



**EGY ÚJ JELÖLŐ NYELV (METAADAT-RENDSZER) KIDOLGOZÁSA  
A SZÁMÍTÓGÉPES BIBLIOGRÁFIAI ADATFELDOLGOZÁSHOZ**

**SALGÁNÉ MEDVEZCKI MARIANNA**

*Doktori (PhD) értekezés*

Debreceni Egyetem  
Informatikai Kar  
Debrecen, 2005



Ezen értekezést a Debreceni Egyetem Matematika- és Számítástudományok Doktori Iskola Informatika programja keretében készítettem a Debreceni Egyetem IK doktori (PhD) fokozatának elnyerése céljából.

Debrecen, 2005. november 25.

a jelölt aláírása

Tanúsítom, hogy Salgáné Medveczki Marianna doktorjelölt 2001-2005 között a fent megnevezett Doktori Iskola Informatika programjának keretében irányításommal végezte munkáját. Az értekezésben foglalt eredményekhez a jelölt önálló alkotói tevékenységével meghatározóan hozzájárult. Az értekezés elfogadását javaslom.

Debrecen, 2005. november 25.

a témavezető aláírása



## **TARTALOMJEGYZÉK**

<b>1. BEVEZETÉS</b>	<b>1</b>
<b>2. KUTATÁSI MÓDSZEREK</b>	<b>4</b>
<b>3. IRODALMI ELŐZMÉNYEK</b>	<b>5</b>
3.1. A könyvtár jövője	5
3.2. Az információ azonosítása	7
3.2.1. Az információ azonosításának elméleti alapjai	7
3.2.2. Az információ azonosítását támogató szabványok és szabályzatok	9
3.3. Az információtechnika hatása a bibliográfiai adatfeldolgozásra	9
3.3.1. Az információtechnika fejlődésének hatása a könyvtári munkafolyamatokra	9
3.3.2. A számítógépes bibliográfiai adatformátumok megjelenése és fejlődése (MARC formátumok és társaik)	10
3.4. A magyar könyvtárak innovációja a számítógépes bibliográfiai adatfeldolgozás területén	14
3.4.1. A magyar könyvtárak automatizálásáról röviden	14
3.4.2. A magyar könyvtárakban használt számítógépes könyvtári rendszerek és az általuk meghatározott bibliográfiai adatformátumok	16
3.4.3. A HUNMARC formátum	17
3.5. Az elektronikus dokumentumok könyvtári feldolgozásának lehetőségei, a feldolgozást támogató formátumok	18
3.5.1. MARC formátum továbbfejlesztései	18
3.5.2. Metaadatokról és metaadat-rendszerekről általában	19
3.5.3. A metaadatok meghatározása a könyvtár- és információtudományban	22
3.5.4. A Dublin Core metaadatok	24
3.5.5. A bibliográfiai rekordok és formátumok funkcionális követelményei (megváltozott elvárások)	25
3.6. Metaadatok alkalmazása magyar könyvtári területen	26
3.6.1. Magyar Elektronikus Könyvtár	27
3.6.2. Elektronikus Periodika Archívum és Adatbázis	28
3.7. Rövid áttekintés az XML-ről	28

<b>3.8. Az XML felhasználási területeiről</b>	<b>32</b>
<b>4. EREDMÉNYEK</b>	<b>35</b>
<b>4.1. Az XML alkalmazásának lehetősége a könyvtárak állományának formai feldolgozásában</b>	<b>35</b>
<b>4.2. A BDML (Bibliographic Description Markup Language) fejlesztési koncepciója, eszközei</b>	<b>40</b>
<b>4.3. A BDML tervezése, tervezési elvei</b>	<b>44</b>
4.3.2. Az egyezményes jelek kezelése	48
4.3.3. A leírás nyelve és írásrendszere	51
4.3.4. A bibliográfiai adatelemeknek megfeleltetett XML-elemek és attribútumaik	52
4.3.5. A bibliográfiai adatelemek forráshelyeinek kezelése (főforrások és megengedett források)	52
4.3.6. A BDML struktúrájának áttekintése	54
<b>4.4. A BDML struktúrájának példákkal illusztrált bemutatása a könyvek monografikus szintű leírása esetén</b>	<b>56</b>
4.4.1. A leírás 1. adatcsoportja, az <area1> elem	56
4.4.2. A leírás 2. adatcsoportja, az <area2> elem	68
4.4.3. A leírás 3. adatcsoportja, az <area3> elem	71
4.4.4. A leírás 4. adatcsoportja, az <area4> elem	74
4.4.5. A leírás 5. adatcsoportja, az <area5> elem	75
4.4.6. A leírás 6. adatcsoportja, az <area6> elem	80
4.4.7. A leírás 7. adatcsoportja, az <area7> elem	82
<b>4.5. Továbbfejlesztési lehetőségek</b>	<b>85</b>
<b>5. ÖSSZEGRÉS</b>	<b>86</b>
<b>IRODALOMJEGYZÉK</b>	<b>89</b>
<b>FÜGGELÉK</b>	<b>I</b>
<b>Glosszárrium</b>	<b>I</b>
<b>1. melléklet: BDML_DTD_AREA1.dtd</b>	<b>III</b>
<b>2. melléklet: BDML_DTD_AREA2.dtd</b>	<b>VI</b>
<b>3. melléklet: BDML_DTD_AREA3.dtd</b>	<b>VII</b>
<b>4. melléklet: BDML_DTD_AREA4.dtd</b>	<b>IX</b>

<b>5. melléklet: BDML_DTD_AREA5.dtd</b>	<b>X</b>
<b>6. melléklet: BDML_DTD_AREA6.dtd</b>	<b>XIV</b>
<b>7. melléklet: BDML_DTD_AREA7.dtd</b>	<b>XV</b>
<b>8. melléklet: A teljes BDML DTD</b>	<b>XVI</b>
<b>9. melléklet: XML-dokumentum érvényességének ellenőrzése Stylus Studio 6 segítségével</b>	<b>XXV</b>
<b>10. melléklet: XML-dokumentum érvényességének ellenőrzése Stylus Studio 6 segítségével</b>	<b>XXVI</b>
<b>11. melléklet: BDML DTD készítése Stylus Studio 6 -ban</b>	<b>XXVII</b>
<b>12. melléklet: XML-dokumentum hierarchikus megjelenítése Stylus Studio 6-ban</b>	<b>XXVIII</b>
<b>13. melléklet: BDML DTD-részlet grafikus megjelenítése Stylus Studio 6 -ban</b>	<b>XXIX</b>
<b>MAGYAR NYELVŰ ÖSSZEFOGLALÁS</b>	<b>XXXI</b>
<b>ANGOL NYELVŰ ÖSSZEFOGLALÁS</b>	<b>XXXIII</b>
<b>PUBLIKÁCIÓS JEGYZÉK</b>	<b>XXXV</b>



„Két dolgot soha nem értünk meg egészen: a mindenséget s önmagunkat. Minden tudománynak elérhető tárgya csak e kettő között fekszik, s éppen ez az ok, melyért a tudományban teljes megnyugvást nem találhatunk. Mert miután minden komoly gondolkozás azon kérdéshez vezet, mily összefüggésben áll azon tárgy, melyről elmélkedünk, az egésszel s minmagunkkal, természetes, hogy minden tudományunk kétellyel vagy hittel végződik.”

- Eötvös József -

## **1. BEVEZETÉS**

A könyvtáros szakemberek tevékenysége (a könyvtártudomány és az általános információtudomány) hosszú-hosszú évekkel ezelőtt összekapcsolódott az információrögzítés és információkeresés problémájával, az információforrások bibliográfiai leírásával és (formai – tartalmi) feldolgozásával, amióta csak könyvtárak léteznek.

A telekommunikációs eszközök fejlődése és könyvtárakban történt megjelenése már néhány évtizede hatást gyakorol a könyvtárak információfeldolgozási és – szervezési technikáira. Az elmúlt húsz, harminc évben úgy tűnt, hogy a számítógépes könyvtári integrált rendszerek megbízható megoldást jelentenek a könyvtárak mindennapos problémáira.

Az 1990-es évek elején azonban egy újabb mérföldkő következett, a Tim Berners-Lee által kitalált web, amely az eredetileg elosztott dokumentumkezelő rendszerből az írásos kommunikáció legforradalmibb eszköze lett a nyomtatás megjelenése óta. (És a legnagyobb információs káosz is a civilizáció történetében.) A web kínálta újfajta technológiák új lehetőségeket kínálnak a könyvtárak és információs intézmények számára is, amelyekkel élniük – tudni – kell. A könyvtári integrált rendszereket új modulokkal, funkciókkal, formátumokkal kell kibővíteni és felvértezni, hogy képesek legyenek a webről elektronikus dokumentumokat, felhasználói kéréseket és kérdéseket fogadni, valamint képesek legyenek nyitni szolgáltatásaikat a web széles használói közönsége felé.

Ma a világon elterjedt és használt integrált könyvtári rendszerek többsége a MARC (Machine Readable Cataloguing) formátumot támogatja, amelynek kialakítása még jóval a web megjelenése előtt, az 1960-as években kezdődött meg a könyvtári szoftverek egy zártabb világa számára.

Szükségesnek látszik az, hogy olyan új bibliográfiai adtaformátumokról és ezek kialakítási lehetőségeiről gondolkodjunk, amelyek képesek alkalmazkodni az új körülményekhez, illetve képesek a webes technológiák nyílt szabványai és nyílt forráskódú szoftverei által nyújtott új lehetőségeket minél jobban kiaknázni.

Dolgozatom indításként röviden áttekintem a számítógépes bibliográfiai adatformátumok kialakulásának és fejlődésének legfontosabb állomásait, összefoglalom a nemzetközi és hazai tendenciákat ezen a területen.

A dolgozat második részében egy **új, saját fejlesztésű, XML-alapú jelölő nyelv** bemutatása következik **a számítógépes bibliográfiai adatfeldolgozás számára**, amely a **BDML (Bibliographic Description Markup Language)** nevet kapta. Indításként felvázolom az XML alkalmazásának azon lehetőségeit a könyvtári dokumentumok formai feldolgozásában, amelyek - véleményem szerint - pozitív előrelépést jelenthetnek ezen a területen. Majd a következő fejezetekben a BDML tervezéséről, fejlesztési elveiről szólok. Ezt követően konkrét, bibliográfiai rekordokat tartalmazó példák segítségével kívánom bemutatni a BDML nyelvet leíró DTD (Document Type Definition) legfontosabbnak ítélt részeit. A tárgyalás során felhasznált bibliográfiai leírási példák kifejezetten úgy kerültek összeválogatásra, hogy előforduljanak ezekben azok a leírási nehézségek, amelyek kezelése a jelenleg, magyar könyvtárakban használt, számítógépes bibliográfiai adatformátumok számára problémát okoz. Ilyen módon a BDML struktúrájának bemutatása egyben egy összehasonlító vizsgálat is, amelynek során az egyes bibliográfiai leírási problémák kezelhetőségét megnézem a USMARC és a HUNMARC formátumokban (amely két formátum alkalmazása a hazai könyvtárakban a legelterjedtebb), illetve bemutatom ezen problémákra a BDML nyújtotta lehetőségeket és pozitív válaszokat is.

Végezetül, a dolgozat zárásaként a BDML nyelv néhány továbbfejlesztési lehetőségét veszem sorra, amelyek a további kutatásaimat fogják meghatározni ebben a témakörben.

A BDML DTD kialakításánál szerzett tapasztalatok hasznosítására, és a további fejlesztésre két aktuálisan futó projektben is lehetőség nyílik. Az egyik az EMIR (Elektronikus Magyar Irodalom) projekt, amely az Eötvös Loránd Tudományegyetem Bölcsészeti Informatika Önálló Program, a Veszprémi Egyetem Magyar Irodalomtudományi Tanszékének és a Debreceni Egyetem könyvtár tanszékének együttműködése. A projekt célja a kortárs magyar irodalom komplex és sokoldalú online feldolgozása és a nagyközönség számára elérhetővé tétele, egyben egy nagyobb léptékű, összefüggő, modulárisan építkező, közös hozzáférésű, elosztott rendszerre épülő szolgáltatás létrehozása, amely egy rendszerben kapcsolná össze a digitálisan feldolgozott katalógusrendszereket, folyóiratokat, kölcsönzési és vásárlási lehetőségeket. A másik kezdeményezés a Debreceni Egyetem Zeneműtárához és a Debreceni Egyetem könyvtár tanszékéhez kapcsolódik. A cél ebben az esetben egy digitális zeneműtár létrehozása, amely nemcsak a hangzó dokumentumok bibliográfiai adatait tárolná elektronikus formában, hanem magukat a zenei dokumentumokat is.

A dolgozat rövid összeggzéssel zárul.

A dolgozatban, a téma tárgyalásakor rövidített formában alkalmazott kifejezések feloldásai a dolgozat végén, a Függelék Glosszáriumában található.

## 2. KUTATÁSI MÓDSZEREK

A BDML kialakításánál a hatályban lévő nemzetközi és nemzeti bibliográfiai leírási szabványok és szabályzatok messzemenő figyelembevételére törekedtem.

A következő ajánlások és szabványok szolgálták a tervezés alapjául:

- ISBD / G(eneral): a nemzetközi szabványos bibliográfiai leírás általános szabályai [26];
- ISBD / M(onographic publications): ajánlás a könyvek nemzetközi szabványos bibliográfiai leírására [27];
- MSZ 3424/1. Bibliográfiai leírás. Könyvek : a könyvek bibliográfiai leírásának magyar szabványa [40];
- Dublin Core Metadata Initiative [13];
- FRBR: IFLA tanulmány a bibliográfiai rekordok funkcionális követelményeiről [22].

A BDML DTD és a BDML-alapú bibliográfiai rekord példák készítésében, ezek jól formázottságának és érvényességének ellenőrzésében a Stylus Studio 6 XML Home Edition környezet volt a segítségemre.

### **3. IRODALMI ELŐZMÉNYEK**

A könyvtárak az új kommunikációs technológiákat alkalmazva folyamatosan új kérdésekkel, lehetőségekkel és kihívásokkal találják szembe magukat. Dolgozatom e fejezetében elektronikus korunk könyvtárakra gyakorolt hatásának legfontosabb aspektusait kívánom kiemelni és összegezni elsősorban az automatizált bibliográfiai adatfeldolgozás szempontjából, valamint áttekinteni a számítógépes bibliográfiai adatformátumok fejlődésének legfontosabb eredményeit. Az első néhány alfejezet a nemzetközi tendenciákat mutatja be, ezt követően a magyar könyvtáros szakma ilyen irányú törekvéseit írom le röviden.

#### **3.1. A könyvtár jövője**

Napjaink könyvtárügyét az erőteljes kettősség jellemzi, amellyel az elektronikus kor jelenlegi szakaszában, az új kommunikációs formák megjelenésével a Gutenberg-galaxist őrző, de a Neumann-galaxis kihívásaira reagáló könyvtáraknak meg kell birkózniuk.

Az információ- és könyvtártudomány, vagy mondhatjuk könyvtár-informatikának is, a nyugat-európai országokban, valamint a tengerentúlon évtizedes hagyományokkal rendelkező, tudományos eredményeket felmutató szakterületnek számít. Interdiszciplináris kapcsolatban áll más tudományterületekkel is (például az informatikával, kommunikációtudománnyal stb.), bár vannak, akik e „multi-szerepét” kétségbe vonják.

