytelen indexelési megoldásokéra is.

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"

- Képernyőképek a gépelési hibák felismeréséről

A cél egy olyan interpreter írása, amely lehetőleg *teljesen szintaktikai úton* elemzi az ETO-jelzeteket, a táblázati számok beazonosítása nélkül, pusztán a központoszási jelek segítségével. Ez azt is jelenti, hogy két, hasonló szintaktikai struktúra szerint felépülő szám esetén a produkált kimenetek is hasonlóak lesznek, csak magukban a táblázati számokban különbözve.

Az interpreter és az algoritmus folyamatos fejlesztés alatt áll, tükrözve a kutatás aktuális állását. Kidolgozás alatt állnak az eredmény felhasználását megkönnyítő szabványos interfészek és konverziós metódusok is. A szoftver aktuális verziója a <http://piros.udc-interpreter.hu#process> címen érhető el.

6.3 További fejlesztési irányok és kutatási tervek

Az ETO számok XML formátumban történő eltárolásának 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 információkat. 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 jelzetalak és az XML alapján lehetőség nyílik olyan KWIC-indexek építésére, amelyek lehetővé teszik a jelzetelemek szerinti böngészést figyelembe véve azok környezetét is. 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 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.

Introduction

Universal Decimal Classification (UDC) is one of the largest and most widely used classification systems in the world. It has been published in different editions in no fewer than 39 languages and is used in at least 124 countries; there are currently over 140,000 collections which have been indexed using UDC in Europe alone.

This prevalence is partially due to its invention more than one hundred years ago and *its having been under continuous review* since then to ensure that it is constantly kept up to date. The other key to its popularity is *its analytico-synthetic nature*; it attempts to collect and organize all branches of human knowledge in a hierarchical structure and provides a set of auxiliaries which allow the user to create new concepts from those already in existence.

Despite its advantages, *its employment is often argued*, not least because *information retrieval software systems haven't reflected the significance of this data by using all the capabilities of UDC*.

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*.

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** attended to prove that such solution can be applied automatically without any additional effort or even retrospectively on existing collections.

The *first part of my dissertation* provides an **overview of the literature** regarding the UDC and its automation. In the *first*, introductory chapter I will review *the basic concepts and terminology of classification theory* I will be used later. The *second chapter* introduces the UDC, by *outlining its history and significance*. The *third chapter* introduces the classification itself, by highlighting *its special characteristics*, those mostly affect its automation. The *questions of the automation* detailed in the *fourth chapter*.

The *second part of the dissertation* contains **my own research results**. In the *fifth chapter* I will provide **the detailed description of a computer readable format to store UDC numbers by keeping all the information hidden in them**. In the sixth chapter I will introduce a **software interpreter** I have developed **to process even the most complicated UDC notations** and convert them into the format described in the previous chapter. The software is also **able to generate outputs in different formats** which support the different approaches. Finally, I will summarize the **possible further research directions and plans** for the future.

1 Basic concepts of classification science

The general meaning of the word “**classification**” is grouping, ordering, or sorting objects based on a few chosen characteristics of them. In the terminology of information retrieval “classification” means the interpretation of contents in a (usually artificial and hierarchical) *information language*.

The classification language must reflect not only the concepts, but their **semantic relationships** as well. These relations provide the structure of the classification system, which grants the possibility of browsing in that. The third component of a classification system is the system of the formal and informal rules and regulations to describe the **syntactic relations**. Their relevance is given by the expectation to explain not only the attributes of the described content, but the conceptual relationships between them as well. It is necessary in order to support the proper decisions about the relevance of the text by explaining its content as precisely as possible.

The permanent, constantly valid relations those cited often in the given collection are usually called as **paradigmatic relations**, to differentiate them from the **syntagmatic relations** that describe occasional, rarely occurrent, ad-hoc concepts. The paradigmatic relations can be registered in the vocabulary, as a conclusion the proper control can grant their effective handling in a mechanized environment as well. However, the automation of syntagmatic relations is problematic and has not been solved in most systems.

The fundamental structure of the UDC is often referred to as **analytico-synthetic**: in addition to its deep and systematic concept hierarchy it also provides a rich set of facilities for building classmarks for both specifying and join concepts. Due to this complex nature, the level of handling syntactic relations is a key question in a system that employs an analytico-synthetic classification for indexing and organizing contents.

2 The Universal Decimal Classification

The UDC is one of the three “classic” classification systems, which have been created at the end of the Nineteenth Century and have become the most prevalent library classifications in the world. It was constructed by having extended the Dewey Decimal Classification, in order to build a universal catalog to register all of the printed documents on the planet. As a result of the extensions mentioned above, the UDC became *the first faceted, analytico-synthetic classification*. Although it was created for bibliographical purposes, it has been used also by libraries from the beginnings. In addition to its direct employment in libraries, due to its flexibility and rich vocabulary, it has been often used as a source for terminology of specialist subject classifications and thesauri and has had influence on other universal classification schemes.

The UDC was managed and maintained by International Federation for Documentation (FID) until the 1990s, when the copyrights of the UDC and the responsibility for its development and dissemination were passed to a self-funded, non-profit Consortium. In the same time the **UDC Master Reference File (MRF)**, a database with the size of the earlier medium editions was compiled. The UDC MRF contains the UDC schedules (about 70,000 records) together with records needed for administration, maintenance, and archiving. It can be employed *with the proper license*, and it is also available online in seven languages, with a commercial license.

In the 1970s the structure and the revision process of UDC were refactored due to the changes in the expectations, the progress in classification science and the experience of the automation of the classification. The revision also affected the common approach of building classmarks. In the recent three decades the *systematical faceted revision* of the system has come to the front.

During the recent 120 years the UDC become *the second most prevalent library classification*. The most recent research shows that it has been used in 124 countries of the world. In Europe it has been

used by at least 144,000 libraries in 28 countries. Its different editions have been published in at least 39 languages, its brief online edition is available free in 57 languages.

In Hungary the UDC has been known and employed for 120 years. By the second half of the twentieth century it became the primary library classification. Although its usage is often argued, the most recent research, conducted in 2016 shows it is still used actively in most of the different libraries.

The last printed Hungarian edition was published in 2006, and arrangements were made to integrate the Hungarian translation of the current MRF into a namespace system that is presently under implementation.

3 The UDC as an Analythico-Synthetic Classification Scheme

As we already mentioned above, the fundamental structure of the UDC is usually referred to as *analytico-synthetic*, because it contains the elements of both the enumerative and the faceted classification schemes.

The enumerative nature lays in the deep concept hierarchy of the *main and the common auxiliary tables*. The UDC also provides a lot of *special auxiliaries* and other solutions to build classmarks, in order to describe facets that are attributes that typically occur within a class. Finally, the synthetic nature is given by the *common auxiliaries*, that make it possible to explain complex and agglomerated subjects as well, even on more levels, in order to describe contents as precise as is possible.

The UDC is a *strongly pre-coordinated* scheme. Its rich, multilevel facilities to express composite subjects mean that interpreting its classmarks requires to take the syntactic relations into account, because the way how the classes join to each other, the type of the relations or even the order of the elements can all determine the meaning. Disregarding this information during the indexing or the retrieval process may affect not only the precision, but the recall too.

4 The Automatization of UDC Numbers

The synthetic principle of Universal Decimal Classification (UDC) is an advantageous feature as it allows the combination of concepts in the process of indexing and thus provides a great power in describing composite subjects in order to support information retrieval. However, the pre-combination of concepts creates a new concept, which not only inherits the semantics of its parts but acquires its own, new meaning. The same UDC number when used in different relationships may represent different subjects. There is a consensus in the literature that the integrity of data elements and UDC semantics depends on better notation management in the process of automation.

This leads to two important *requirements* with respect to the management of UDC numbers within an information retrieval system:

- a) the integrity of the semantic components of the UDC data elements must be preserved, i.e. it should be possible to access and search for each individual component of a complex UDC number and to trace its original verbal description in the process of information retrieval
- b) combined UDC numbers must be presented and treated as a whole together with the information regarding their syntactic structure and the verbal expression containing additional meaning that the number as a whole represents, and which cannot be implicated from the individual parts.

In the past the following solutions were devised and implemented to store UDC numbers:

- 1) dealing with UDC notation "as given"
 - a) simpler approach – pre-coordinated UDC notations were entered in the bibliographic databases as **simple text**
 - b) **KWOC-index** – UDC numbers were split into individual components to allow for notation searching

- c) **KWIC-index** – the individual components can be retrieved within their context
- 2) representing and storing a notation as structured and coded data
 - a) **authority control** – i.e. notations are stored in classification authorities which would give provision for each individual notation from the pre-coordinated notation statement to be identified, accessed, and linked to its verbal representation, while the authority file can also reveal the semantical relationships between the records.

A common attribute of the first group of solutions is the low level of precision due to the undeveloped support of the syntactic relations. On the other hand, losing information is also common because of the lack of the consistency in the representations and the unsatisfactory identification of the data elements.

An *authority record* may contain not only the classification number and its metadata, but also additional semantic information. This may include broader, narrower, and related classes, mappings to thesauri, subject heading lists or other classifications and access points using either the codes of the classification or keywords in natural languages. Due to the developed level of exploring the semantic relationships, authority control can effectively support information retrieval and subject browsing.

However, creation and maintenance of an authority control is *very costly* in time and effort consuming and requires expertise that many libraries do not have. Consequently, the practice of authority control in managing UDC numbers can only be found in well managed national and university libraries. The majority of bibliographic data that can be found in library networks is not maintained in authorities and encouraging the authorities to be created requires some kind of robust automated solution. Another disadvantage of the method is that *it cannot be employed in the case of syntagmatic concepts*.

5 The representation of UDC notations

In the course of this research into improving the ways of storing UDC notations, it became clear that almost *every UDC number can be represented as a tree* with the same hierarchical structure. This is determined by the conceptual precedence of the operators and the definitions of common and special auxiliaries (the exceptions are mainly caused by breaking the rules concerning combining numbers) which describes their syntactic structure keeping all relevant information regarding its parts, the way they join to each other, their role in the compound, or perhaps their order if it is relevant.

A great advantage to the representation mentioned above is its *independence from citation order*. If two numbers differ only in the citation order of their parts, they have the same logical structure, which will result in the same representation. However, order can be added as an attribute of the parts of the numbers in order to make a difference if the citation order has a semantic significance.

5.1 The syntactic structure of UDC numbers

The **precedence order** of the operations is as follows.

6. + Coordination
7. : Simple relation
8. :: Order-fixing
9. [] Subgrouping
10. ‘ Synthesis (within special auxiliaries)

The above operators *can join* numbers – considering the priority order – or ranges of numbers in the schedules. The top three can also join *subgroupings* inside which the same hierarchy operates (although it cannot contain additional subgroupings). In addition to the common auxiliary symbols, there is also *synthesis* within special auxiliaries which make it possible to join numbers of a given part of a schedule (such as numbers for chemical compounds or political parties). It is also possible to use *operators inside common auxiliaries*; the rules depend on the type of the auxiliary.

The concepts joined by operators can contain zero or one main table number (or a range of main table numbers) with zero or more (or a range of) special auxiliaries joined to this and zero or more common auxiliaries (which can also contain numbers, ranges, operands, and special auxiliaries).

The *special auxiliaries* in UDC contain very different solutions in order to describe the most specific subject facets (such as properties, processes, operations, techniques, instruments) which have to be handled in different ways.

5.2 Using XML to describe the structure of the UDC numbers

The simple and composite UDC numbers can be represented and possibly stored in a database as a tree. As with every tree, this structure can be described with XML.

In my dissertation I will discuss a feasible way of building a UDC-specific **XML schema** for describing the most detailed and complicated UDC numbers (containing not only the common auxiliary signs and numbers, but also the different types of special auxiliaries). The schema definition is available online at: <http://piros.udc-interpreter.hu#xsd>.

The XML file generated based on the schema definition *can be used to describe every possible form of the UDC numbers* which observe the rules of UDC, and due to its comprehensiveness *can be converted into different formats* automatically.

The **complex types** of the XSD describe the possible elements of UDC numbers: the *numbers of the schedules*, the *auxiliary signs* and the *UDC number itself*. For example, the following complex type can be employed to describe common auxiliaries of form.

```
<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>

```

An operator can be described by a complex type containing its possible operands. The following complex type describes the addition of main table numbers.

```

<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>

```

The **simple types** of the XSD were introduced *for validation purposes*; the constraints defined for them contain regular expressions to describe the possible formats of the numbers in the different schedules. The following type was constructed for describing the numbers in common auxiliaries of time.

```

<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:restriction>
  <xsd:pattern value="(-|\+)?([0-2]\d{3}(\.\d{2}(\.\d{2}(\.\d{2}(\.\d{2}(\.\d{2})?)?)?)?)?)?)" />
  <xsd:pattern value="[3-9]((\.)?\d{1,4})(\.\d{1,4})*" />
</xsd:simpleType>

```

Each of the UDC concepts is described by a complex element. In accordance with the schema XML must contain one element to describe a UDC concept. This element contains the other elements as can be seen in the hierarchy above. It also contains the notation of the number as an attribute and the *descriptions of the concept in different languages* as elements too. The following example demonstrates the XML representation of the UDC-number of a football match between Germany and England.

```

<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="796.332.093.434 (430:410.1) &quot;1990\&quot;;">
<ns:description xml:lang="EN">
  Germany-England football semifinal (1990).
</ns:description>
<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:udc_concept>

```

6 The automatic interpretation of pre-coordinated UDC numbers

Converting UDC numbers manually to a complex format such as the one mentioned above is an unrealistic expectation; supporting building these representations, as far as possible automatically, is a well-founded requirement. An additional advantage of this approach is that the existing records could also be processed and converted. In my dissertation I would like to prove also that it is possible to design and implement an algorithm that is able to convert pre-coordinated UDC numbers into the introduced format by identifying all their elements and revealing their whole syntactic structure as well.