A könyvtárak és a könyvtárosság jövőjéről – érthető okokból – igen sok tanulmány jelent meg az utóbbi években. Az egyik közismert alapműben *Buckland, M.* (1992) – magyar fordításban is megjelent könyvében – a könyvtárak fejlődésének három szakaszát írta le: az első szakaszban a könyvtári

dokumentumok papír alapúak, a könyvtári segédeszközök, munkafolyamatok hasonlóan papírhoz kötöttek. A fejlődés következő fokozatát automatizált könyvtárnak nevezte. Ebben a könyvtárban a dokumentumok még döntően papír alapúak, de a könyvtári munkafolyamatok már automatizáltak. A harmadik szakaszban az elektronikus könyvtár fogalmát annak az állapotnak a leírására használja, amelyben a dokumentumokat elektronikus formában tárolják, tehát nem papíron vagy egyéb, helyhez kötött médiumon és a munkafolyamatok, illetve a segédeszközök természetesen szintén automatizáltak [6]. Buckland megállapításait a könyvtáros szakma ma is helytállóan ítéli.

A papír nélküli könyvtárról szóló negatív jóslatok nyilvánvalóan nem váltak be, és nem valószínű, hogy a jövő könyvtáraiból hiányozni fognak a papír alapú dokumentumok. Ebből következően állítható, hogy nem az elektronikus könyvtárak megteremtése a cél, hanem az olyan integrált könyvtáraké, amelyekben egyaránt megtalálhatók a hagyományos publikációk és az elektronikus dokumentumok, amelyekben a felhasználó hozzájut a fizikai valójában jelenlévő állományhoz éppúgy, mint – a hálózat révén – a világ bármely sarkában vagy éppen hálózaton található elektronikus és nem-elektronikus forrásokhoz. Vagyis az igazi kihívás – az integrálás.

Az integrált könyvtár fogalmát *Poprády G.* (2003) a könyvtári trendekről készült tanulmányában így határozza meg: „Az integrált könyvtár olyan osztott információs rendszer, amely mind a hagyományos, mind a rendkívül változatos elektronikus dokumentumokat tekintve biztonságos tárolást, a felhasználóknak pedig kényelmes és hatékony hozzáférést biztosít a legfejlettebb digitális technológiákkal és adatátviteli hálózatokkal támogatott könyvtári munkafolyamatok révén” [43].

## 3.2. Az információ azonosítása

### 3.2.1. Az információ azonosításának elméleti alapjai

*Burkart, M.* (1990) szerint a könyvtárnak, mint a társadalom egyik információs alrendszerének elsődleges feladata a kommunikáció, az információcsere, információáramlás biztosítása olyan partnerek között, akik ugyanazon téma iránt érdeklődnek, de nem tudnak közvetlen kommunikációs kapcsolatba lépni egymással [8].

A könyvtár feladata, hogy a társadalom információkincsét, azaz a közös tudásvagyont megőrizze és gondoskodjon arról, hogy ez a vagyon bekerüljön a kommunikáció csatornájába.

A könyvtár inputját tehát a dokumentumokban (információforrásokban) megjelenő információ/tudás képezi. Bármilyen eljárást, technológiát alkalmaznak is a könyvtárak a feltárási folyamatok során, első lépésként az információnak, az ismeretnek azt a mennyiségét/részét kell meghatározni, kiválasztani, amely a folyamatok további szakaszaiban egy egységként jelenik meg.

Az információtudomány szűken értelmezett feladata nem más, mint hogy létrehozza az eredeti szövegek reprezentációját, szurrogátumát. Ez a szurrogátum kerül az adatbázisokba, katalógusokba, információ-kereséskor a szurrogátum jelenik meg találatként, s ebből lehet eljutni az eredeti szöveghez. A tömörítés mértékétől függően a szurrogátum sokféle lehet. Eredeti dokumentumot képviselhet akár egy kódszám is, bár ennek információtartalma minimális. Másik végletként megvalósulhat a teljes szöveg tárolása, amely a továbbiakban kiegészülhet más adatokkal is. Egy tipikus szurrogátum részei *Horváth T.* (2001) összefoglalását alkalmazva a következők [20]:

- azonosító(k);
- **bibliográfiai leírás;**
- szabad szöveges ismertetés, szöveges tömörítés, kivonat referátum, annotáció;

- generalizáló osztályozási jelzet, kód;
- individualizáló osztályozási jelzet;
- felhasználói kódok;
- alkalmazási kódok.

A könyvtár a feltárás munkafolyamata során a dokumentumokról, információforrásokról egy reprezentációt készít, amelynek egyik legelterjedtebb formája a bibliográfiai leírás, illetve a bibliográfiai tétel (számítógépes környezetben: bibliográfiai rekord). A bibliográfiai tétel önmagában is megjelenhet szurrogátumként.

A feltárás célja mindig az, hogy a majdani kommunikációs folyamatban a felhasználó a leggyorsabban és leghatékonyabban találja meg a releváns információt problémája megoldásához, ezért a könyvtáraknak gondoskodniuk kell a használói igényekhez kapcsolódó információk magas színvonalú, korszerű feldolgozásáról.

A könyvtárak automatizálásával együtt a hagyományos bibliográfiai feltárási eszközök újabb, hatékonyabb eszközökkel, megoldásokkal cserélődnek le. Változik a technológia és változik a szellemi eszközök (szabványok, szabályzatok, metaadatok stb.) köre is.

*Hagler, R.* (1997) szerint a bibliográfiai adatok illetve adatelemek jelentik azokat a „legkisebb építőköveket”, amelyek együttesen azonosítanak egy dokumentumot vagy a dokumentum bibliográfiai rekordját. Ezek a – szabványosan összeállított és használt – építőelemek összefüggő és széles körű struktúrát alkotnak, bármely dokumentum legeredményesebb meghatározásához, és a bibliográfiai adatbázis leghatékonyabb kezeléséhez [19].

### **3.2.2. Az információ azonosítását támogató szabványok és szabályzatok**

Nemzetközi és nemzeti szabványok egész sora rendelkezik ezekről a kérdésekről.

Az IFLA koordinálásával és közreműködésével számos nemzetközi program indult el, s eredményükként készültek el az ISBD-ajánlások elsősorban a bibliográfiai információk nemzetközi felhasználásának és a bibliográfiai adatok cseréjének segítésére, valamint további nemzetközi és nemzeti szabványok, szakmai szabályzatokat dolgoztak ki. Az ISBD-hez hasonló szerepet hivatott betölteni, de a besorolási adatok terén, az 1984-ben megjelent GARE útmutató, amelyet az IFLA katalogizálási és információtechnológiai bizottsága dolgozott ki, megcélozva egy nemzetközi authority-rendszer (egységesített besorolási adatállomány) kialakítását. A GARE meghatározza az egységesített besorolási adatok adatalemeit, azok sorrendjét, központosítási jeleit, az adathoz tartozó utalókat, kapcsolódó adatokat és megjegyzéseket, azaz egy nemzetközi szinten egységes szabályrendszer felállítására vállalkozik [43].

### **3.3. Az információtechnika hatása a bibliográfiai adatfeldolgozásra**

#### **3.3.1. Az információtechnika fejlődésének hatása a könyvtári munkafolyamatokra**

Az információtechnika változásai, mint például az internet használata; bibliográfiai adatokat kezelő könyvtári számítógépes rendszerek elterjedése, a helyi és nemzeti könyvtári hálózatok létrejötte, az elektronikus központi katalógusok megjelenése, az országos közös katalogizálás megszervezése, az információkeresési és dokumentumtovábbítási protokollok kidolgozása, a világszabvánnyá váló Z39.50-en alapuló ügyfélkiszolgáló technológia

megteremtése; a digitalizálás térhódítása, s a retrospektív konverziók (hagyományosan már feltárt könyvtári állományrészek számítógépes újrakatalogizálása) megindulása – hogy csak néhány fontos tényezőt emeljünk ki – alapvetően átalakították a könyvtárak működését.

A digitalizálás eredményeként folyamatosan bővül a könyvtárakban gyűjtendő és feldolgozandó dokumentumok köre, ami azt is jelenti, hogy megnő a katalogizálást végzők szerepe a világhálón elérhető információforrások feldolgozásában, hiszen a helyi és távoli elérésű elektronikus dokumentumokat ugyanúgy be kell vonni a bibliográfiai számbavétel körébe, mint a hagyományos dokumentumokat.

A retrospektív konverzió révén megvalósulhat a széles körű hozzáférés a korábbi cédulakatalogusokhoz, azaz a korábban hagyományos eszközökkel feltárt könyvtári állományrészek sem válnak „holttá” elektronikus korunkban.

Egy mű tetszőleges megjelenési formájának, illetve e megjelenési forma tartalmának és bibliográfiai kapcsolatainak integrált feltárásához szükségessé vált a meglévő katalogizálási szabványok és adatsere-formátumok továbbfejlesztése, valamint a katalogizálási szabályok újragondolása.

### **3.3.2. A számítógépes bibliográfiai adatformátumok megjelenése és fejlődése (MARC formátumok és társaik)**

Az első bibliográfiai adatformátumok, amelyek a bibliográfiai és kapcsolódó információk számítógépes formában történő cseréjét célozták különböző típusú számítógépes rendszerek között, több mint 40 évvel ezelőtt készültek.

Az amerikai Library of Congress 1962-ben indított MARC projektje új fejezetet indított el a géppel olvasható katalogizálás területén [33]. A projekt eredményeként készült el az első MARC formátum, az LCMARC. Ez alapján különféle MARC formátum változatokat, verziókat fejlesztettek ki az egyes országok és bibliográfiai rendszerek eltérő gyakorlatának és szükségletének megfelelően, többek között például a USMARC-ot az Egyesült Államokban, a

CANMARC-ot Kanadában (ennek alapján készült a MARC21 formátum 1999-ben), a UKMARC-ot Nagy-Britanniában, valamint a HUNMARC-ot Magyarországon 1994-ben.

Az IFLA egységesítési törekvései a katalogizálás területén megkívánták, hogy az eredményes információcsere érdekében ne csak a bibliográfiai tételek tartalmával és formájával foglalkozzanak, hanem a gépi rekordok formátumával is. Az 1981-ben megjelent ISO 2709 szabvány rögzítette a szabványos rekordszerkezetet [29], és a továbbiakban minden MARC formátumot erre a szabványra építettek, megnövelve ezzel a könyvtári rendszerek közötti rekordcsere lehetőségét.

A nemzeti MARC formátumokat az egyes országok katalogizálási szabályzatainak megfelelően alakították ki, így – bár sok szempontból hasonlítanak egymásra – lényeges különbségek is vannak közöttük. Ez okozza azt, hogy az egyes MARC formátumok között nem valósulhat meg közvetlen adatcsere, hanem konverziós programokra van szükség [54]. Az országos jelentőségű MARC-okon kívül egyes szoftverek vagy rendszerek saját formátumokat is kialakítottak, amelyek tovább színesítették/színesítik a használatban lévő formátumok körét.

A hatékony információcsere érdekében a katalogizálási szabályok nemzetközi egységesítésének az adatcsere-formátumokra is ki kellett terjednie. Szükségessé vált egy nemzetközi csereformátum kidolgozása, amely közvetítőként működhet az egyes MARC formátumok között, és segítheti a konverziós programok kialakítását. Ez lett a UNIMARC (Universal MARC) [52, 53], amelyet az IFLA elsősorban a nemzeti könyvtári rendszerek közötti adatcserére szánt, bár több európai országban is ezt a formátumot alkalmazzák, mint nemzeti MARC formátumot is, például Horvátországban, Portugáliában, Görögországban és Olaszországban, valamint említsünk meg példaként egy nem európai térségbe tartozó, de UNIMARC-ot használó országot, ez pedig India.

Az országonként eltérő könyvtári szabványok, illetve az egyes könyvtárakban használt különböző számítógépes könyvtári rendszerek nyomán kialakult nemzeti MARC formátumok közötti konverzió a katalogizálás jelenlegi legnagyobb problémája. Az IFLA 1987-ben indította el a UBCIM programját, melynek

célkitűzései (a bibliográfiai számbavételt biztosító szabványok és rendszerek nemzeti szintű fejlesztésének, valamint a bibliográfiai adatok nemzetközi szintű cseréjének koordinálása, a géppel olvasható bibliográfiai adatok cseréjének szabványosítása, a UNIMARC formátum fejlesztése) illeszkednek a kialakult kedvezőtlen helyzethez.

Léteznek más egységesítési törekvések is, amelyek a különböző formátumok összehangolását célozzák.

Az UNESCO által 1974-ben közreadott UNISIST formátum a referáló és indexelő szolgáltatások adatbázisai közötti információcsere céljait szolgálja. A nemzetközi egységesítés érdekében 1978-ban megkezdődött a UNIMARC és UNISIST összehangolása, és kidolgoztak egy közös adatcsere-formátumot, a CCF-t. Célja az adatkommunikáció a könyvtárak és a tájékoztató szolgáltatások között oly módon, hogy meghatározták a minden rekordban kötelező és a dokumentum azonosításához nélkülözhetetlen adatelemek összességét.

A MARC formátumok egységesítése nemcsak az IFLA keretében folyik. A MARC-harmonizációnak elkeresztelt együttműködési programtól azt várják, hogy megkönnyítse a rekordok átvételén alapuló katalogizálást, valamint a konverzió redukálása, illetve elmaradása következtében egymás rekordjainak gyorsabb hozzáférhetőségét, s így olcsóbb katalogizálást. A program keretében az AACR katalogizálási szabványra támaszkodó USMARC, CANMARC, UKMARC formátumok összehangolására törekednek.

USEMARCON az Európai Unió Könyvtári Programján belül 1995-ben elindított program, amelynek célja egy generikus MARC konverziós rendszer létrehozása. A jelenleg használt mintegy 50 különféle MARC formátum között tett/tesz különféle fokú konverziót lehetővé, amely folyamatokban a UNIMARC formátumot használja belső (köztes) formátumként. A USMARCON során használt szoftver első változata 1996. februárjában készült el a francia Jouve cég fejlesztésében. Ennek a konverziós alkalmazásnak a továbbfejlesztése azóta is folyamatos a British Library és CERL koordinálásával, a közös munka során a USMARCON jelentése Universal MARC Record Converterre módosult. A USMARCON legutolsó verziója, a v1.51 a British Library és a Helsinki University

Library közös fejlesztése, az alkalmazás és a kapcsolódó dokumentáció a könyvtáros szakma számára szabadon hozzáférhető [54].

Az utóbbi néhány évben erőteljes elmozdulás érezhető a UNIMARC formátum felől a MARC21 formátum felé. A MARC21 a Library of Congress és a National Library of Canada együttműködésének eredménye, amelyet saját, elfogadott nemzeti formátumaik összefésülésével, harmonizációjával hoztak létre [33]. A MARC21 azóta élvezi más nemzeti könyvtárak, online bibliográfiai szolgáltatások támogatását és végfelhasználók támogatását, valamint a könyvtári rendszerek többsége is képes kezelni. Ennek köszönhetően úgy tűnik, hogy a MARC21 fokozatosan átveszi majd a UNIMARC szerepét a nemzetközi együttműködés feladataiban.