Processing the UDC numbers directly has the following advantages:

- Although the databases and the formats used in information retrieval systems vary, the UDC number itself is a stable point in all of them.
- In this way the system would not require the indexers and the maintainers of the authority files to utilize special, system related knowledge, consequently they can concentrate on building the numbers taking maximum advantage of their skills and competencies related to the classification,
- Future research can be conducted based on the processing of UDC numbers so as to enhance the efficiency of searching and browsing for example by discovering the connections between the complex and compound concepts.

The parsing of pre-coordinated UDC numbers is a topic which has been subject to thorough investigation since the 1960s, however, earlier research has focused on the identification of the independent parts of UDC classmarks without keeping their syntactic context.

6.1 The underlying principles of the interpreter

The proposed algorithm for UDC notation interpretation must follow the syntax rules for UDC subject synthesis, as far as possible, with no loss of information. This specifically means the following:

- The algorithm must recognize those numbers and only those numbers which keep to the rules for synthesizing UDC numbers regardless the UDC edition used to build them.
- The algorithm must retain all the information stored by the number.
- The parsing method must be fully syntactic as far as is possible.
- The process must be fully automated.

The **novelty** of a program such satisfies the above expectations is due to *its ability to identify even the most specific solutions* for building UDC classmarks by retaining all of the syntactic information stored by the number, and *to react to the changes of the schema*.

With the evolution of the Internet and related infrastructure, the **Software-as-a-Service (SaaS)** model has been on the increase significantly. Using SaaS delivery model allows the UDC numbers to be analyzed by a remote service and implement only a thin middleware to join its output to the client software. In this way the results would be able to be utilized for the purposes of managing authority files as well as by systems employing other methods to access UDC numbers.

Considering the above requirements, **Java** was chosen as the programming language, due to its excellent support for related technologies.

6.2 The basics of how the interpreter works

The rules governing the usage of UDC determine a *formal language* over an alphabet of decimal digits, dots, and the characters of UDC metalanguage symbols. The interpreter is an *automaton* which recognizes this formal language.

The input words are UDC numbers and the states of the automata describe the type and the role of that part of the number to which the letter being processed belongs. While processing the number, the automaton generates a tree, which contains the parts of a number, based on the rules of precedence regarding auxiliary signs and joining auxiliary numbers as explained in the previous sections. It also contains some additional information which may be necessary during a search.

If the automata does not recognize the input word, then the given number is invalid or cannot be interpreted in the given UDC version. This means that in principle the language recognized by the automata specified by a year is just the UDC edition for that year.

The *inputs* of the algorithm are the UDC number and the year of the UDC edition which was used to build it; the *output* will be the XML representation of the number, or an error message which describes the problem if it cannot be interpreted in the given version. The interpreter *is able to identify* problems that caused by either *typing errors or indexing practices that didn't exactly match the rules* of UDC.

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"092"

- Screenshots of the recognition of typing errors

The proposition is to construct an algorithm which is able to parse UDC numbers *in a syntactic way* as far as possible. This would mean that parsing two different numbers having the same structure would result in similar representations, differing only in the numbers themselves. It would also require every UDC number to be parsed without the need for identifying the numbers, based on the punctuation signs and rules only.

The interpreter is under continuous development; the algorithm has been improved depending on the status of the research; standard interfaces and conversion methods have started to be worked out as well. The current version of the software is available at: <http://piros.udc-interpreter.hu#process>.

6.3 Future research directions and plans

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 do not require special algorithms and programming effort. This may improve the efficiency of those methods which apply 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** those subjects inclusively.

The generated XML may also serve as the bases to the development of outputs and methods that can support implementing **intelligent classification interfaces**. Generating the notation that contains the proper tags to arrange the numbers in the correct filing order have been implemented. Building 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 too. 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 changing the notations of the numbers automatically and displaying them in a form that fits to the supposed target of the searching process and the cognitive status of the searcher during subject browsing.

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 of examining how the revisions and appropriate changes performed in recent years have helped in handling UDC notations.

Irodalomjegyzék / Bibliography

- [1] „About Universal Decimal Classification (UDC).” UDCC (UDC Consortium).
<http://udcc.org/index.php/site/page?view=about>.
- [2] Alexander, Fran, Kathryn Stickley, Vicky Buser, és Libby Miller. 2010. UDC at the BBC.
Extensions and Corrections to the UDC 32: 21-24.
- [3] 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>.
- [4] „ASCII History.” ASCII World. <http://ascii-world.wikidot.com/history>.
- [5] 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.
- [6] Babiczky 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.
- [7] Babiczky Béla. 1975. *Könyvtári feldolgozó munka: A tanítóképző intézet II. évfolyama számára (népművelő-könyvtáros szak). 2. rész. Osztályozás és szakkatalógus szerkesztés*. Budapest: Tankönyvkiadó.
- [8] Babiczky Béla. 1978. „Új szakasz az Egyetemes Tizedes Osztályozás fejlődésében.” *Könyvtári figyelő* 24 (6): 659-671.
- [9] Babiczky Béla. 1985. „Szemléletváltozás az ETO jelzetszerkesztésében.” *Könyvtári Figyelő* 31 (1): 17–27.
- [10] Babiczky 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.
- [11] Babiczky Béla, és Barátné Hajdu Ágnes. 1998. *Bevezetés az információkereső nyelvek elméletébe és gyakorlatába*. Budapest: Universitas.
- [12] Balázs Sándor. 1960. „Magyarország képviselete és szerepe a FID-ben.” *Magyar könyvszemle* 76 (1): 58-64.
- [13] 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>.
- [14] Balíková, Marie. 2009. „The role of UDC classification in the Czech Subject Authority File.” *Extensions and Corrections to the UDC* 31: 253-274.
- [15] Balogh Béla. 2014. „ETO és Google.” Katalist (levelezőlista).
<https://listserv.niif.hu/pipermail/katalist/2014-January/030127.html>.
- [16] 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.
- [17] Barátné Hajdu Ágnes. 2006. „Usability and Responsibility.” *Extensions and Corrections to the UDC* 28: 46-55.
- [18] 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>.
- [19] 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>.

- [20] Barátné Hajdu Ágnes. 2007c. „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.
- [21] 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>.
- [22] Barátné Hajdu Ágnes. 2011. „Az ismeretszervezés dinamizmusa, különös tekintettel az Egyetemes Tizedes Osztályozásra.” *Könyv és nevelés* 13 (1): 63-72.
- [23] Barátné Hajdu Ágnes, és Fejős László. 2005. „New Hungarian UDC edition.” *Extensions and corrections to the UDC* 27: 8-12.
- [24] 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.
- [25] 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.
- [26] Barki katalin, és Prejczér Paula. 2003. *Számítógépes katalogizálás hunmarc formátumban*. Szombathely: Berzsényi Dániel Főiskola Központi Könyvtár, Berzsényi Dániel Főiskola Könyvtár- és Információtudományi Tanszék. <http://mek.oszk.hu/05200/05287/html/index.htm>.
- [27] 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>.
- [28] Batley, Susan. 2014. *Classification in Theory and Practice*. Amsterdam: Elsevier Science.
- [29] 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.
- [30] 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>.
- [31] Berners-Lee, Tim, James Hendler, és Ora Lassila. 2001. „The Semantic Web” *Scientific American* 284 (5): 29-37.
- [32] 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>.
- [33] „A biterminal symbol for non-UDC notation.” Notes & Queries. UDC Consortium (UDCC). <http://www.udcc.org/notes.htm#bitsymbol>.
- [34] Bliss Classification Association (BCA). 2015. „The Bliss Bibliographic Classification: history & description.” <http://www.blissclassification.org.uk/bchist.shtml>.
- [35] Bliss, Henry Evelyn. 1929. *The organization of knowledge and the system of the sciences*. New York, NY: Henry Holt and Company.
- [36] 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/>.
- [37] Bourret, Ronald. 2011. „XML Data Binding Resources.” Legutóbb módosítva 2011. október 7-én. <http://www.rpbourret.com/xml/XMLDataBinding.htm>.

- [38] Bourne, Charles P., és Trudi Bellardo Hahn. 2003. *A History of Online Information Services, 1963-1976*. Cambridge, MA, and London: MIT Press.
- [39] British Standards Institution (BSI). 2005. *Universal Decimal Classification: standard edition: volume 1: systematic tables*. London: British Standards Institution.
- [40] Bourne, Charles P., és Trudi Bellardo Hahn. 2003. *A History of Online Information Services, 1963-1976*. Cambridge, MA, and London: MIT Press
- [41] Broughton, Vanda. 2010. „Concepts and Terms in the Faceted Classification: the Case of UDC.” *Knowledge Organization* 37 (4): 270-279.
- [42] Broughton, Vanda. 2015. *Essential Classification. Second Edition*. London: Facet Publishing.
- [43] Buxton, Andrew. 1990. „Computer Searching of UDC Numbers.” *Journal of Documentation* 46 (3): 193-217.
- [44] Buxton, Andrew. 2011. „Duplication of Concepts in UDC.” *Extensions & Corrections to the UDC* 33: 13-17. <http://hdl.handle.net/10150/232711>.
- [45] „Cancellations to the UDC.” UDCC (UDC Consortium). <http://udcc.org/index.php/site/page?view=cancellations>.
- [46] „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/>.
- [47] 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>.
- [48] Chatterjee, Amitabha. 2016. *Elements of Information Organization and Dissemination*. Amsterdam: Elsevier.
- [49] Civalero, 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.
- [50] „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.
- [51] 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
- [52] „Collections indexed by UDC.” UDCC (UDC Consortium). <http://udcc.org/index.php/site/page?view=collections>.
- [53] „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/>.
- [54] „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/>.
- [55] 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.
- [56] Csík Tibor. 1995. „Ismeretek és könyvtári osztályozás.” *Könyv, Könyvtár, Könyvtáros* 4 (4): 13-24. <http://epa.oszk.hu/01300/01367/00163/pdf/02muhelykerdesek.pdf>.
- [57] Dahlberg, Ingetraut. 1972. „Az ETO újjáalakításának lehetőségeiről” [Fordította Horváth Magda]. *Könyvtári Figyelő* 18 (4): 412-430.
- [58] Dahlberg, Ingetraut. 1978. *Ontical Structures and Universal Classification*. Bangalore. Sarada Ranganathan Endowment for Library Science.
- [59] 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.

- [60] Dahlberg, Ingetraut. 2017. „Why a New Universal Classification System is Needed.” *Knowledge Organization* 44 (1): 67-71.
- [61] de Grolier, Eric. 1962. *A study of general categories applicable to classification and coding in documentation*. Paris: UNESCO.
- [62] 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.
- [63] „Dewey Decimal Classification summaries.” Online Computer Library Center (OCLC).
<https://www.oclc.org/dewey/features/summaries.en.html>.
- [64] „Digital National Library of Portugal.” The European Library. National Library of Portugal.
<http://www.theeuropeanlibrary.org/tel4/collection/a0005>.
- [65] 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.
- [66] Ellis, David. 1989. „A Behavioural Approach to Information Retrieval System Design.” *Journal of Documentation* 45 (3): 171-212.
- [67] 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.
- [68] „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.
- [69] „Extensible Markup Language (XML).” World Wide Web Consortium (W3C). Legutóbb módosítva 2016. október 11-én. <https://www.w3.org/XML/>.
- [70] „Extensions and Corrections to the UDC.” UDCC (UDC Consortium).
<http://udcc.org/index.php/site/page?view=ec>.
- [71] 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.
- [72] 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>.
- [73] 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>.
- [74] 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>.
- [75] 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>.
- [76] 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>.
- [77] 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.
- [78] 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.