A nemzetközi forrásmegosztás célját két kifejezéssel lehet tömören megfogalmazni: „megtalálni” és „megkapni”. A végfelhasználót valójában jobban érdekli, hogy mikor kapja meg a számára szükséges (releváns) információt, mint hogy hogyan és honnan. Bár ez egyszerűen hangzik, megvalósítása meglehetősen bonyolult. A globális könyvtári forrásmegosztás négy fő akadályként a következő tényezőket említhetjük:

- az internet és a felhasználói viselkedés;
- a nyelvi és betűproblémák;
- az online metaadatok kritikus tömegének hiánya;
- a könyvtárak közötti nemzetközi együttműködési hagyományok hiánya.

E problémák leküzdésére az OCLC különböző projektek indítását kezdeményezte. Az OCLC több mint húsz évvel ezelőtt kezdte hálózatépítő tevékenységét a különböző helyeken lévő források összekapcsolására.

A számítógépes bibliográfiai adatformátumokra fókuszálva az említett akadályok közül a nyelvi és betűproblémák, valamint az online metaadatok hiánya okozza a legnagyobb gondot.

A nyelvi és betűproblémák az informatikában használt karakterkészletek sokféleségéből fakad. A világon körülbelül négyszáz nyelven megjelenő dokumentumok betűkészletének kódolására sokféle szabvány és gyakorlat létezik, vannak közöttük nyelvre épülők, és léteznek írásjelekre (például latin, cirill betű) alapozottak. Az első szabvány az ASCII karakterkészlet volt, amely az angol nyelvű írógép-billentyűzeten lévő betűket tartalmazta, kibővített változatába bekerültek a diakritikus jelek is. A legtöbb nyugati könyvtárban a nem latin betűs írásjeleket transliterálják, ennek során értékes adatok veszhetnek el. Sokéves vita után a világ írott nyelveinek nagy részét támogató UNICODE [51] rendszer használata mellett döntöttek. Az UNICODE-ot a Nemzetközi Szabványügyi Szervezet, az ISO is elfogadta, beépítette a vonatkozó szabványba. A fejlesztést a legnagyobb számítástechnikai cégek is támogatják, így várhatóan nem lesz majd szükség a szoftvercsomagok különböző nyelvi változatainál sem kiegészítő munkákra.

(A metaadatokról a 3.5. alfejezetben részletesen szólok.)

A nemzeti formátumok közelítését az említett törekvéseken túlmenően számos más projekt is sürgeti.

### **3.4. A magyar könyvtárak innovációja a számítógépes bibliográfiai adatfeldolgozás területén**

#### **3.4.1. A magyar könyvtárak automatizálásáról röviden**

Magyarországon alig múlt tíz-tizenöt éve, hogy a nagyobb könyvtárak megvásárolták első integrált számítógépes rendszerüket, s lassan valamennyi könyvtártípusban megkezdődött a munkafolyamatok automatizálása. A könyvtárak erősödő informatikai irányultságának köszönhetően a nyugati gyakorlathoz képest fennálló, mintegy 30 éves kezdeti lemaradásunk mára már jelentősen csökkent.

Magyarországon a könyvtárgépesítés komoly lendületet a rendszerváltás után, az 1990-es években kapott. A tudományos - felsőoktatási és akadémiai - könyvtárak léptek leghamarabb. Nemzetközi és hazai szakértők bevonásával megtörtént az egyetemi könyvtárak fejlesztési stratégiájának megfogalmazása, rendszerterv született a könyvtárak széles körét bevonó országos szakirodalmi információs rendszer létrehozására, és a közgyűjtemények modern és új feladatairól, biztosításuk feltételeiről is új szellemű törvény rendelkezik.

A fejlesztések megvalósításába - a pályázati formákat preferálva - hazai és külföldi források egyaránt beléptek. A kilencvenes évek elején az e területeken is jelentkező - egyedi s általában egyszeri - fejlesztéseket támogató külföldi (PHARE, Mellon Alapítvány, TEMPUS stb.) források mellett a későbbiekben megjelentek a hazaiak is (FEFA, OMF B IKTA, NKA), amelyek hosszabb távra terveznek rendszeres fejlesztéseket. Jelentős támogatásként indult az Nemzeti Kulturális Örökség Minisztériuma telematikai fejlesztési programja is. A közgyűjtemények, különösen a könyvtárak fejlesztésében meghatározó és kiemelkedő jelentősége volt (és van) az Információs Infrastruktúra Fejlesztési Program (IIFP), majd Nemzeti Információs Infrastruktúra Fejlesztési Programnak (NIIFP), amely a hazai kutatási szféra információs infrastruktúrájának fejlesztését és alkalmazásait támogatta/támogatja, amelybe sikerült bekapcsolni a majdnem teljes közgyűjteményi rendszert. A technológia, technika, eszközfejlesztés támogatása mellett - s ez külön érdeme - nagy gondot fordított az elektronikus tartalom-előállításra is.

A nemzetközi trendeket követő fejlesztések fő irányai az alábbiak voltak: az egyedi intézményeket tekintve a számítógépes, automatizált katalogizálás kialakítása; a bibliográfiai feltárás, adatszétsugárzás elektronikus megoldása; az új típusú adathordozók nyújtotta új információs források rendszeres gyűjtése és szolgáltatása; elektronikus dokumentumok előállítása, gyűjtése és szolgáltatása, hagyományos dokumentumok felállítása. Az intézményrendszer egészét tekintve pedig az együttműködés:

- az elektronikus bibliográfiai adatfeldolgozás nemzeti szintű megteremtésében és működtetésében (nemzeti közös katalogizálás – MOKKA-projekt, központi elektronikus katalógus létrehozása);

- a (hagyományos és elektronikus) dokumentumok tartalmi feltárásában, az információkeresés új megoldásainak kimunkálásában, közös felületek kialakításában;
- a retrospektív bibliográfiai feldolgozásban (a hagyományos katalógusrekordok elektronikus konverziójának megoldásaiban);
- a virtuális könyvtárak létrehozásában;
- az elektronikus könyvtárak megteremtésében (Magyar Elektronikus Könyvtár, stb.);
- a nemzeti digitális könyvtár létrehozásában.

### **3.4.2. A magyar könyvtárakban használt számítógépes könyvtári rendszerek és az általuk meghatározott bibliográfiai adatformátumok**

Mára lényegében minden jelentősebb felsőoktatási, megyei és városi könyvtárban találunk valamilyen rendszert, sőt a kisebb könyvtárak is rendelkeznek olyan szoftverekkel, amelyek a legfontosabb könyvtári munkafolyamatok automatizálását legalább részben segítik.

A magyar könyvtárakban használt számítógépes könyvtári rendszerek egyik nagy csoportját azok a külföldön kifejlesztett, integrált rendszerek jelentik, amelyeket a nagyobb felsőoktatási és szakkönyvtárak vásároltak meg és vettek használatba, mint például az Aleph, TINLIB, a ma már magyar fejlesztésű Corvina (korábban Voyager), Oracle Libraries, Horizon (korábban Dynix), Amicus.

A másik jelentős könyvtári szoftver-bázist a Fővárosi Szabó Ervin Könyvtár és nyolc nagy megyei könyvtár vezetésével az INFOKER cég által a közkönyvtárak számára kifejlesztett Textar, majd TextLib adja, amelyet nagyon sok közkönyvtárban megvettek és telepítettek.

Ezen túlmenően további kisebb méretű, hazai fejlesztésű könyvtári rendszerekkel találkozhatunk a kisebb könyvtárakban, mint például a Bibliotéka (cikk-, mikrofilm-, kép-, irodalmielemzés-feltáró programcsomag), ELKA, GEORGE, HunTéka, KisTéka, koLIBRI (könyvtárállomány nyilvántartó program), S-LIB, SZIRÉN (iskolai könyvtári program), NanLIB, Otéka, SZIKLA stb.

Ebből következően az egyes könyvtárakban más és más, MARC és nem MARC alapú formátumokat használnak (a magyar bibliográfiai szabványokon alapuló, a magyar nemzeti könyvtár által kidolgozott HUNMARC-nak komoly ellenfele a legtöbb magyar könyvtárban alkalmazott USMARC alapú rendszerek tekintélyes rekordszáma).

### **3.4.3. A HUNMARC formátum**

Magyarország hivatalos MARC formátuma a HUNMARC, amely a nemzeti MARC formátumok rendszeréhez csatlakozva, az MSZ ISO 2709 szabvány által megadott keretekben határozza meg a bibliográfiai adatok tartalmi jelölőit és egyéb jellemzőit. Az adatcserében résztvevő adatok közlésmódjára a magyar könyvtári szabványok, szabályzatok előírásai az irányadók. A HUNMARC formátum kialakítása figyelembe vette a nemzetközi szolgáltatásokban elérhető UNIMARC vagy USMARC alapú nemzeti MARC formátumok gyakorlatát, valamint támaszkodik a MARC21 bibliográfiai formátum útmutatóira. A HUNMARC formátumot az Országos Széchényi Könyvtár készítette. Első kiadása 1994-ben jelent meg, amely a könyvek és az időszaki kiadványok feldolgozására vonatkozott. Az 1998-as kiadása már valamennyi dokumentumtípusra kiterjedő előírásokat tartalmazott. A HUNMARC kézikönyv legfrissebb, revideált változata 2002-ben jelent meg [21].

A napjainkra kialakult helyzetről elmondható, hogy a magyar könyvtárakban használt könyvtári rendszerek nagy többsége (beleértve a hazai és külföldi fejlesztésű könyvtári szoftvereket egyaránt) képes bibliográfiai rekordokat HUNMARC formátumban küldeni és fogadni. Abban az esetben, ha az adott könyvtári rendszernek nem a HUNMARC a belső (tárolási) formátuma, megfelelő konverziós programok segítségével éri el mindezt.

### **3.5. Az elektronikus dokumentumok könyvtári feldolgozásának lehetőségei, a feldolgozást támogató formátumok**

Napjainkban már nem kétséges, hogy a könyvtárak információ feltáró munkájának az elektronizáció korában megjelent új dokumentumtípusokra, az elektronikus dokumentumokra (és új információhordozókra) is ki kell terjednie. Azokat a formátumbővívéseket és -továbbfejlesztéseket, amelyek az elektronikus dokumentumok katalogizálhatósága érdekében történtek, a létező formátumok többségén az 1990-es évek folyamán végezték el.

#### **3.5.1. MARC formátum továbbfejlesztései**

A MARC formátum egyik komoly hátrányát jelenti, hogy az ember számára olvashatatlan, a formátum egyes szerkezeti részeit, adatmezőit, almezőit, indikátorait számjegyekből álló kódértékek azonosítják, és rendkívül összetett programokra van szükség a rekordokban tárolt tartalom feldolgozásához.

A MARC formátumok további fejlesztésének következő jelentős állomását az 1990-es évek közepén a Network Development and MARC Standards Office által kifejlesztett két SGML DTD jelentette, amely támogatta az adatkonverziót a MARC adatstruktúrából SGML-be és vissza adatvesztés nélkül. 2001-ben az előbb említett SGML DTD-eket teljes mértékben átkonvertálták XML nyelvre. Ez az XML alapú DTD a MARC formátum szerkezetén nem változtatott, sőt a DTD-ben szereplő egyes elemek elnevezéseiben is megmaradtak a korábbi számjegy alapú azonosítók [7]. Ez utóbbi területen változás 2002-ben következett be, amikor is a Library of Congress elkészített egy újabb XML sémát a bibliográfiai elemek számára. Ez lett a MODS [38], amely egyrészt a MARC adatmezőknek csak egy részhalmazát tartalmazza (tehát némi egyszerűsítés is történt), illetve felhasználóbarátabbá vált annak következtében, hogy az egyes bibliográfiai elemek természetes nyelvű azonosítókat kaptak.

Az elmúlt időszakban, legalábbis a könyvtári szakmában evidenciának tűnt, hogy a különféle dokumentumtípusok leírásának természetes kerete, szerkezete a MARC formátum. S bár a könyvtári rendszerek évtizedek óta a MARC rekordok készítését és karbantartását támogatják, viszonylag új keletű a más adatformátumok kialakításának, létrehozásának az igénye.

Érdemes még talán megemlíteni azt, hogy a nemzetközi szakirodalomban, elsősorban az Egyesült Államokban, a közelmúltban megjelentek olyan vélemények is, amelyek a MARC formátum közelgő „halálát” jósolták [49]. Más szakemberek viszont a MARC formátum hosszútávú együttélését vetítik előre a webes bibliográfiai adatformátumokkal. A legfontosabb érvei ez utóbbi elképzelésnek, a világszerte MARC formátumban hozzáférhető bibliográfiai adatok óriási méretű adatbázisainak ténye, illetve az, hogy a jelenleg létező metaadat-rendszerek egyike sem képes felvenni a versenyt a MARC formátum nyújtotta bibliográfiai feltárési minőséggel.

A könyvtárak jövőjét inkább abban lehet/kell látni, hogy képesek legyenek elfogadni és számítógépes rendszerükbe fogadni bibliográfiai rekordstruktúrák különböző formáit, a Dublin Core-től az ONIX-ig, az ONIX-tól a BDML-ig, mindazokat, amelyek használható információkat tartalmaznak. A legfontosabb tehát a különböző bibliográfiai környezetek és formátumok közötti átjárhatóság biztosításának kérdése.

### **3.5.2. Metaadatokról és metaadat-rendszerekről általában**

Bár a metaadatok problematikája a könyvtár- és információtudomány jelenlegi érdeklődésének középpontjában áll, maga a terminus technicus a számítástechnika tudományából származik.

A metaadat adat az adatról. Specifikusan, a metaadat nem más, mint olyan adat, amely az információs objektumokat reprezentáló forrásadatokat jellemzi, viszonyaikat leírja, elősegíti az objektumok fellelhetőségét és hatékony felhasználását, s ezek az objektumok (információforrások) akár fizikai, akár elektronikus megjelenésűek lehetnek [9, 10].

Amíg a számítógépekkel feldolgozott strukturált metaadat-rendszerek még viszonylag új keletűek a könyvtári területen, addig a metaadatok alapvető koncepciója jól ismert a könyvtáros szakemberek számára nagy méretű információs gyűjteményeik kezeléséből és használatából fakadóan. A könyvtári cédulakatalógusok mindenki által jól ismert példái a metaadatoknak.

A metaadatok lehetnek intellektuálisan vagy automatikusan létrehozott másodlagos adatok, amelyek vagy magába az elsődleges dokumentumba ágyazhatók, vagy csatolókkal kapcsolhatók hozzá.

A számítástechnikai alapokon történő adatfeldolgozás, a hálózati információforrások elérhetőségének egyre sürgetőbb igénye azonban új módszereket is kifejlesztett, létrejöttek a különböző metaadat-rendszerek, melyeknek száma az 1990-es években szinte sohasem látott módon, robbanásszerűen megnövekedett. Ezek a rendszerek nem a MARC kiváltására születtek (sőt ezek közül számosat eredetileg nem is könyvtári alkalmazásra fejlesztettek ki), hanem arra, hogy a weben fellelhető, egymástól nagymértékben különböző elektronikus dokumentumokat valamilyen szinten le lehessen írni, és a visszakeresésüket könnyebbé lehessen tenni.

Napjainkra számos különböző metaadat-rendszer jött létre a hálózati információk feldolgozására, mégis ma még a szerteágazó igények miatt egységes, témafüggetlen, könyvtári területen jól használható metaadat-formátumról nem beszélhetünk. (A különböző változatok a leírás mélysége szerint csoportosíthatók; a metaadatok választéka, komplexitása messze elmarad a bibliográfiai formátumokban rögzített adatelemekétől.) Általánosságban is elmondható, hogy a létező metaadat szabványok értékészlete és érvényességi köre jelentős mértékben különbözik egymástól.