- [79] 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.
- [80] FID (International Federation de Documentation). 1983. *Az Egyetemes Tizedes Osztályozás (ETO) alapelvei, revíziójának és kiadásának szabályai (FID 603)*. Budapest: OMIKK.
- [81] FID (International Federation de Documentation). 1990. *Task Force For System Development: final report*. Hága: International Federation de Documentation.
- [82] 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.
- [83] 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).
- [84] Frâncu, Victoria. 2003. „Multilingual access to information using an intermediate language.” PhD diss., Antwerpen: Universiteit Antwerpen. <http://hdl.handle.net/10150/106319>.
- [85] 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.
- [86] 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.
- [87] Frické, Martin H. 2016. „Logical division.” *Knowledge Organization* 43 (7): 539-549.
- [88] 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.
- [89] 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
- [90] 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.
- [91] Gendler-Szabó Zoltán. 2004. „Compositionality.” In *Stanford Encyclopedia of Philosophy*. Stanford, CA: Stanford University. <https://plato.stanford.edu/entries/compositionality>.
- [92] 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>.
- [93] „Global Forest Decimal Classification (GFDC) Database.” International Union of Forest Research Organizations (IUFRO). <http://iufro.forintek.ca/GFDCDefault.aspx>.
- [94] Gnoli, Claudio. 2009. „The UDC Philosophy Revision: First Report.” *Extensions and Corrections to the UDC* 31: 25-31. <http://hdl.handle.net/10150/200633>.
- [95] 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.
- [96] 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>.
- [97] 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.
- [98] 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.
- [99] Gnoli, Claudio. 2018. „Notation.” *Encyclopedia of Knowledge Organization*. ISKO. <http://www.isko.org/cyclo/notation>.
- [100] 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.
- [101] 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.
- [102] 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).
- [103] 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>.
- [104] Guenther, Rebecca S. 1992. „The Development and Implementation of the USMARC Format for Classification Data.” *Information Technology and Libraries* 11 (2): 120-31.
- [105] Gulati, Dipti. 2013. „Use of library classification schemes in the ICT environment in selected libraries in national capital region: a study.” PhD diss., Delhi: University of Delhi. <http://shodhganga.inflibnet.ac.in:8080/jspui/handle/10603/31769>.
- [106] 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.
- [107] Hiroshi. 2007. „Advantages & Disadvantages Of XML.” <http://www.techmynd.com/advantages-disadvantages-of-xml/>
- [108] Hjørland, Birger. 2013. „Facet analysis: The logical approach to knowledge organization.” *Information Processing and Management: an International Journal* 49: 545–557.
- [109] Hjørland, Birger. 2017a. „Subject (of Documents).” *Knowledge Organization* 44 (1): 55-64. <http://www.isko.org/cyclo/subject>.
- [110] 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.
- [111] 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.
- [112] 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>.
- [113] Hutchins, Williams John. 2001. „Indexelő és osztályozó nyelvek. Nyelvészeti tanulmány struktúráról 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.
- [114] International Federation of Library Associations (IFLA). 2000. „Concise UNIMARC Classification Format.” http://softsbgp.free.fr/bibliotheque/Unimarc_Format_US.pdf.
 - [115] International Federation of Library Associations (IFLA). 2001. „UNIMARC Manual. Authorities Format.” <https://www.ifla.org/files/assets/uca/unimarc-authorities-format.pdf>.
 - [116] International Federation of Library Associations (IFLA). 2008. „UNIMARC concise bibliographic Format.” <http://archive.ifla.org/VI/8/unimarc-concise-bibliographic-format-2008.pdf>.
 - [117] International Federation of Library Associations (IFLA). 2010. „Functional Requirements for Subject Authority Data (FRSAD): Final Report.” <https://www.ifla.org/node/5849>.
 - [118] „International UDC Seminar 2015.” UDCC (UDC Consortium). <http://seminar.udcc.org/2015/>.
 - [119] „International UDC Seminar 2017.” UDCC (UDC Consortium). <http://seminar.udcc.org/2017/>.
 - [120] „International UDC Seminars.” UDCC (UDC Consortium). <http://seminar.udcc.org/>.
 - [121] „ISO 2709.” International Federation of Library Associations (IFLA). <http://www.ifla.org/best-practice-for-national-bibliographic-agencies-in-a-digital-age/node/8811>.
 - [122] „ISSN Manual test.” International Standard Serial Number (ISSN) International Center. Utolsó módosítás 2010. <http://www.issn.org/issn-manual-test/>.
 - [123] Iyer, Hemalata. 2012. *Classificatory Structures: Concepts, Relations and Representation*. Würzburg: Ergon Verlag.
 - [124] Horváth Tibor (1968). „A második cranfieldi jelentés.” *Könyvtári Figyelő* 14 (5): 351-369.
 - [125] Horváth Tibor. 1997. „Don Quijoték útjain.” *Könyvtári Figyelő* 44 (4): 665-674. <http://ki.oszk.hu/kf/kfarchiv/1997/4/horvath.html>.
 - [126] 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.
 - [127] „Katalist. 2014 januári archívum.” Katalist. <https://listserv.niif.hu/pipermail/katalist/2014-January/>.
 - [128] 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.
 - [129] Kleineberg, Michael. 2017. „Integrative levels.” *Knowledge Organization* 44 (5): 349-379.
 - [130] 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>.
 - [131] Kumar, T.V.R., és Parameswaran, M. 1999. „Chain indexing and UDC.” *Annals of Library Science and Documentation* 46 (2): 53–58.
 - [132] „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.
 - [133] 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>.
 - [134] 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>.
 - [135] 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>.

- [136] 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>.
- [137] „Library of Congress Classification.” The Library of Congress.
<https://www.loc.gov/catdir/cpsol/lcc.html>
- [138] 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.
- [139] 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.
- [140] „Major Changes to the UDC 1993-2013.” UDCC (UDC Consortium).
http://www.udcc.org/index.php/site/page?view=major_revisions.
- [141] Mándy Gábor. 2011. „Régi ETO, új ETO.” *Könyvtári Figyelő* 57 (2): 317-332.
<http://ki.oszk.hu/kf/2011/07/regi-eto-uj-eto/>.
- [142] 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.
- [143] „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>.
- [144] 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.
- [145] 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.
- [146] McIlwaine, I. C. 2000. *The Universal Decimal Classification – a Guide to its Use*. The Hague: UDC Consortium.
- [147] McIlwaine, I. C. 2006. „The new ecumenism: Exploration of a DDC/UDC view of religion.” *Extensions and Corrections to the UDC* 28: 9-16.
- [148] 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.
- [149] „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/>.
- [150] Mitchell, Joan S., Marcia Lei Zeng, és Maja Žumer. 2011. „Extending models for controlled vocabularies to classification systems: modelling DDC with FRISAD.” 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.
- [151] 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.
- [152] Murell, Paul. 2009. *Introduction to Data Technologies*. London: Chapman & Hall / CRC
<https://www.stat.auckland.ac.nz/~paul/ItDT/>
- [153] Nabhan, Yulia. 2010. UDC News in Ukraine. *Extensions and Corrections to the UDC* 32: 31-34.

- [154] Neelameghan Arashanipalai. 1973. „Agglomerate basic subjects.” *Library Science with a Slant to Documentation* 10 (2): 202-206.
- [155] Noreikiene, Dalia. 2008. UDC In Lithuania. *Extensions and Corrections to the UDC* 30: 53-54. <http://hdl.handle.net/10150/110844>.
- [156] Nyamisa, Mang'are Fridah, Waweru Mwangi, és Wilson Cheruiyot. 2017. „A Survey of Information Retrieval Techniques.” *Advances in Networks* 5 (2): 40-46.
- [157] 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.
- [158] 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.
- [159] O'Neill, Edward T., Rick Bennett, és Kerre Kammerer. 2014. „Using Authorities to Improve Subject Searches.” *Cataloguing & Classification Quarterly* 52 (1): 6-19.
- [160] 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.
- [161] Orosz Gábor. 1977. „Az Egyetemes Tizedes osztályozás tárgymutatójának összeállítása számítógéppel.” *Tudományos és műszaki tájékoztatás* 24 (7-8): 271-293.
- [162] 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.
- [163] 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.
- [164] „OSZK Tezaurusz / Köztaurusz.” Országos Széchényi Könyvtár (OSZK). Utolsó módosítás 2013. szeptember. <http://mek.oszk.hu/adatbazis/thes.htm>.
- [165] 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.
- [166] „OWL 2 Web Ontology Language Document Overview (Second Edition).” World Wide Web Consortium (W3C). <https://www.w3.org/TR/owl2-overview/>.
- [167] Pálvölgyi Mihály. 1998. *Információfeldolgozás, információkereső nyelvek*. Szombathely: Berzsényi Dániel Tanárképző Főiskola.
- [168] Perrault, Jean Martin. 1974. „Az ETO alapelveiről.” [Tömörítette Horváth Magda] *Könyvtári Figyelő* 20 (2): 126-131.
- [169] 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.
- [170] 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.
- [171] 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.
- [172] 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>

- [173] 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.
- [174] 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.
- [175] 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.
- [176] 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.
- [177] Piros Attila. 2015c. „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.
- [178] 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.
- [179] Piros Attila. 2017b. „A facettás osztályozás napjainkban: elmélet, technológia és a végfelhasználók.” *Könyv, könyvtár, könyvtáros* 26 (11): 22-30.
- [180] Piros Attila. 2018. „New automatic interpreter for complex UDC numbers.” *Extensions and Corrections to the UDC* 36-37 (2014-2015). [megjelenés alatt]
- [181] 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/>.
- [182] 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>.
- [183] 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.
- [184] 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>.
- [185] 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>.
- [186] Ranganathan, S. R. 1967. *Prolegomena to library classification. 3rd ed.* New York: Asia Publishing House. <http://hdl.handle.net/10150/106370>.
- [187] 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>.

- [188] 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>.
- [189] 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>.
- [190] „Rayward’s Otlet Page: Paul Otlet and Documentation.” Rayward, W. Boyd.
<http://people.ischool.illinois.edu/~wrayward/otlet/otletpage.htm>.
- [191] „Referativni Zhurnal Classification Scheme.” legutóbb módosítva 2011. június 1-én.
<http://www.mathontheweb.org/mathweb/Classif/RZhClassification.html>.
- [192] „Revised UDC tables 2003.” *Extensions & Corrections to the UDC* 25.
- [193] Riesthuis, Gerhard. 1997. „Decomposition of complex UDC notations.” *Extensions and Corrections to the UDC* 19: 13-19.
- [194] Riesthuis, Gerhard. 1999. „Searching with words: re-use of subject indexing.” *Extensions & Corrections to the UDC* 21: 24-32.
- [195] Riesthuis, Gerhard. 2003. „A Revised Format for the Master Reference File Database.” *Extensions and Corrections to the UDC* 25: 11-18.
- [196] 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. Koppenhága: Danish Centre for Documentation, 1971.
- [197] Rigby, Malcolm. 1974. *Computers and the UDC: a decade of progress 1963-1973. (FID 523)*. The Hague: International Federation of Documentation (FID)
- [198] 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>.
- [199] Robinson, Geoffrey. 2003. „Citation Order in UDC.” *Extensions and Corrections to the UDC* 25: 19-27.
- [200] 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.
- [201] „RUSMARC.” Российская Национальная Библиотека.
<http://www.rusmarc.info/2017/auth/content.htm>.
- [202] 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>.
- [203] 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.
- [204] 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.
- [205] 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>.
- [206] 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>.

- [207] Satija, Mohinder P. 2017. „Colon Classification (CC).” *Knowledge Organization* 44 (4): 291-307. http://www.isko.org/cyclo/colon_classification.
- [208] 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>.
- [209] 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 diss., Mangalore: Mangalore University. <http://shodhganga.inflibnet.ac.in/handle/10603/131771>.
- [210] 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.
- [211] Shimmer, 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>.
- [212] „Should the slash / be abolished?” Notes & Queries. UDC Consortium (UDCC). <http://www.udcc.org/notes.htm#stroke>.
- [213] 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>.
- [214] Schulte-Albert, Hans G. 2009. „Cyprian Kinner and the Idea of a Faceted Classification.” *Libri* 24 (4): 324-337.
- [215] „SKOS Simple Knowledge Organization System - Home Page.” Semantic Web. World Wide Web Consortium (W3C). <https://www.w3.org/2004/02/skos>.
- [216] 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>.
- [217] Slavić, Aida, és Maria Inês Cordeiro. 2004. „Core requirements for automation of analytical-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>.
- [218] 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.
- [219] 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>.
- [220] Slavić, Aida. 2006. „Interface to classification some objectives and options.” *Extensions and Corrections to the UDC* 28: 24-45.
- [221] Slavić, Aida, Jiri Pika, Gerhard Riesthuis, és Chris Overfield. 2007. „German UDC Translation Project.” *Extensions and Corrections to the UDC* 29: 11-15.
- [222] 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>.
- [223] Slavić, Aida. 2008a. „Faceted classification: management and use.” *Axiomathes* 18 (2): 257-271.
- [224] Slavić, Aida. 2008b. „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>.

- [225] Slavić, Aida. 2008c. „Use of the Universal Decimal Classification: A worldwide survey.” *Journal of Documentation* 64 (2): 211-228.
- [226] 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>.
- [227] 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>.
- [228] 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>.
- [229] Slavić, Aida. 2009c. „UDC SKOSification.” *Universal Decimal Classification* (blog), UDC Consortium, 2009. március 22. <http://universaldecimalclassification.blogspot.hu/2009/03/udc-skosification.html>.
- [230] 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.
- [231] 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.
- [232] 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>.
- [233] 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>.
- [234] Slavić, Aida. 2012c. „UDC as Linked Data.” *Universal Decimal Classification* (blog), UDC Consortium, 2012. augusztus 26. <http://universaldecimalclassification.blogspot.hu/2012/08/udc-as-linked-data.html>.
- [235] 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>.
- [236] 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.
- [237] 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>.
- [238] 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>.
- [239] Tartaglia, S. 2004. „Authority Control and Subject Indexing Languages.” *Cataloging & Classification Quarterly* 39 (1/2): 365-377.
- [240] „The Nobel Peace Prize 1913.” Nobel Media AB. http://www.nobelprize.org/nobel_prizes/peace/laureates/1913/.