A nemzetközi szakirodalom a metaadatok három típusát különbözteti meg (hozzá kell azonban tenni ehhez a felosztási rendszerhez, hogy csak olyan metaadat-rendszerekre vonatkozik, amelyeket digitális könyvtárak objektumainak feldolgozására használnak):

- leíró metaadatok: az objektumban szereplő szellemi tartalom leíró, bibliográfiai adatait tartalmazza, azaz az információk leírására és azonosítására szolgálnak;
- adminisztratív metaadatok: az objektum kezeléséhez szükséges adatok rögzítésére, mint például szkennelés módja, a tárolási formátum, stb. (ezeket szokás technikai metaadatoknak is nevezni), továbbá idetartoznak azok a metaadat-készletek is, amelyek az objektum szerzői jogi, licenz és a hosszútávú megőrzéséhez kapcsolódó információk rögzítésére szolgálnak (ezek a prezervációs metaadatok); vagyis a legfontosab cél a tájékozódás és megjelenítés;
- strukturális metaadatok: az egymástól független objektumok közötti logikai kapcsolatok leírására, irányításra és feldolgozásra szolgálnak [10].

Általában, egy információs rendszerben kereső vagy böngésző végfelhasználók számára csak a leíró metaadatok láthatók.

A teljesség igénye nélkül következzen néhány ismert könyvtári és egyéb kulturális területen alkalmazott metaadat-rendszer:

- CIMI (Computer Interchange of Museum Information) (2003);
- DC (Dublin Core Metadata Initiative) (1995);
- DOI (Digital Object Identifier) (1998);
- DDI (Data Documentation Initiative) (1997);
- EAD (Encoded Archival Description) (1996);
- GEM (Gateway to Educational Materials) (1998);
- GILS (Government Information Locator Service) (1994);
- IAFA (internet Anonymous FTP Archive) /whois++ templates (1995);
- INDECS (Interoperability of Data in E-Commerce Systems) (2000);
- NCITS (National Committee on Information Technology Standards Technical Committee L8, Metadata) L8 javasolt tervezet ANSI szabvány metaadatokhoz (1998);
- ONIX (Guidelines for Online Information Exchange) (2000);

- TEI (Text Encoding Initiative) – fejléc (1990);
- URC (Uniform Resources Characteristic or Uniform Resources Citation);
- VRA (Visual Resources Association) (1997);
- WWW Semantic Header.

A digitális állományok tartalmának leírására szolgáló metaadatok használata elengedhetetlen a nagy gyűjtemények pontos és megfelelő megismeréséhez. A metaadatok segítenek a szabadszöveges keresés félreérthetőségének kiküszöbölésében és olyan szemantikai szempontok megadásában, amelyek szűkítik és konkretizálják az információkeresést.

### **3.5.3. A metaadatok meghatározása a könyvtár- és információtudományban**

A metaadat terminus technicus jelentése és használata nem egyértelmű és következetes azonban még a nemzetközi könyvtáros szakmában sem. Egyesek szakemberek használják mind a digitális, mind a nem digitális információforrások leírásakor is, mások azonban szűkített értelmezésben csak az elektronikus források leírásakor.

Ha közelebbről megvizsgáljuk a kérdést, azt találjuk, hogy a számítástudományban használatos metaadat fogalom a legszűkebb értelmezés. Valószínűleg azonban célravezetőbb úgy tekinteni a metaadatokat, mint az információforrások valamennyi típusának leírását, beleértve a nyomtatott (hagyományos) dokumentumokat is. Hiszen számos gyűjtemény tartalmaz együttesen mind elektronikus, mind pedig hagyományos dokumentumokat, és sok esetben egy-egy elektronikus dokumentum őse, alapja egy hagyományos formában megjelent mű.

Egy másik, gyakran említett megszorítás, hogy magának a metaadatnak kell elektronikusnak lennie függetlenül a leírt objektum természetétől, megjelenési formájától. A gyakorlatban ez a kikötés könnyen teljesíthető, hiszen napjainkban

a legtöbb információforrásról készített leírás már elektronikus formában készül és tárolódik.

A metaadat kifejezést talán a legszűkebben a W3C értelmezi: „a metaadat nem más, mint számítógéppel olvasható információ a web számára” [36]. Itt a géppel olvashatóság követelménye szinte teljes mértékben ellentmond annak az eredeti és határozott igénynek, amelyet a természet- és társadalomtudósok fogalmaztak meg numerikus adathalmazokkal kapcsolatban, miszerint olyan metaadatokra van szükségük, amelyek a számítógépes adataikat az ember számára érthetővé alakítják. A definíció utolsó része egy még kényesebb kikötést tartalmaz: „a web számára”. Ezzel azonban nemcsak a cédula- és a könyvkatalógusokat zárja ki ebből a körből, hanem azokat a forrás leírásokat is, amelyek más internet protokollon vagy nem web-alapú számítógépes rendszereken keresztül érhetők el.

A könyvtár- és információtudományban egy jóval liberálisabb meghatározását alkalmazzák a metaadatoknak. A metaadat strukturált adat tetszőleges dokumentumtípusú, vagy formátumú információs forrásról. Ebben az értelmezésben két megszorítás teljesítését várjuk el egy metaadattól. Az első, hogy strukturált legyen. Vagyis adatelemeknek egy véletlenszerűen összegyűjtött és megjelenített halmaza esetén nem beszélhetünk metaadatokról, csak abban az esetben, ha az adatelemeket egy előzetesen felállított metaadat sémával összhangban rögzítjük. A második pedig, a metaadatnak információforrást kell leírnia.

Az UKOLN definíciója összecseng az előzővel: „ a metaadat strukturált adat az elektronikus (és nem elektronikus) források adatairól, amelyek segítségével különféle tevékenységek széles skálája támogatható és végezhető el, mint például a források leírása és feltárása, az információs források menedzselése (beleértve a szerzői jogi kérdéseket) és hosszútávú megőrzése ” [35].

Átfogó, leginkább használatos, a konkrét alkalmazási területtől független definíció ezidáig nem készült.

A könyvtárak számára mindez vegyes haszonnal jár, lehetőségeket és változásokat, változtatásokat indukál. A pozitív oldalon azok az új lehetőségek

állnak, amelyek a könyvtári anyagok, források leírásában egy új, egy megújult szemléletmódot adnak az eddigi MARC hegemoniával szemben. Ugyanakkor ezek az új metaadat formátumok egyfajta új „teherként” és elvárásként is nehezednek a könyvtáros szakmára.

Az elektronikus dokumentumok katalógusba történő integrálását illetően jelenleg még két szemléleti irányzat uralkodik. Az egyik értelmezésben a bibliográfiai leírás nemzetközi szabványaival (ISBD) összhangban alakítható ki az elektronikus források ún. metaadatainak beépítése az online katalógusba. Itt a cél az, hogy adatcsere esetén kiderüljön, miről szól a dokumentum. Míg a másik, „minimalista” nézet szerint az internetforrások feltárása a hagyományos katalógustétel struktúrájával való szakítást igényel, a lehető legegyszerűbb formára törekszik, hogy az információkeresést minél könnyebbé tegye.

#### **3.5.4. A Dublin Core metaadatok**

Az említett metaadat-rendszerek közül a DC struktúrájáról lehet leginkább elmondani, hogy nem teljesen idegen a MARC-tól, megfeleltetésükről konverziós táblák segítségével lehet gondoskodni.

A DC az OCLC és a National Center for Supercomputing Application által finanszírozott műhelytanácskozásból nőtt ki 1995-ben [1, 2, 13]. A DC az elektronikus dokumentumok leírásának modelljeként 13 metaadat elemet határozott meg, amelyeket mint legfontosabbakat, magának (core) neveztek el, és a hálózaton lévő szöveges források leírására szánták. Ez a metaelemek dublini magja (DCME) néven vált közismertté. Az egy évvel később, 1996-ban Warwick-ban rendezett konferencia eredménye a Dublin Core szintaxisának, a használat irányelveinek, valamint az ún. Warwick Framework-nek a kidolgozása volt, amely a metaadat elemek elhelyezésére szánt modul rendszerű keret [58]. Ezt követően meghatározták a DC nem-szöveges, tehát főleg képi dokumentumokra való kiterjesztését, később pedig sor került a HTML-hez kapcsolódó szintaxis kidolgozására, amelyet Canberra Qualifiers-nek (módosító) neveznek. A DCME formátum továbbfejlesztéséről a DCMI gondoskodik, amely

1999-ben 15 azonosító felvételére, nevezetesen a DCMES alkalmazására tett javaslatot.

A DC-nak eredetileg a metaleírásokban a „lingua franca” szerepét szánták, később azonban széles körű érdeklődés középpontjába került a „felső szintű” egyszerűsített forrás (metaadat) leírások fejlesztésében. Úgy jellemezhetjük, mint az információforrások leírására szolgáló egyszerű, leíró jellegű, a keresés céljaira is releváns elemsorozatot, amely mögött bármely objektum leírását lehetővé tevő általános szemlélet áll. Ez az általánosság teszi lehetővé, hogy jól lehessen alkalmazni heterogén, sokszor amorf dokumentumok leírásánál és adataik kereshetőségénél [3].

A DC elterjedését elősegíti, hogy adatait az európai szabványosítási szervezet, a CEN is jóváhagyta.

### **3.5.5. A bibliográfiai rekordok és formátumok funkcionális követelményei (megváltozott elvárások)**

1990-ben Stockholmban szemináriumot rendeztek a bibliográfiai rekordok kérdésével kapcsolatban, ahol a különböző célú felhasználás szempontjából tárgyalták a bibliográfiai rekordok sajátosságait.

A szeminárium határozatainak egyike a bibliográfiai rekordok funkcionális követelményeinek kutatását tűzte ki célul. A projekt a bibliográfiai adatok megváltozott környezetéből indul ki. Az elkészült és 1998-ban megjelent FRBR-tanulmánynak kettős célja van: egyrészt definiálja azokat a funkciókat, amelyeket a bibliográfiai rekord betölt a különböző kiadványtípusokat, a különböző alkalmazásokat és a különböző felhasználói igényeket illetően; másrészt ajánlást tesz a rekordok funkcióinak alapszintjére és javasolt minimális elemkészletére [22]. Ezen túlmenően az FRBR-tanulmány felállítja a bibliográfiai adatok entitás-kapcsolat modelljét.

A tanulmány a legszélesebb értelemben tekinti át a bibliográfiai rekord funkcióit. Az elemzés gyakorlati célja a katalogizálás racionalizálása, elméleti célja pedig, a

bibliográfiai univerzumnak, céljainak és kapcsolat-rendszerének leírása, kiindulási alapot szolgáltatva a további formátum-fejlesztésekhez és kutatásokhoz.

Mindmellett új lendületet kapott az egységesített besorolási adatok kérdése is a FRANAR program elindításával [60]. Célja, hogy elősegítse a virtuális, osztott besorolási-adatállományok gyakorlati megvalósulását, valamint feladatai közé tartozik, hogy az FRBR-ben megkezdett munkát folytatva kidolgozza a besorolási rekordok funkcionális követelményeit, megvalósítási tanulmányt készítsen az ISADN-ről [32].

Természetesen további tendenciák is láthatók már, de ezek részben az előzőeken alapulnak, mint például a tetszőleges típusú hálózati információforrásokat feldolgozó RDF W3C technológián alapuló megoldások [46]. Ez egy olyan infrastruktúra, amely lehetővé teszi az internet-tartalom és a metaadatok egymáshoz rendelését, mégpedig úgy, hogy a metaadatok szabványos szolgáltatásának rugalmas infrastruktúráját hozza létre. Általánosságban jellemezve, az RDF olyan forrásleírási mechanizmus definiálását tűzte ki célul, amely nem tartalmaz egyedi felhasználási feltételezéseket, és nem definiálja egyetlen alkalmazási terület szemantikáját sem.

### **3.6. Metaadatok alkalmazása magyar könyvtári területen**

Magyarországi könyvtárakban metaadat-rendszereket alig-alig használnak, és különösen nem szabványos adatcsere-formátumban. Ha körülnézünk a magyar interneten, alig találunk néhány olyan könyvtári oldalt, ahol például a legegyszerűbb Dublin Core metaadatok benne lennének a dokumentumok fejlécében.

Az elmúlt években öröndetesen megszapordó hazai digitalizálási pályázatok és projektek is csak a digitális másolatok előállítását szorgalmazzák, és nem terjednek ki ezek szolgáltatásának, visszakereshetőségének módjára.

Hazánkban az ún. metaadatok vagy metaadat-rendszerek könyvtári alkalmazásával két jelentős, futó projektet érdemes megemlíteni: a Magyar Elektronikus Könyvtár (MEK), illetve az Elektronikus Periodika Archívum és Adatbázis (EPA) projekteket.

### **3.6.1. Magyar Elektronikus Könyvtár**

Az elmúlt néhány évben a MEK az egyik legfontosabb magyar oktatási és kulturális információforrássá vált az interneten (ittthon és külföldön egyaránt). Folyamatosan növekedik olvasótábor, a látogatók száma ma már havi 300-350 ezer főre tehető.

A MEK 1994-ben jelent meg az interneten. Az online információforrások iránt rohamosan növekvő igények és a gyűjtemény gyors bővülése miatt 2000-ben egy új „könyvtár” építésébe kezdtek: létrehozták a MEK 2.0 verzióját. A fejlesztés során mind a dokumentumok, mind pedig a kezelőfelület oldalai teljesen megújultak, korszerűbb formát kaptak.

A mi szempontunkból a legfontosabb újítások a műveket kísérő metaadatok terén történtek. A MEK új szolgáltatása használta a gyakorlatban Magyarországon először a metaadatok Dublin Core formátumát. A MEK fejlesztése során kidolgozták a MEK rekordstruktúrájának megfeleltetését mind a HUNMARC, illetve USMARC könyvtári szabványhoz, mind a Dublin Core nemzetközi szabványhoz. Az utóbbin alapulva kidolgozták ennek szerkezetleíró tábláját (DTD) a metaadatok XML-alapú szolgáltatásához.

A MEK-be feltöltött minden dokumentumról egyszerűsített bibliográfiai leírás készül, és ezek a bibliográfiai adatok az interneten elterjedt Dublin Core és a könyvtári világban használatos MARC szabványok szerint is rendelkezésre állnak, így az online forrásokat indexelő keresőgépek és az integrált könyvtári

rendszerek számára egyaránt átvehető az elektronikus dokumentumok metaadatai. A hozzáférést és a keresést további technikai megoldások is segítik: saját fejlesztésű, PHP-alapú katalógus; XML adatcsere, stabil URN azonosítók, Z39.50 és OAI protokoll szerinti lekérdezési lehetőség.

2004-ben a MEK 2.0 fejlesztés során szerzett tapasztalatokat és az ott alkalmazott technológiákat egy ingyenesen terjesztett, integrált e-könyvtári rendszerben tették közzé, eleMEK néven [31].

### **3.6.2. Elektronikus Periodika Archívum és Adatbázis**

Az EPA elektronikus időszaki kiadványokat kezelő archívum és forrás, illetve leőhely-nyilvántartó rendszer, mely egyrészt a MEK időszaki kiadványokat tároló különgyűjteménye, másrészt a MEK által működtetett országos internetes leőhely-adatbázis. Az EPA online és offline elektronikus időszaki kiadványok leírására, nyilvántartására szolgál.