- [241] Thék György. 2014. „tárgyszó vagy ETO (is).” Katalist (levelezőlista). <https://listserv.niif.hu/pipermail/katalist/2014-April/030774.html>.
- [242] 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>.
- [243] Tillett, 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.
- [244] „UDC Consortium.” UDCC (UDC Consortium). http://udcc.org/index.php/site/page?view=about_udcc.
- [245] „UDC Editions.” UDCC (UDC Consortium). <http://www.udcc.org/index.php/site/page?view=editions>.
- [246] „UDC Editorial Team & Advisory Board.” UDCC (UDC Consortium). http://udcc.org/index.php/site/page?view=editorial_team.
- [247] „UDC Licences.” UDCC (UDC Consortium). <http://udcc.org/index.php/site/page?view=licences>.
- [248] „UDC Master Reference File (MRF).” UDCC (UDC Consortium). <http://www.udcc.org/index.php/site/page?view=mrf>.
- [249] „UDC MRF Use Licences.” UDCC (UDC Consortium). http://udcc.org/index.php/site/page?view=mrf_licence.
- [250] „UDC Online.” UDCC (UDC Consortium). <http://www.udc-hub.com/>.
- [251] „UDC Online English – Prices.” UDCC (UDC Consortium). <http://www.udc-hub.com/en/prices.php>.
- [252] „UDC Publishing Licences.” UDCC (UDC Consortium). http://udcc.org/index.php/site/page?view=publishing_licence.
- [253] „UDC Subject Coverage.” UDCC (UDC Consortium). http://udcc.org/index.php/site/page?view=subject_coverage.
- [254] „UDC Summary. Multilingual Universal Decimal Classification Summary. UDCC Publication No. 088.” UDCC (UDC Consortium). <http://www.udcc.org/udcsummary/php/index.php>.
- [255] „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>.
- [256] „UDC Summary Exports.” UDCC (UDC Consortium). <http://www.udcsummary.info/exports.htm>.
- [257] „UDC Summary Linked Data.” UDCC (UDC Consortium). <http://udcdata.info/>.
- [258] „UDC Summary Translations.” UDCC (UDC Consortium). <http://www.udcsummary.info/translation.htm>.
- [259] UDCC (UDC Consortium). 2005. „Guidelines for the revision and maintenance of the UDC.” *Extensions and Corrections to the UDC 27*: 19-38.
- [260] „UDC Training Service.” UDCC (UDC Consortium). http://www.udcc.org/index.php/site/page?view=training_service.
- [261] „UIA’s History.” Union of International Associations (UIA). <http://www.uia.org/history>.
- [262] 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
- [263] UNESCO. 2015. *Interoperability and Retrieval*. Paris: UNESCO. <http://unesdoc.unesco.org/images/0023/002321/232199E.pdf>.
- [264] Ungváry Rudolf. 2000. „Az ETO szükségessége.” *Iskolakultúra* 10 (4): 27-31. http://real.mtak.hu/61411/1/EPA00011_iskolakultura_2000_04_027-031.pdf.

- [265] 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>.
- [266] 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>.
- [267] Ungváry Rudolf, és Vajda Erik. 2002. *Könyvtári információkeresés*. Budapest: Typotex. <http://mek.oszk.hu/05000/05030/05030.pdf>.
- [268] 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/>.
- [269] Ungváry Rudolf. 2013. „A Relex névtérkezelő.” (előadás) Networkshop 2013 konferencia, Sopron, 2013. március 26-28. <https://niif.videotorium.hu/hu/recordings/6195/relex-nevterkezelo>.
- [270] „Universal Bibliographic Repertory.” *Mundaneum*. <http://archives.mundaneum.org/en/universal-bibliographic-repertory>.
- [271] „Universella Decimalklassifikationen. Svensk Elektronisk Utgåva.” Miguel Benito. <http://www.taranco.eu/udk/>.
- [272] 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>.
- [273] „VIAF: The Virtual International Authority File.” OCLC. <https://viaf.org/>.
- [274] 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.
- [275] Wellisch, Hans H. 1969. *Bonds and Bonding of UDC Notational Symbols*. Collage Park, Md.: University of Maryland.
- [276] Wellisch, Hans. 1971. „Reorganization of the UDC.” *Nachrichten für Dokumentation* 22 (2): 55-63.
- [277] 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.
- [278] „What is UDC Sumary?” UDCC (UDC Consortium). <http://www.udcsummary.info/about.htm>.
- [279] Wimmer, F. 2000. „The Apostrophe in UDC: A critical examination of an unwelcome development.” *Extensions and Corrections to the UDC* 22: 19-27.
- [280] „XML Schema.” World Wide Web Consortium (W3C). Legutóbb módosítva 2004. október 28-án. <https://www.w3.org/XML/Schema>.
- [281] 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.
- [282] 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.



Nyilvántartási szám: DEENK/69/2018.PL
Tárgy: PhD Publikációs Lista

Jelölt: Piros Attila

Neptun kód: LT2KJ9

Doktori Iskola: Matematika- és Számítástudományok Doktori Iskola

A PhD értekezés alapjául szolgáló közlemények

Idegen nyelvű, hazai könyvrészletek (3)

1. Piros, A.: A different approach to Universal Decimal Classification in a mechanized retrieval system.
In: Proceedings of the 9th International Conference on Applied Informatics January 29 - Februar 1, 2014. Eger, Hungary Volume II [elektronikus dokumentum]. Ed.: by Kovács Emőd, Kusper Gábor, Kunkli Roland, Tómacs Tibor, Eszterházy Károly Főiskola, Eger, 29-39, 2015. ISBN: 9786155297182
2. Piros, A.: A new approach to Universal Decimal Classification as the indexing language of a mechanized information retrieval system.
In: Agria Media 2014 : XI. Információtechnikai és Oktatástechnológiai Konferencia és Kiállítás. Szerk.: Dr. Nádasi András, Eszterházy Károly Főiskola Médiainformatikai Intézet, Eger, 228-235, 2015. ISBN: 9786155621154
3. Piros, A.: Aspects of automatic information retrieval using Universal Decimal Classification.
In: Tavasz Szél 2014 Konferencia VI.: Spring Wind 2014. Szerk.: Csiszár Imre, Kőmíves Péter Miklós, DOSZ, Debrecen, 223-231, 2014. ISBN: 9786158004442

Idegen nyelvű, külföldi könyvrészletek (1)

4. Piros, A.: Automatic interpretation of complex UDC numbers: towards support for library systems.
In: Classification & authority control: expanding resource discovery: proceedings of the International UDC Seminar 2015, 29-30 October 2015, Lisbon, Portugal. Ed.: Aida Slavic, Maria Inês Cordeiro, Ergon-Verlag, Würzburg, 177-193, 2015. ISBN: 9783956501241

Magyar nyelvű tudományos közlemények hazai folyóiratban (2)

5. Piros, A.: Az ETO-jelzetek automatikus elemzésének kérdései.
Tud. műsz. táj. 62 (4), 127-141, 2015. ISSN: 0041-3917.
6. Piros, A.: Az ETO alapú számítógépes információkeresés egy új megközelítése.
Kvt. figy. 60 (2), 156-166, 2014. ISSN: 0023-3773.