Az EPA az archivált időszaki kiadványokat metaadatokkal látja el, azaz bibliográfiai adatokat HUNMARC formátumban és DC-XML formátumban szolgáltatja. A használt metaadat-struktúra és a DC-, valamint a MARC-megfeleltetések a rendelkezésre állnak [14].

### **3.7. Rövid áttekintés az XML-ről**

„Ahogyan a HTML lehetővé tette, hogy minden számítógép-felhasználó olvashassa az internetes dokumentumokat, úgy az XML – az inkompatibilis számítógépes rendszerek bábeli zűrzavarában is – az eszperantót jelenti, melyen mindenki olvashat és írhat.”<sup>1</sup>

---

<sup>1</sup> Bosak, J. - Bray, T.: Az XML és a második generációs web. [http://www.scala.hu/XML\\_DAT\\_hun.htm](http://www.scala.hu/XML_DAT_hun.htm)

Az XML szabvány és a hozzá kapcsolódó szabványos eljárások egy új lehetőséget kínálnak a szöveg és más médiák együttes kezeléséhez [15, 16, 55, 56]. Az XML a W3C által kifejlesztett új, hatékony dokumentumleírási és megjelenítési formátum. Az XML két meglévő nyelv, a HTML és az SGML elvei és szabályai alapján egyszerű, de mégis hatékony információátviteli, -feldolgozási és -megjelenítési rendszert hoz létre. Az XML elsődlegesen az SGML-re épül [25], de átvette a HTML néhány jellemzőjét is, valamint kiegészül olyan további lehetőségekkel, amelyek az interneten való felhasználását könnyítik meg.

Az XML-ajánlás 1998-ban jelent meg, amelyet az informatikus szakma nagy tetszéssel fogadott, és azóta széles körben elterjedt. Az XML sikerében használhatósága mellett további jelentős szerepet játszott az is, hogy az XML nyelvnek nincs tulajdonosa, és egyetlen gazdasági érdek sincs rá kizárólagos befolyással.

Az XML egy metanyelv, amelynek segítségével újabb jelölőnyelvek készíthetők. Az XML kiterjeszhetősége lehetőséget ad a különböző adatstruktúrák tartalmi és nem formai szempontok alapján történő leírására, amelyeket elektronikus úton tárolunk olyan formában, hogy egyaránt érthető legyen a számítógépes programok és az emberek számára is. Követelmény továbbá, hogy az XML dokumentumok emberi olvasásra is alkalmasak legyenek. Ebből következően az XML-fájlok nem binárisak. Egy XML-formátumú adatállomány az XML-dokumentum<sup>2</sup> szövegének és az XML-dokumentum egyes részeit összerendező és azonosító XML-jelöléseknek a keveréke.

Az XML ideális adatformátum a különböző médiákon terjesztett, közvetlenül megjelenített strukturált vagy félig strukturált szövegek számára. Az XML

---

<sup>2</sup> A könyvtár- és információtudományban a dokumentum fogalmán a könyvtár által állományba vett, alap- és kiegészítő feladatai ellátásához szükséges tudományos, oktatási, művészeti, közművelődési vagy történeti értékű könyv, időszaki kiadvány, egyéb kiadvány, illetőleg minden szöveg-, kép-, adat- és hangrögzítés, beleértve a könyvtár állományába vett elektronikus dokumentumot is, kivéve az levéltári törvény hatálya alá tartozó, irattári jellegű levéltári anyagnak minősülő dokumentumot. Egy XML dokumentum szövegből áll, Unicode karakterek sorozata, amely vegyesen tartalmaz jelölést és karakteres adatot.

A két dokumentum fogalom formai elkülönítése érdekében, amikor könyvtári dokumentumról beszélünk, akkor egyszerűen a dokumentum szót használjuk, amikor az XML felől közelítünk, ebben az esetben az XML-dokumentum kifejezést.

formában megjelenő információ „önleíró” – a megcélzott közönség és a használt média követelményeihez igazodóan használható, alakítható és formázható. A megjelenítésen túl, a különböző programok és számítógépes rendszerek közti információcsere egyre gazdagabbá és bonyolultabbá válik az új, hálózatalapú alkalmazások fejlődésével. Az ilyen információkkal szemben követelmény, hogy önleíróak legyenek annak érdekében, hogy a kliens program tudja értelmezni a kapott adatokat, és végrehajthassa a felhasználótól kapott feladatokat anélkül, hogy újra a kiszolgálóhoz kellene kapcsolódnia.

Egy XML-alapú dokumentumnak van logikai és fizikai szerkezete is. A logikai szerkezet lehetővé teszi, hogy a dokumentumot nevet viselő egységekre, alegységekre bontsuk, amelyeket elemnek neveznek. A fizikai felépítés során pedig lehetőség van arra, hogy a dokumentum részeit, ún. egyedeit elnevezzük és külön tároljuk általában olyan adatállományokban, amelyekből az információ újra kinyerhető; ezen túlmenően hivatkozásokkal és nem XML-típusú adatokkal is (például képekkel) bővíthetjük.

Az XML segítségével definiált nyelvek bővíthetők, ugyanis ezekben nincs előre definiált elemlista (mint például a HTML-ben), hanem saját elemeket lehet tetszőlegesen deklarálni. A deklarált elemek szabadon egymásba ágyazhatóak, s a legmagasabb szinten valamennyi elemet egyetlen dokumentumelembe kell zárni. A dokumentumelem egy konkrét előfordulását nevezzük XML-dokumentumnak. Az egy osztályba tartozó XML-dokumentumok elemei alkotják a definiált nyelv szókincsét, és az elemek tartalmát, valamint egymáshoz való kapcsolatát rögzítő szabályok adják a definiált nyelv nyelvtanát. Az XML specifikáció megad egy szintaxist mind az XML-dokumentumokra – vagyis az elemek jelölésére -, mind pedig a szabályok leírására (DTD). A megadott szintaktikai szabályok betartásával bárki saját nyelvet (XML-dokumentumtípust) készíthet, s azt a megfelelő XML-konform eszközzel ellenőrizheti, feldolgozhatja.

A DTD formális szabályokat definiál, amely egyebek közt lehetőséget nyújt egy adott dokumentum-osztályban (XML-dokumentumtípusban) felhasználható elemek előzetes megadására. A DTD meghatározza a megengedett elemek

egymáshoz fűződő viszonyát, meghatározza az elemek jellemzőit, az ún. attributumokat és az általuk felvehető értékeket stb., elkészítve ezzel az XML-dokumentumok logikai szerkezetének a sémáját. Ezzel gyakorlatilag kialakítja a dokumentumelem hierarchiáját és tagoltságát. Tagoltságnak nevezzük annak a mértékét, hogy az elemek tartalma mennyire felosztott gyermek elemekre. Ahol sok a gyermek elem vagy más néven leszármazott, finom tagoltságról, ahol pedig kevés a leszármazott, ott durva tagoltságról beszélünk.

Az XML-ben rejlő lehetőségek néhány jól megválasztott szabályból erednek. Ezek egyike az, hogy a jelölőelemek szinte mindig párosával szerepelnek, mintegy zárójelként veszik körül a szöveget, amelyre vonatkoznak. A jelölőelemek általában közvetlenül nem határozzák meg a megjelenítés stílusát, hanem csak azonosítják a különböző szövegrészek jellegét. A jelölőelem nyitó része és a jelölőelem záró része, a benne lévő adattal együtt jelent egy elemi egységet. Az XML-elemek tartalmazhatnak további, beágyazott elemeket tetszőleges mélységben, és végül az egész dokumentumot egyetlen dokumentumelembe kell zárni. Az elkészült XML-dokumentum szerkezetének felépítése sokféle grafikus módszerrel ábrázolható, az egyik lehetőség például a fa struktúra alkalmazása.

Az XML ilyen módon az egyes szakterületek specialistái számára lehetővé teszi új és egyedi nyelvek tervezését, azonban valóban jó nyelvet létrehozni korántsem lebecsülendő feladat. Egy-egy új XML alapú jelölőnyelv kialakítása előtt lényeges végiggondolni: milyen jelölőelemeket használjunk, a kódolt elemek hogyan ágyazódhatnak egymásba, és hogyan történjen a kódok végrehajtása. Az első kettőt (a nyelv „szókincsét” és „nyelvtanát”) általában a DTD szabályozza, a harmadikról az XML-hez kapcsolódó, kiegészítő szabványok és a megfelelő alkalmazás gondoskodik. Az XML szabvány nem kényszeríti egy új nyelv tervezőit a DTD kötelező használatára, azonban a legtöbb nyelv feltehetően mégis tartalmazni fogja, mert segíti a későbbiekben a programozókat helyesen működő szoftverek elkészítésében.

Az XML nem elszigetelt szabvány. Az alapvető szintaxis-specifikáción túl az XML-ben megtalálhatók a hipertextkapcsolatokat létrehozó (XLink és XPointer) és

dokumentumformázó (XSL és XSLT) szabványok is [16, 55, 56]. Az XPath lehetővé teszi az egyes komponensekre való hivatkozást egy XML-dokumentumon belül. Ez a képesség lehetővé teszi - többek között - az XSL és az XSLT stíluslapokon belül az XML-dokumentum egyes részeinek dinamikus kiválasztását olyan sorrendben, ahogy azt a kimenet megköveteli. Az XML névterek lehetőséget teremtenek arra, hogy egy XML-dokumentum több szótárból tartalmazzon XML-elemeket és attribútumokat névütközések nélkül. Az XML aláírás digitális aláírások létrehozására szolgáló, az XML titkosítás pedig XML-dokumentumok titkosítására szolgáló szintaxist és feldolgozási szabályt definiál.

Az XML-dokumentumok feldolgozásának problémáját ún. programozói interfész – amely az XML-dokumentumokat a leíró szabványoknak megfelelően értelmezi, és a benne tárolt adatokat a felhasználói rutinoknak átadja – kialakításával kívánják megoldani. A fejlesztések két irányban indultak el: a SAX szekvenciális, eseményvezérelt feldolgozást lehetővé tevő API, míg a DOM egy logikai adatmodell XML-dokumentumok leírására, amely az adatmodellen értelmezett műveleteket definiál, lehetőséget teremtve az XML-dokumentum manipulálására [5].

Ezen túlmenően létező kidolgozott és fejlesztés alatt álló szabványok sora egészíti és terjeszti ki az XML hatókörét (például ASCII, UNICODE, ISO/IEC 10464, HTTP, CSS, SPDL).

### **3.8. Az XML felhasználási területeiről**

Az XML megjelenése óta komoly sikereket ért el az elektronikus kereskedelemben, a tudományban, a számítástudományban és valamennyi iparágban, a gyártástól kezdve a gyógyításig. A legjelentősebb programgyártó cégek is behódoltak a szabvány előtt, és számos XML-t támogató termékkel jelentek meg.

Minden esetben, ha bonyolult adatformátumot kell két program között mozgatni, az XML megfelelő formátum lehet. Az XML már bizonyított bizonyos területeken, mint például az információszolgáltató technológiák (push technology), EDI és köztes adatok generálása MCF, XML-adatok és RDF.

Sokféle, ún. ipari DTD létezik, amelyeket sokan használnak, így egyfajta szabvánnyá válhat egy adott felhasználói körben. Az ipari DTD esetenként kifejezetten megengedő szabályokat tartalmaz annak érdekében, hogy sokféle projekthez alkalmazható legyen. Leggyakrabban ezért ezek az ipari DTD-k testre szabhatók, bővíthetők. Ilyen DTD például a TEI DTD, amelyet elsősorban szépirodalmi szövegek digitalizálásánál használnak.

Napjainkban már különféle adatok leírásához léteznek DTD-k, néhány ezek közül:

- XHTML, amely a HTML webes formátumleíró nyelv XML változata;
- DocBook, dokumentációk, könyvek leírásához;
- XFRML, amely pénzügyi jelentések leírására szolgál;
- XMLNews, a hírek tartalmának és meta-adatainak leírását teszi lehetővé;
- WML, a WAP mobil telefonok nyelve;
- MathML matematikai jelölések leírására alkalmas szerkezet és tartalom szempontjából egyaránt;
- SVG a skálázható vektorgrafikák jellegzetességeit és használati módját írja le.

Az XML alkalmazható relációs adatbázisok mozgatására is, ebben az esetben az XML elemek csupán határolóként használatosak a rendszerek közti rekordok, mezők és kapcsolatok átvitelekor.

Az XML szabványok fejlődésük során elérték a könyvtári világot is. Több projekt is elindult a MARC formátum XML-ben történő leírására (például XMLMARC). Ennek jövőbeli bevezetésével és elterjedésével lehetővé válhat az így feldolgozott rekordok kezelése, visszakeresése XML vagy XML-re épülő

eszközökkel. Teljes szövegű elektronikus dokumentumok esetén pedig közvetlen kapcsolat teremthető a bibliográfiai leírás és a teljes szöveg között.

Az XML-t alkalmazzák már a teauruszépítésben és –szerkesztésben is, mint például a Virtual HyperGlossary egy XML-formátumú teaurusz-szoftver.

Hazánkban, könyvtári területen a Szegedi Egyetem Egyetemi Könyvtárában Bakonyi Géza vezetésével indult egy program, amelyben digitalizált szöveget XML eszközökkel alakítanak át strukturált szöveggé. Az általuk kidolgozott módszer lényege, hogy egy digitális formában létező folyóirat esetén milyen módon lehet a strukturált szövegfeldolgozás eszközeit hasznosítani úgy, hogy az adott folyóirat szövegéből mintegy „automatikusan” előállítható legyen egy virtuális repertórium (lehetőleg a nyomtatott változatnál nagyobb lehetőségekkel) [5].

A Neumann Ház gondozásában futó Digitális Irodalmi Akadémia névre keresztelt projekt a TEI kódolást alkalmazza az irodalmi szövegek rögzítésére, amelyet a hazai szélesebb körű elterjedés végett magyarítottak.

Ne feledkezzünk meg a Magyar Elektronikus Könyvtár új rendszeréről sem. Míg az előző két példa a dokumentumok szintjén alkalmazza az XML/SGML technológiát, addig a MEK 2.0 a bibliográfiai leírások szintjén biztosítja az XML formátumban való letöltést. Az alkalmazott DTD-t az Empolis cég közreműködésével fejlesztették ki a TEI alapján, azt szűkítve.

## **4. EREDMÉNYEK**

### **4.1. Az XML alkalmazásának lehetősége a könyvtárak állományának formai feldolgozásában**

A (félig) strukturált adatok, strukturált szövegek (azaz esetünkben a bibliográfiai rekordok) meghatározásában és számítógépes környezetben történő kezelésében új lehetőségeket teremt az XML.

Az XML érdekes lehet mindazok számára, akik XML-dokumentumaik kialakításában és kezelésében nagyobb rugalmasságot kívánnak elérni, akik sokoldalúan támogatott rugalmas tárolási vagy általános adatformátumot, valamint fejlett, ám ezzel együtt nem túl drága alkalmazásokat keresnek.

Az XML nem egy programozási nyelv, és használatához vagy megtanulásához nem szükséges programozónak lenni (bár egy „kis” informatikai ismeret sosem árt). A legtöbb számítógépes adatformátumtól eltérően az XML-jelölések az ember (azaz könyvtáros szakemberek) számára is érthetőek, mivel semmi más nem tartalmaznak, mint közönséges szöveget.

Az XML szabvány elsődleges feladata ugyanaz, mint a MARC formátumoké, szintaktikai és szűkebb értelemben szemantikai szabályokat biztosít a bibliográfiai leírás (rekord) hierarchikusan rendeződő elemeinek formális leírásához. Az alapvető különbség a MARC formátumokhoz képest azonban, hogy az XML segítségével bibliográfiai adataink kialakításában és kezelésében nagyobb rugalmasság érhető el, ugyanazon (félig)strukturált adataink számára különféle, konkurens szerkezetek is kialakíthatók.

Az XML biztosítja az információ tervezését, autonómiáját, többszöri felhasználását, kereshetőségét, számítógépes feldolgozhatóságát, az információk összekötését, karbantartását és a minőséget, azaz a tudatos bánásmódot. A könyvtári dokumentumok feldolgozása során, illetve a különféle feldolgozási

problémákkal kapcsolatban oly sokszor emlegetett „egyszeri feldolgozás, többszöri felhasználás elve” összecseng ezekkel az elvárásokkal.

XML-dokumentumokat tetszőleges szövegszerkesztő programmal létrehozhatunk, de léteznek kifejezetten XML-érzékeny szerkesztők is. Ez utóbbiak képesek „olvasni” a DTD-t, ezzel segítve a szerkesztő munkáját, így elkerülhetők az XML-dokumentum létrehozása során a szintaktikai és a logikai hibák.

Az XML-jelölőnyelv használata lehetővé teszi a bibliográfiai leírásban rejlő lehetőségek jobb kihasználását a leírás azonosító funkciójának csorbítása nélkül. A katalogizálási szabvány a bibliográfiai leírás fogalmát a következőképpen határozza meg: a bibliográfiai leírás meghatározott szabályok szerinti egységes szerkezetben, formában és sorrendben leírt adatok összessége, amelyek a dokumentumok azonosítására, nyilvántartásba vételére és a róluk szóló tájékoztatásra szolgálnak. Ebből következően a bibliográfiai leírás formálisan értelmezhető, az adatelemek formálisan felismerhetők. Ha a könyvtáros szakember rápillant a leírás adott pontjára, máris tudja, hogy ott például egy kiadó megnevezéséről van szó, hiszen az adatelem környezetéből, illetve a megelőző egyezményes jelből ez egyértelműen kiderül. A számítógépek nem ilyen „okosak”: pontosan meg kell mondanunk a számukra, mi mit jelent, az hol található, és mi a teendő vele. Az XML olyan új nyelv, amelynek kifejezett célja az, hogy az adatok önmagukért beszéljenek.

A DTD-ben tudunk rendelkezni az elemek sorrendjéről, amely a bibliográfiai adatformátumoknál egy kritikus szempont. A DTD-ben megadható az elemekhez szöveggörnyezettől függően sorrendi szabályozás (az elemek XML-ben két logikai művelettel kapcsolhatók össze, a sorozat illetve a választás összekapcsolóval), és megadhatók előfordulási szabályok (az elemek kötelezőségére és ismételhetőségére vonatkozóan) az ún. mennyiségi jelzők alkalmazásával. Mivel XML-ben lehetőség van az elemek tetszőleges mélységű egymásba ágyazására, nagyfokú rugalmasság érhető el. Ennek köszönhetően a bibliográfiai leírás adatelemeinek, vagy néhány összetartozó adatelemének ismételhetősége, opcionálitása, feltételektől függő kötelezősége (például a

bibliográfiai leírás 1. adatcsoportjában a különböző párhuzamosan meglévő adatelemek kezelése, akár szerzői vagy vegyes gyűjtemények esetén is) XML-formátumban nem jelent problémát, rugalmasan kezelhetők a különböző előfordulás-variációk.

A bibliográfiai adatfeldolgozás során a bibliográfiai leírást a leírandó dokumentum nyelvén kell elkészíteni, eltekintve néhány, a katalógus vagy a bibliográfia nyelvén megadandó adatcsoportot és adatelemet. A dokumentum nyelvén kell rögzíteni az adatokat az azonosítás szempontjából legfontosabb adatcsoportokban (mint például a cím és szerzőségi közlés, kiadás, megjelenés, sorozat). A katalogizálás nyelve és a dokumentum nyelve az esetek nagy százalékában eltér egymástól.

Nehezebb a problémával állunk szemben akkor, ha a dokumentum nem latin betűs írásrendszerű betűkészlettel készült. A szabványos eljárás ebben az esetben a transliteráció, azaz a szabványos átírási táblázatok alkalmazása, amelyek segítségével az egyik írásrendszer minden egyes betűjének (jelének) a másik írásrendszernek pontosan egy betűje (jele) felel meg és fordítva. Ha az átírandó írásrendszer jelkészlete szükségessé teszi, az ISO-szabványok a latin ábécé betűkészletét kiegészítő jelek, azaz mellékjelek (vagy idegen szóval diakritikus jelek) alkalmazásával bővítik. Mivel a MARC formátum az ASCII kódrendszeren alapszik, a diakritikus jelek komoly gondot jelentenek mind a bibliográfiai rekordok megjelenítésénél, mind a bibliográfiai adatok indexelésénél és rendszerből történő visszakeresésénél. Éppen ezért a MARC alapú könyvtári rendszerrel dolgozó könyvtárak többsége kényszerűségből „eltekint” a diakritikus jelek használatától, egyszerűen csak a megfeleltetett alapkaraktert alkalmazza. Ezzel azonban a transliteráció elveszíti igazi célját és értelmét: a nemzetköziességet és az eltérő írásrendszerek betűi közötti kölcsönösen egyértelmű megfeleltetést.

Az XML elkerüli a programozási nyelvekben gyakran előforduló csapdákat: ugyanis az XML bővíthető, platform-független, támogatja a nemzetköziesítést. Az XML jelentőségének egyik záloga, hogy az új szabványon, a UNICODE nevű karakterkódolási rendszeren alapul, mely lehetővé teszi a világ főbb nyelveiben

készült szövegek kevert használatát. A UNICODE megengedi a diakritikus és egyéb speciális jelek, valamint a nem latin betűk alkalmazását, normál szövegként való kezelését. Az XML nyelvet értelmezni képes szoftver e karakterkészletek bármely kombinációját kezelni tudja, így az XML az adatcserét nemcsak a különböző számítógépes rendszerek, hanem az országok és kultúrák között is lehetővé teszi.

Az UNICODE használata a bibliográfiai rekordokban lehetővé teszi az idegen nyelvű bibliográfiai adatok bevitelét is az elektronikus adatbázisokba transliterálás nélkül. Ez azt jelenti, hogy a végfelhasználók a könyvtári katalógusokban bármilyen nyelven kereshetnek. Az Európai Unió is támogatja az UNICODE fejlesztését, és egy projekt keretében finanszírozza elterjesztését már a gyakorlatban használt automatizált könyvtári rendszerekben is azzal a céllal, hogy a világ bármely nyelvén írt bibliográfiai rekordokat vagy teljes szövegű dokumentumokat képesek legyenek tárolni.

Az XML-jelölések önmagukban nem tartalmaznak utasítást arra vonatkozóan, hogyan jelenjenek meg az adatok a képernyőn vagy nyomtatásban, azaz a jelölések a tartalom jellegére koncentrálnak függetlenül a megjelenítő közegtől. A kódolt tartalomhoz tetszőleges számú ún. stíluslap készíthető, s az ezekben a stíluslapokban szereplő szabályokat alkalmazva az anyag a különböző eszközöknek megfelelően automatikusan formázható újra és újra.

A bibliográfiai leírások formalizálásánál jól kihasználható ez az XML jellemző: a tartalom, a megjelenítés és a kapcsolások különválasztása, amelyek így külön-külön optimálisan kezelhetők. A felhasználást tekintve ez azt jelenti, hogy a bibliográfiai leírások különböző megjelenítési (output) formátumai külön-külön egy-egy stíluslap elkészítésével kényelmesen megoldhatók (figyelembe véve, hogy az XML-elemek, amelyek megfeleltethetők például a bibliográfiai leírás adatelemeinek, adacsoportjainak és így tovább, sorrendje XML-ben egyszerűen előírható).

A gyengén látó emberek számára készített szolgáltatásoknál is előnyös ez a megközelítés. A stíluslap illetve megfelelő szoftver segítségével átalakítható az XML-dokumentum Braille írássá vagy beszéddé.

Egy bibliográfiai leírás DTD megalkotásával és használatával számos előnyhöz juthatunk. A programozók készíthetnek adatkezelő és feldolgozó szűrőket anélkül, hogy a programokat nem várt adatok kezelésére fel kellene készíteni. Ugyanilyen nagy biztonsággal készíthetők stíluslapok is, és a DTD minden eleméhez stílusszabályok rendelhetők. XML-kompatibilis szerkesztő használatával a feldolgozó könyvtárosok irányíthatók, és egységes szabályoknak megfelelő XML-dokumentumok (bibliográfiai leírások) készíthetők (bár az XML megengedi a felhasználó által definiált újabb elemek létrehozását, a szabályozott környezetben használt szerkesztőprogramokban ez a lehetőség letiltható).

Mivel az XML nem előre definiált elemkészlettel dolgozik, lehetőség van arra, hogy az XML alapú bibliográfiai metaadat-rendszereket az egyes könyvtárak a helyi igényeinek és szükségleteiknek megfelelően testreszabják. Erre korábban a MARC formátumok esetében minimális lehetőség volt csak, a rekordfejben egy-egy előre nem definiált karakterpozíció, illetve a rekord bibliográfiai adatmezők részében előre nem definiált hívójel-tartományok jelentették a formátum rugalmasságát, flexibilitását.

Természetesen annak érdekében, hogy az ilyen módon testreszabott metaadat-rendszerek átjárhatók legyenek, szükséges, hogy egy közös DTD-nek vagy XML-sémának feleljenek meg.

Az XML számos olyan további előnyös tulajdonsággal rendelkezik, amellyel kivívhatja a könyvtáros szakma figyelmét is, elsősorban a formátum-fejlesztések területén. Olyan valós problémákra kínál megoldást, amelyekkel a napjainkban használt könyvtári számítógépes rendszerek adatformátumai jelenleg is küzdenek. Ezen túlmenően az XML könyvtári területen történő alkalmazása megoldást jelenthet az internet előretöréséből fakadó problémákra, valamint nyíltabb rendszer létrehozását teszi lehetővé a jelenlegi, MARC-alapú zártabb könyvtári rendszerekhez képest. (Nyílt rendszer az, amely olyan részekből áll, melyek egységes, gyártó független és nemzetközileg elfogadott szabványokon alapuló felületeket nyújtanak. Ennek köszönhetően az egyes részek rugalmasan összekapcsolhatók, valamint segítenek összehangolni a különböző hardver és szoftver termékek használatát.) Az internet globális világában egyre fontosabbá

váló követelmény, hogy a könyvtári OPAC-okhoz hasonló közszolgáltatások nyíltak, átjárhatók, más keresők által elérhetők, mélyen linkelhetők legyenek. Ma ennek a követelménynek egy átlagos könyvtári rendszer maradéktalanul nem képes megfelelni, amelynek egyik oka a nem megfelelő bibliográfiai adatformátum használatában rejlik.

Az XML-alapú metaadatokat használó katalógusok teljes mértékben integrálhatók a weboldallal, a weboldalakra, és nem szükséges a könyvtárnak egy izolált OPAC-ot elszigetelten működtetni. A kérdés már csak az, milyen is legyen ez az XML-alapú metadatrendszer?

## **4.2. A BDML (Bibliographic Description Markup Language) fejlesztési koncepciója, eszközei**

Fiander 2001-ben megjelent tanulmányában három különböző, lehetséges kiindulási pontot vázol fel bibliográfiai adatok XML DTD-jének kifejlesztéséhez:

1. MARC formátum a kiindulási alap, és ezt áttranszformálni XML-be,
2. a nemzeti katalogizálási elveket és szabályokat figyelembe venni, és ezeknek megfelelő struktúrát kialakítani az XML segítségével,
3. mielőtt az XML struktúrát elkészítenénk, éljünk a technológiai vívmányok adta új lehetőségekkel, és ötvözzük ezeket a katalogizálás területén jelenleg folyó kutatások legfrissebb eredményeivel [18].

A nemzetközi és hazai tendenciákat megvizsgálva arra a következtetésre juthatunk, hogy a napjainkra erőteljesen informatizálódott könyvtári világ a bibliográfiai információk kezelése, kezelhetősége kapcsán komoly érdeklődéssel fordul a metaadat-rendszerek és az XML technológia felé. Ezzel együtt azt is megállapíthatjuk, hogy a metaadatok könyvtári használata területén ma még a MARC formátum és ennek nemzeti variánsai az egyeduralgok, illetve a digitális (elektronikus) könyvtárak a DC vagy TEI helyi adaptációval dolgoznak.

MARC formátum robosztus adatszerkezete, a feldolgozást segítő kódolt információk azonban túlterhelik a bibliográfiai rekordokat, a MARC webes környezetben nehézkesen alkalmazható, és mint ahogy azt már korábban láttuk (a 3.5.1. alfejezetben), a MARC XML alapú továbbfejlesztései sem szakítanak a korábbi szerkezettel.

Az internet és az elektronikus dokumentumok előretörésének köszönhetően azonban a könyvtárak figyelme a metaadat-rendszerek felé fordult, sajnos azonban kifejezetten a könyvtári igényekhez, illetve bibliográfiai leírási elvárásokhoz alkalmazkodó metaadat-rendszerek jelenleg még nincsenek. A létező metaadat-rendszerek döntő többsége kizárólag elektronikus dokumentumok feltárására alkalmazható (kialakításuk is ennek megfelelően történt), és az ezek által támogatott bibliográfiai leírási szint és minőség nem éri el az ISBD/ER (Electronic Resources) [28] nemzetközi és a KSZ/2 Bibliográfiai leírás. Elektronikus dokumentumok [30] hazai szabályozásokban megfogalmazott elvárásokat.

Másrészt, ha az integrált könyvtár eszményéből indulunk ki, akkor olyan metaadat-rendszerekre van/lesz szükségünk, amelyek mind a hagyományos mind az elektronikus dokumentumok bibliográfiai feldolgozását támogatják. Érdemes lenne tehát a hagyományos dokumentumok irányába is kiterjeszteni a metaadat-rendszerek hatósugarát, amelyek bibliográfiai leírása - sok szempontból - jóval bonyolultabb az információhordozóik sokszínűsége és az azonosító funkció teljesítése miatt, éppen ezért innen lenne szerencsés elindulni.

Az XML egy metanyelv, amelynek segítségével újabb jelölőnyelvek készíthetők. Ebből következően a fejlesztés olyan formális szabálygyűjtemény kialakítására irányult, amely XML nyelven megadja a szabványos bibliográfiai leírás szerkezetének sémáját. A készülő nyelv a BDML nevet kapta.

A BDML nyelv egy kísérlet arra vonatkozóan, lehetséges-e az XML segítségével olyan új jelölőnyelvet kifejleszteni a bibliográfiai leírások összeállítására, amely lehetővé teszi a bibliográfiai leírásban rejlő lehetőségek jobb kihasználását a leírás azonosító funkciójának csorbítása nélkül.

A ma létező metaadat-rendszerek többségében az elemek mindegyike szabadon választható és ismételhető, ezek sorrendje nem meghatározott, azonban az általános érthetőség miatt mégis kötött szókészlet alkalmazása ajánlott vagy szükséges. Az egyes elemeket különböző minősítőkkal lehet rendszerint finomítani, jelentésüket szűkíteni.