Idegen nyelvű tudományos közlemények külföldi folyóiratban (1)

7. Piros, A.: The thought behind the symbol: about the automatic interpretation and representation of UDC numbers.
Knowl. organ. 44 (6), 416-424, 2017. ISSN: 0943-7444.
DOI: <http://dx.doi.org/10.5771/0943-7444-2017-6-416>
IF: 0.831 (2016)

További közlemények

Idegen nyelvű, külföldi könyvrészletek (1)

8. Piros, A.: The thought behind the symbol: about the automatic interpretation and representation of UDC numbers. Utánközlés másodközlés,
In: Faceted classification today : theory, technology and end users : proceedings of the International UDC Seminar 2017, London (UK), 14-15 September. Eds.: Aida Slavic, Claudio Gnoli, Ergon Verlag, Würzburg, 203-218, 2017. ISBN: 9783956502699

Magyar nyelvű tudományos közlemények hazai folyóiratban (2)

9. Piros, A.: A gondolat a szimbólum mögött: Az ETO-jelzetek automatikus interpretálásáról és reprezentációjáról. Utánközlés nyelvi változat,
Tud. műsz. táj. 65 (2), 86-96, 2018. ISSN: 0041-3917.
10. Piros, A.: A facettás osztályozás napjainkban: elmélet, technológia és a végfelhasználók.
Kv. kvt. 26 (11), 22-30, 2017. ISSN: 1216-6804.

A közlő folyóiratok összesített impakt faktora: 0,831

A közlő folyóiratok összesített impakt faktora (az értekezés alapjául szolgáló közleményekre):
0,831

A DEENK a Jelölt által az iDEa Tudóstérbe feltöltött adatok bibliográfiai és tudásmetriai ellenőrzését a tudományos adatbázisok és a Journal Citation Reports Impact Factor lista alapján elvégezte.





Registry number: DEENK/69/2018.PL
Subject: PhD Publikációs Lista

Candidate: Attila Piros

Neptun ID: LT2KJ9

Doctoral School: Doctoral School of Mathematical and Computational Sciences

List of publications related to the dissertation

Foreign language Hungarian book chapters (3)

1. Piros, A.: A different approach to Universal Decimal Classification in a mechanized retrieval system.
In: Proceedings of the 9th International Conference on Applied Informatics January 29 - Februar 1, 2014. Eger, Hungary Volume II [elektronikus dokumentum]. Ed.: by Kovács Emőd, Kusper Gábor, Kunkli Roland, Tómacs Tibor, Eszterházy Károly Főiskola, Eger, 29-39, 2015. ISBN: 9786155297182
2. Piros, A.: A new approach to Universal Decimal Classification as the indexing language of a mechanized information retrieval system.
In: Agria Media 2014 : XI. Információtechnikai és Oktatástechnológiai Konferencia és Kiállítás. Szerk.: Dr. Nádas András, Eszterházy Károly Főiskola Médiainformatikai Intézet, Eger, 228-235, 2015. ISBN: 9786155621154
3. Piros, A.: Aspects of automatic information retrieval using Universal Decimal Classification.
In: Tavasz Szél 2014 Konferencia VI.: Spring Wind 2014. Szerk.: Csiszár Imre, Kőmíves Péter Miklós, DOSZ, Debrecen, 223-231, 2014. ISBN: 9786158004442

Foreign language international book chapters (1)

4. Piros, A.: Automatic interpretation of complex UDC numbers: towards support for library systems.
In: Classification & authority control: expanding resource discovery: proceedings of the International UDC Seminar 2015, 29-30 October 2015, Lisbon, Portugal. Ed.: Aida Slavic, Maria Inês Cordeiro, Ergon-Verlag, Würzburg, 177-193, 2015. ISBN: 9783956501241

Hungarian scientific articles in Hungarian journals (2)

5. Piros, A.: Az ETO-jelzetek automatikus elemzésének kérdései.
Tud. műsz. táj. 62 (4), 127-141, 2015. ISSN: 0041-3917.
6. Piros, A.: Az ETO alapú számítógépes információkeresés egy új megközelítése.
Kvt. figy. 60 (2), 156-166, 2014. ISSN: 0023-3773.





Foreign language scientific articles in international journals (1)

7. Piros, A.: The thought behind the symbol: about the automatic interpretation and representation of UDC numbers.
Knowl. organ. 44 (6), 416-424, 2017. ISSN: 0943-7444.
DOI: <http://dx.doi.org/10.5771/0943-7444-2017-6-416>
IF: 0.831 (2016)

List of other publications

Foreign language international book chapters (1)

8. Piros, A.: The thought behind the symbol: about the automatic interpretation and representation of UDC numbers. Utánközlés másodközlés,
In: Faceted classification today : theory, technology and end users : proceedings of the International UDC Seminar 2017, London (UK), 14-15 September. Eds.: Aida Slavic, Claudio Gnoli, Ergon Verlag, Würzburg, 203-218, 2017. ISBN: 9783956502699

Hungarian scientific articles in Hungarian journals (2)

9. Piros, A.: A gondolat a szimbólum mögött: Az ETO-jelzetek automatikus interpretálásáról és reprezentációjáról. Utánközlés nyelvi változat,
Tud. műsz. táj. 65 (2), 86-96, 2018. ISSN: 0041-3917.
10. Piros, A.: A facettás osztályozás napjainkban: elmélet, technológia és a végfelhasználók.
Kv. kv. kv. 26 (11), 22-30, 2017. ISSN: 1216-6804.

Total IF of journals (all publications): 0,831

Total IF of journals (publications related to the dissertation): 0,831

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of Web of Science, Scopus and Journal Citation Report (Impact Factor) databases.



12 March, 2018