A BDML elsősorban a hatályos magyar katalogizálási szabványokra, szabályzatokra, illetve előírásokra támaszkodik. Illeszkedve a bibliográfiai leírás definíciójában szereplő „meghatározott szabályok szerinti egységes szerkezetben, formában és sorrendben leírt adatok összessége” elváráshoz, valamint a nemzetközi érthetőség érdekében, az ISBD szabályzatokban elfogadott adatcsoportokat, adatelemeket megjelenítő, a bibliográfiai adatok előfordulás-variációinak megfelelő, kötött sorrendű elemkészletet deklará a BDML nyelv.

A bibliográfiai leírás fogalmából következik, hogy magának a leírásnak formálisan értelmezhetőnek, az adatelemeknek formálisan felismerhetőeknek kell lenniük. Ha a könyvtáros szakember rápillant a leírás adott pontjára, máris tudja, hogy ott például egy kiadó megnevezéséről van szó, hiszen az adatelem környezetéből, illetve a megelőző egyezményes jelből ez egyértelműen kiderül. A számítógépek nem ilyen „okosak”: pontosan meg kell mondanunk a számukra, mi mit jelent, az hol található, és mi a teendő vele. Az XML olyan új nyelv, amelynek kifejezett célja, hogy az adatok önmagukért beszéljenek.

A BDML rendelkezik az elemek sorrendjéről, amely a bibliográfiai adatformátumoknál a kritikus szempontok egyike. Az elemekhez szöveggörnyezettől függően sorrendi szabályozást ad (az elemek XML-ben két logikai művelettel kapcsolhatók össze: a sorozat, illetve a választás összekapcsolóval), és megad előfordulási szabályokat (az elemek kötelezőségére és ismételhetőségére vonatkozóan) az ún. mennyiségi jelzők alkalmazásával.

Mivel XML-ben lehetőség van az elemek tetszőleges mélységű egymásba ágyazására, nagyfokú rugalmasság érhető el. Ennek köszönhetően a bibliográfiai leírás adatelemeinek, vagy néhány összetartozó adatelemének ismételhetősége, opcionálitása, feltételektől függő kötelezősége (például a bibliográfiai leírás 1. adatcsoportjában a különböző, párhuzamosan meglévő adatelemek kezelése,

akár szerzői vagy vegyes gyűjtemények esetén is) BDML-formátumban nem jelent problémát, rugalmasan kezelhetők a különböző előfordulás-variációk.

Összefoglalva, a BDML-től várt legfontosabb jellemvonások a következők:

- rugalmas, tetszőlegesen bővíthető, platform-független metanyelv, illetve adatformátum (az adatstruktúra tartalmi és nem formai szempontok alapján történő leírása, amelyet elektronikus úton tárolunk olyan formában, hogy egyaránt érthető legyen a számítógépes programok és az emberek számára is);
- metaadatok különválasztása a megjelenítési formátumoktól;
- „önleíró” adatelemek (a megcélzott közönség és a használt média követelményeihez igazodóan használható, alakítható és formázható);
- rugalmasan illeszkedik a feldolgozandó dokumentumhoz (az XML segítségével definiált nyelvek szükség szerint bővíthetők, ugyanis ezekben nincs előre definiált elemlista, saját elemeket tetszőlegesen lehet deklarálni);
- különböző nyelveken írt szövegek kevert használata (az XML által támogatott UNICODE karakterkódolási-rendszer megengedi a diakritikus és egyéb speciális jelek, valamint a nem latin betűk alkalmazását, normál szöveggént való kezelését; XML nyelvet értelmezni képes szoftver e karakterkészletek bármely kombinációját kezelni tudja, így az XML az adatcserét nemcsak a különböző számítógépes rendszerek, hanem az országok és kultúrák között is lehetővé teszi);
- költséghatékony eszköz (az XML nyelv bárki által szabadon hozzáférhető, nyílt szabvány, nincs tulajdonosa, és egyetlen gazdasági érdek sincs rá kizárólagos befolyással).

A BDML fejlesztéséhez alapul az MSZ 3424/1 szabvány szolgált, amely a könyvek bibliográfiai leírását szabályozza [40]. A választást a következők indokolják:

- Az MSZ 3424/1-es szabvány az IFLA ISBD/M ajánlása alapján készült, azaz illeszkedik a nemzetközi szakma ajánlásaihoz és elvárásaihoz a könyvek bibliográfiai információinak feldolgozását tekintve.
- Az ISBD/M, és ennek következtében az MSZ 3424/1 valamennyi dokumentumtípus leírásának az alapját jelentik: időben a legkorábban jelentek meg a leírási szabványok sorában; olyan dokumentumtípusról szólnak, amely jól behatárolható jellemvonásokkal bír, ezen jellemvonások többsége jellemző más dokumentumtípusokra is (például címekkel, szerzőségi adatokkal a legtöbb dokumentum rendelkezik); valamint ha a további ISBD, illetve magyar szabványokat és szabályzatokat tekintjük, azok is kiindulási és viszonyítási alapként ezeket tekintik, ezekre hivatkoznak (például a főcím kiválasztásának részletes szabályait a könyvek leírási szabványa taglalja, minden más dokumentumtípus leírási szabványa ezt alapnak, ismertnek tekinti, és csak hivatkozik rá; természetesen a főcím és más címadatok forráshelyeinek kérdése és kezelése dokumentumtípusonként eltérő).
- A BDML továbbfejlesztése, kiterjesztése a további dokumentumtípusok feldolgozására abban áll, hogy az egyes dokumentumtípusok elérő formai és egyéb sajátosságaiból fakadó, és azokat megragadó/leíró, újabb XML-elemekkel bővítjük a BDML-t, illetve már létező XML-elemekhez újabb, alternatív elemtartalom-struktúrákat rendelünk.

A BDML DTD és a BDML-alapú bibliográfiai rekord példák készítésében, ezek jól formázottságának és érvényességének ellenőrzésében a Stylus Studio 6 XML Home Edition környezetet alkalmaztam.

### **4.3. A BDML tervezése, tervezési elvei**

A BDML tervezésénél fontos szereppel bírtak az elmúlt években a metaadatokkal és az egyes elemekkel, a qualifierekkel kapcsolatos megjelent tanulmányok.

A dokumentumok, információforrások sajátosságainak megértéséhez érdemes visszanyúlni és megvizsgálni a metaadatok elméletét és a bibliográfiai leírás alapjait is.

A problémák első csoportja az információforrásokhoz kapcsolódik. A metaadat adat az adatról, azonban az adatok nem egy légüres térben lebegnek, hanem minden esetben valamilyen hordozóhoz kapcsolódnak. Ezért is szerencsésebb, ha általában információforrásokról beszélünk, és nem adatokról. Mit is tekinthetünk információforrásnak, milyen adatokat kell leírni? A szakirodalom jelenleg eléggé egyértelmű választ ad erre: mindent fel kell dolgozni, amire igény van. Azaz marketing központú a válasz: amire az információs piacon igény van, azt fel is kell dolgozni.

A következő terület az adatok csoportja, tehát amelyekről a tényleges metaadat leírás majd elkészül. Ebből a szempontból az adat nem más, mint az információforrás egy absztrahálható sajátossága. Ilyen sajátosságok a következők (csak a szempontomból legfontosabbakat említem): megnevezés, létrehozó, keletkezés körülményei, időbeli és földrajzi kiterjedés, megtalálhatóság. A könyv dokumentumok szempontjából ezek többnyire jól kiforrottak, bár léteznek speciális, nehezen kezelhető esetek, valamint a kiadói hanyagság és pontatlan adatközlés is megnehezítheti a dolgunkat.

A megtalálhatóság két sajátosságot is magában rejt: az elérhetőséget és a hozzáférhetőséget. Ez mind a hagyományos, mind az elektronikus dokumentumok esetében fontos: a használó arra kíváncsi, hogy az adott dokumentumhoz hol és hogyan, milyen körülmények között férhet hozzá.

Maga a bibliográfiai leírás ezeknek a sajátosságoknak a számbavétele, s általánosságban három csoportba osztjuk őket: szerzőségi/tulajdoni adatok; szellemi tartalomra vonatkozó adatok; formai/fizikai megjelenés leírására alkalmas adatok összessége.

Az egyes mezők, az egyes elemek szerkezetében fontos szerepet játszanak a kvalifikerek, azaz a metaadat elemeket kiegészítő, illetve azok értéket pontosító attribútumok, amelyek lehetővé teszik, hogy az adatelem rekord készítője a használt elemek szemantikáját finomítsa, illetve azok értékeit ellenőrzöten és pontosan megadhassa.

Természetesen az adatelemek komponensei is sokszor elegendően (azaz például a cím esetében ilyen komponensek a főcím, alcím, alternatív cím, stb.), de bizonyos esetekben ennél pontosabb attribútumokat is társítani kell az egyes elemekhez. Ezek az attribútumok kétfélék: az egyik az adatelemek szemantikájának pontosítását (element qualifiers), a másik az adatelem értékének a pontosítását (value qualifiers) szolgálja. Az első eset azt jelenti, hogy bizonyos elemeknél szükséges az adatelem szemantikájának további finomítása az adatelem értéke és az információforrás közti kapcsolat pontosítása érdekében.

Az adatelem értékének pontosítása esetében ezeknek a qualifiereknek a segítségével tölthetjük ki a nem szabadszöveges mezőket; listákban vagy adat- és fogalomtárakban (tezauruszok, X500, stb.) megadott értékek kiválasztásával, illetve szabványokban, ajánlásokban rögzített sémák (ISO, URI, ISBN, RFC, stb.) alapján. Ez nagyon hasonlít a könyvtári adatbázisokban használatos *authority control*a, de annál általánosabb és messzemenően figyelembe veszi a hálózati információforrásokhoz rendelt ajánlásokat.

Azokban a kérdésekben, amelyeket dolgozatomban a BDML kialakításának tárgyalása során nem említek, az MSZ 3424/1, az ISBD/M és a további kapcsolódó szabványok és könyvtári szabályzatok útmutatásai az irányadók, ezeket figyelembe véve alakult ki a DTD.

A DTD-ben szereplő elemek funkcionalitás szerint 3 csoportba oszthatók:

- a szabványos bibliográfiai leírás adatelemeinek megfeleltethető,
- a bibliográfiai leírás szerkezetét leíró,
- a bibliográfiai leírás egyezményes jeleinek megfelelő

DTD elemek.

A BDML formátum a megválasztott adatelemek köre szerint a teljes bibliográfiai leírás készítésének támogatására készült. (Természetesen szükség esetén a BDML DTD alkalmas lehet egyszerűsített leírások készítésére is, ebben az

esetben a DTD-ben az érintett részfák lecserélhetők, elhagyhatók, vagy a szükséges elemeknél a kötelezőséget jelölő gyakoriság jelző lecserélhető az opcionálisra az MSZ 3424/1 útmutatása alapján.)

A bibliográfiai leírás szintje szerint a BDML támogatja a monografikus, az összefoglaló és az analitikus szintű leírások készítését. A leírás szerkezete szerint monografikus és az analitikus szintű leírások esetén az egylépcsős, összefoglaló szintű leírások esetén többlépcsős leírás készítését teszi lehetővé.

A szabványos bibliográfiai leírásban az adatcsoportok és az adatelemek sorrendje a leírás tárgyától (azaz bibliográfiai szintjétől) és szerkezetétől függ. A leírás adatelemeinek meghatározott körülményektől függő kötött sorrendje, valamint az egyes adatelemeket megelőző nemzetközi szinten egységes és elfogadott egyezményes jelek biztosítják a bibliográfiai leírások nemzetközi szintű érthetőségét és értelmezhetőségét (abban az esetben is, ha a könyvtáros nem beszéli azt a nyelvet, amely nyelven a könyv dokumentum információ-tartalma megjelent).

A BDML rendelkezik az elemek sorrendjéről, amely a bibliográfiai adatformátumoknál egy kritikus szempont. Az elemekhez szövegkörnyezettől függően sorrendi szabályozást ad (az elemek XML-ben két logikai művelettel kapcsolhatók össze, a sorozat illetve a választás összekapcsolóval), és megad előfordulási szabályokat (az elemek kötelezőségére és ismételhetőségére vonatkozóan) az ún. mennyiségi jelzők alkalmazásával. Mivel XML-ben lehetőség van az elemek tetszőleges mélységű egymásba ágyazására, nagyfokú rugalmasság érhető el. Ennek köszönhetően a bibliográfiai leírás adatelemeinek, vagy néhány összetartozó adatelemének ismételhetősége, opcionálisága, feltételektől függő kötelezősége (például a bibliográfiai leírás 1. adatcsoportjában a különböző párhuzamosan meglévő adatelemek kezelése, akár – könyvek esetében, szerzői vagy vegyes gyűjtemények esetén is) BDML-formátumban nem jelent problémát, rugalmasan kezelhetők a különböző előfordulás-variációk.

### 4.3.2. Az egyezményes jelek kezelése

A bibliográfiai leírás egyezményes jeleit jelölő elemek üres elemekként lettek definiálva, azaz ezek olyan elemek, amelyeknek nincs tartalmuk. Az ilyen tartalom nélküli elemek esetén létezik a jelölésnek egy másik formája is. Ekkor a zárórész jelölőelem elmarad, és a nyitó (illetve üres) jelölőelemnek kicsit eltérő a szintaxisa, '>' helyett '/>' zárja.

A leírás kötelezően használandó egyezményes jeleinek sajátos jelentésük és céljuk van:

- az adatcsoportok elkülönítése;
- az adatelemek típusának felismerhetővé tétele;
- kiegészítő információk közlése az adatelemekről, illetve az adatról, és az adatok tagolása az adatelemen belül;
- az adatforrások jelölése.

Az egyezményes jelek első funkciójának teljesítését segíti az XML azon jellemvonása, mely szerint az elemek tetszőleges mélységben egymásba ágyazhatók, így az egyes adatcsoportok tartománya jól elkülöníthető. Az egyes bibliográfiai adatelemek felismerhetőségét segítik kisebb részben a nekik megfeleltethető BDML elemek megnevezései, nagyobb részben ezen elemek caption attribútuma, amely tartalmazza a célzott bibliográfiai adatelem szabványos magyar nyelvű megnevezését. Ez a későbbiekben a beviteli képernyők tervezésénél, illetve az output formátumok készítésénél is segítségünkre lehet.

Az egyes bibliográfiai adatelemek kötelezőségének feltételeiről és megismételhetőségéről az MSZ 3424/1-es szabvány részletes előírásokat tartalmaz, a BDML kialakításánál ezeket vettem figyelembe.

A bibliográfiai leírás egyezményes jelei üres, önálló XML-elemekként definiáltak. Tehát az egyezményes jelek nem az elemtartalom részei (mint például a USMARC formátum adatmezőiben), hiszen ez ellentmondana az egyezményes jelek definíciójának, amely szerint az egyezményes jel mintegy bevezeti és

megelőzi a bibliográfiai adatelemet, így minősítve azt. Ebből következően nem az XML-elem egy konkrét előfordulásához kapcsolódik, hanem magához az XML-elemhez, vagyis magához a bibliográfiai adatelemhez.

Az egyezményes jelekhez kapcsolódóan egy másik megfontolandó probléma: vajon a BDML DTD-nek kell-e tartalmaznia ezeket, vagy ráérünk később foglalkozni velük a megjelenítés során.

Úgy vélem, mindkét út járható, és találkozunk mind a két megoldási elvvel és módszerrel a korábban készült, napjainkban használatos bibliográfiai adatformátumoknál. Az előbb említett USMARC az adattartalom részének tekinti az egyezményes jeleket. Ez a megoldás - bár teljesen ellentmond a bibliográfiai leírás elméletének és lényegének - a különféle megjelenítési (output) formátumok készítését megkönnyíti. A HUNMARC formátum ellenben nem foglalkozik az egyezményes jelekkel, szabályzata szerint a megjelenítéshez a megfelelő egyezményes jeleket az adatmezőkben alkalmazott hívójelek, indikátorok és almezőazonosítók alapján kell generálni. Ennek következtében itt az egyes bibliográfiai leírások (rekordok) korrekt megjelenítése meglehetősen bonyolult és nehézkes feladat.

A fentiekből kiindulva a BDML DTD tervezésénél a következő döntésre jutottam: mivel az egyezményes jelek az adatelemek nemzetközi értelmezhetőségét hivatottak elsősorban szolgálni, és a bibliográfiai leírás szerkezetét határozzák meg, annak szerves részét képezik, ezért az egyezményes jeleknek a DTD-ben kell megjeleníteniük. Ha mint önálló elemeket definiáljuk az egyezményes jeleket, akkor hozzárendelhetők a bibliográfiai adatelemek különféle előfordulás-variációihoz. Így gyakorlatilag megkönnyíthetjük a későbbi megjelenítési formátumok készítését, és ami talán fontosabb, garantálhatjuk a szabványos bibliográfiai leírások készítését is.

Az egyezményes jelek többsége egykarakteres, de vannak közöttük olyanok is, amelyek két vagy több karakterből állnak. Az egykarakteres egyezményes jeleket, illetve azokat a többkarakteres egyezményes jeleket, amelyeknél ezek a karakterek szorosan egymást követik, egymás mellett állnak, és így hordoznak egy jól meghatározott jelentést, definiálhatjuk egy-egy XML-elemmel. Ilyen például a „,” (vessző), amelyet a <colon> elem definiál, vagy “. -” (pont szóköz

gondolatjel, azaz adatcsoport elválasztó jel), amelyet a <pointspacedashspace> elem.

A zárójelek viszont két karakterből állnak, s bár egyetlen jelpárnak minősülnek, egyetlen jelentés kapcsolódik hozzájuk, a jelpárt alkotó karakterek távol kerülnek egymástól, s lényeges az, hogy milyen adateleme(ke)t fognak közre. Ezért a zárójelek esetében a zárójelet alkotó nyitó és záró karaktereknek külön, egy-egy XML-elemet feleltetünk meg. A kerek zárójelpár két tagjához rendelt XML-elemek: az <o\_squarebracket> a nyitó elem (az o rövidítés az open angol szóra utal), <c\_squarebracket> a záró elem (a c rövidítés a close angol szóra utal). Hasonlóan járunk el a szögletes zárójel esetén: az <o\_curvedbracket> és a <c\_curvedbracket> XML-elemek definiálják.

Azokhoz az egyezményes jelekhez, amelyek a konkrét adatokról (elemtartalmakról) közölnek kiegészítő információkat, illetve amelyek az adatokat az adatelemen belül tagolják, nem rendelünk XML-elemeket, hiszen ezek az egyezményes jelek a leírásban megjelenő konkrét bibliográfiai adatokhoz, azaz egy konkrét feldolgozott könyvhöz kapcsolódnak. Ezek helyes használatát a BDML DTD-hez illeszkedő bibliográfiai rekordokat készítő feldolgozó könyvtárosra bízuk, illetve tőle várjuk el. Ide tartoznak a következők (részben az MSZ 3424/1 2. táblázat alapján):

- felkiáltójel (a könyv valamely címét záró vagy tagoló felkiáltójel);
- három pont (a szövegelhagyás jelölése a Cím és szerzőségi közlés valamint a Sorozat adatcsoportban);
- idézőjel (annak tükrözése a leírásban, hogy valamely címet vagy szerzőségi adatot vagy ezeknek valamely részét a könyv idézőjelben közölte);
- kérdőjel (a könyv valamely címét záró vagy tagoló kérdőjel);
- pont (abban az esetben, ha a pont szórövidítések vagy sorszámok után áll);
- szögletes zárójelbe tett felkiáltójel (sajtóhiba jelölése a megfelelő szó, kifejezés vagy egyéb jelcsoport után);
- vessző (abban az esetben, ha cím szövegét tagolja szükség szerint, illetve ugyanazon szerzőségi adatban felsorolt szerzők nevének elkülönítésekor);

- [et al.], [etc.] (adatelhagyás jelölése abban az esetben, amikor az azonos szerzőségi vagy közreműködői funkciót betöltő személyek vagy testületek száma a hármat meghaladja).

Az egyezményes jelek megadása a DTD-ben tehát a következő:

```
<!-- Definition of Punctuation Symbols -->
<!ELEMENT c_curvedbracket EMPTY>
<!ELEMENT char_position3. EMPTY>
<!ELEMENT colon EMPTY>
<!ELEMENT comma EMPTY>
<!ELEMENT c_squarebracket EMPTY>
<!ELEMENT diagonal_slash EMPTY>
<!ELEMENT newindent EMPTY>
<!ELEMENT o_curvedbracket EMPTY>
<!ELEMENT o_squarebracket EMPTY>
<!ELEMENT plus EMPTY>
<!ELEMENT point EMPTY>
<!ELEMENT pointspacedashspace EMPTY>
<!ELEMENT equals_sign EMPTY>
<!ELEMENT semicolon EMPTY>
```

#### 4.3.3. A leírás nyelve és írásrendszere

Az XML által támogatott UNICODE karakterkódolási rendszernek köszönhetően nem jelent gondot, ha idegennyelvű dokumentum feldolgozása esetén a dokumentum szellemi tartalmát közvetítő nyelvben az angol ábécé betűitől különböző nemzeti karakterek szerepelnek, és természetesen ezek a bibliográfiai adatokban is meg kell, hogy jelenjenek. Sőt, abban az esetben sincs nagyobb probléma, ha a dokumentum nem latin betűs íráskészletet használó nyelven íródott, ugyanis a UNICODE rendszer használatából következően kezelhetők a transliteráció során (a nem latin betűs írásrendszer minden egyes betűjének szabványos átírási táblázatok alapján pontosan egy diakritikus, vagy más néven mellékjelekkel módosított latin betűket feleltetünk meg) megjelenő mellékjeles latin betűk.

#### **4.3.4. A bibliográfiai adatelemeknek megfeleltetett XML-elemek és attribútumaik**

Az egyes bibliográfiai adatelemeknek megfeleltetett BDML DTD elemek megnevezése/elnevezése az ISBD/M alapján, az angol nyelvű (azaz nemzetközi) terminológiából kiindulva történt a nemzetközi felhasználhatóság érdekében.

A bibliográfiai adatelemeknek megfeleltethető XML-elemek többsége rendelkezik a `retrieve` és `caption` attribútumokkal. Az első arról ad információt, hogy az XML-elem értéke indexbe bekerüljön-e, azaz az információ-visszakeresés folyamatában kereshetővé akarjuk-e tenni. A második attribútum az XML-elemhez kapcsolódó bibliográfiai adatelem magyar nyelvű megnevezését tartalmazza, amely segítségünkre lehet majd az adatbeviteli képernyők, illetve az output (megjelenítési) formátumok tervezésénél és kialakításánál.

A párhuzamos bibliográfiai adatelemeknek megfeleltetett XML-elemek megnevezése rendre p kezdőbetűvel indul az angol parallel szó rövidítéséből adódóan.

#### **4.3.5. A bibliográfiai adatelemek forráshelyeinek kezelése (főforrások és megengedett források)**

A bibliográfiai leírás készítésénél és értelmezésénél fontos, jelentéssel bíró szerepe van annak, hogy az egyes bibliográfiai adatelemek értékei/előfordulásai az leírt/feldolgozott dokumentum mely részéből, azaz a dokumentum mely forráshelyéről származik. A szabvány minden egyes bibliográfiai adatelem vonatkozásában meghatározza az elsődleges forráshelyeket (főforrásokat), ahonnan az adatot elsősorban kell venni a bibliográfiai leírás számára, illetve a másodlagos forráshelyeket (megengedett források), amelyek akkor vehetők figyelembe, ha a főforrásban nem találjuk meg a szükséges adatot. A hagyományos bibliográfiai leírásban a megengedett forrásból származó adatokat szögletes zárójelpárba, mint egyezményes jelbe foglaljuk. Ugyanezt tesszük akkor is, ha MARC formátumban dolgozunk, a megadott bibliográfiai almezőben

szögletes zárójelbe tesszük az adatot. Ennek a technikának az alkalmazásával a MARC almezőkben keverednek a bibliográfiai adatelemek előfordulásai az egyezményes jelekkel, holott az egyezményes jel nem az előforduláshoz kapcsolódik.

A BDML tervezésekor azoknál a bibliográfiai adatelemeknél, amelyeknél megkülönböztetünk fő- és megengedett forrásokat, ezeket XML-elem szinten különítjük el, és ezektől teljesen függetlenül kezeljük az egyezményes jeleket.

Például a következő DTD-részlet a bibliográfiai leírás 1. adatcsoportjába tartozó egyéb címadat definiálását mutatja, amely fő- és megengedett forrásból is származhat:

```
<!ELEMENT othertitle (#PCDATA|othertitle.osource)*>
<!ATTLIST othertitle retrieve (y|n) "n"
           caption CDATA #FIXED "egyéb címadat">

<!ELEMENT othertitle.osource (o_squarebracket,othertitle.os,c_squarebracket)>

<!ELEMENT othertitle.os (#PCDATA)>
<!ATTLIST othertitle.os retrieve (y|n) "n"
           caption CDATA #FIXED "egyéb címadat megengedett
forrásból">
```

Az <othertitle> jelölések közé kerül az egyéb címadat konkrét értéke, ha főforrásból származik, és <othertitle.os> jelölések közé, ha megengedett forrásból, amelyet közrefognak a szögletes zárójelpárnak megfelelő XML-elemek.

A névadási konvenció tehát a következő: a főforrású bibliográfiai adatnak megfelelő XML-elem nevét kiegészítjük a .os utótaggal, amely az angol „other sources”, azaz megengedett források kifejezés rövidítéséből ered.

### 4.3.6. A BDML struktúrájának áttekintése

A következő táblázat tartalmazza az egykötetes könyvek (monografikus szintű) leírásának MSZ szabvány szerinti adatelemeit és az azoknak megfelelő BDML elemeket:

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- XML BIBLIOGRAPHIC DESCRIPTION DTD
      Bibliographic Description Markup Language (BDML)
      AUTHOR: Medveczki M. (marianna@inf.unideb.hu)
      DOCUMENT TYPE: MONOGRAPHIC PUBLICATIONS (based on MSZ 3424/1)
      VERSION: 2.5 (2005.08.15.) -->

<!-- MAIN STRUCTURE -->

<!-- Bibliographic description is document element -->
<!ELEMENT mp_bibl_desc (single_level|multi_level|component_parts)>
<!ATTLIST mp_bibl_desc
      id ID #REQUIRED
      autopsy (y|n) "y">

<!ELEMENT single_level (single|single_volume)>

<!ELEMENT single (area1,(area2)?,area3,area4,(area5)?,area6,area7)>
```

A `<single>` elem tartalmazza a könyvek monografikus szintű leírását, amelyet olyan könyvekről készítünk, amely egy kötetben, azaz egy fizikai egységben jelenik meg. Maga a `<single>` elem további hét nagyobb egységre tagolódik a bibliográfiai leírás egyes adatszoportjainak megfelelően:

`<area1>` a cím és szerzőségi adatoknak megfeleltett XML-elem, ahol az intellektuális tartalom kifejező címek és a létrehozásáért felelős, abban közreműködő természetes és jogi személyiségek megadása történhet;

```
<!ENTITY % repeated_first_area "(point, first_area)*, (colon,
subtitle)*, (colon, othertitle)*">
<!-- Publication can contain individual works by different authors,
unless a linking word or phrase is given in the publication -->
<!ELEMENT area1 (newindent, char_position3., first_area,
%repeated_first_area;)>
```

<area2> a könyv kiadásával kapcsolatos információkat leíró adatelemeket tartalmazza;

```
<!ELEMENT area2 (pointspacedashspace, edition_statement,  
%opt_statements_of_responsibility;)>
```

<area3> a könyv megjelenéséhez kapcsolható adatokat tartalmazza;

```
<!ELEMENT area3 (pointspacedashspace, %publisher;,  
%dates;, %printer;)>
```

<area4> a könyv terjedelmével és fizikai kiterjedésével kapcsolatos információk;

```
<!ELEMENT area4 (pointspacedashspace, extent, (colon,  
illustration)?, semicolon, dimensions, (plus,  
accompanying_material)?)>
```

<area5> a sorozat adatainak megadására szolgál, amennyiben a könyv sorozat tagjaként jelent meg;

```
<!ELEMENT area5 (pointspacedashspace, o_curvedbracket, series,  
c_curvedbracket, (point, o_curvedbracket, series,  
c_curvedbracket)*)>
```

<area6> a könyvre, az abban megjelent szellemi alkotásra illetve a bibliográfiai leírás adatait kiegészítő, módosító megjegyzésekre szolgál;

```
<!ELEMENT area6 ((pointspacedashspace, note, (pointspacedashspace,  
note)*) | ((newindent, char_position3.), note, ((newindent,  
char_position3.), note)*)>
```

<area7> a könyv terjesztésével, elérhetőségével és az ISBN (azaz a könyvek nemzetközi azonosító száma) számmal kapcsolatos információk.

```
<!ELEMENT area7 ((pointspacedashspace, elements, (point, elements)*) | ((newindent, char_position3.), elements, ((newindent, char_position3.), elements)*))>
```

A teljes BDML DTD struktúrát a Függelék 8. számú melléklete tartalmazza.

## **4.4. A BDML struktúrájának példákkal illusztrált bemutatása a könyvek monografikus szintű leírása esetén**

### **4.4.1. A leírás 1. adatszoportja, az <area1> elem**

A bibliográfiai leírás 1. adatszoportjában a létező formátumok esetén a legnagyobb problémát a párhuzamos adatok, valamint gyűjteményes kiadványok esetén a gyűjteményt alkotó művek cím és szerzőségi adatainak helyes kezelése jelenti.

A katalogizálási szabvány szerint a párhuzamos adat a bibliográfiai adat megismétlése más nyelven, vagy más írásrendszerben, az utóbbi esetben a párhuzamos bibliográfiai adat transliterációjára is szükség van a bibliográfiai leírás készítésekor. A szabvány előírja, hogy amennyiben az alcím, egyéb címadat vagy szerzőségi adatok párhuzamosan szerepelnek a feltárandó dokumentum forráshelyein, akkor ezeket a velük megegyező nyelvű és írásrendszerű főcím, illetve párhuzamos főcím után kell közölni a leírásban. Ezt nevezzük a nyelvenkénti csoportképzés elvének.

Ez az elv viszonylag egyszerűen alkalmazható abban az esetben, ha az adatszoport egy-egy adateleme van csak meg párhuzamosan. Az igazán

problémás eseteket a leírásban a teljes és a részleges párhuzamosság, illetve ennek helyes kezelése jelenti.

Nézzünk egy példát a részleges párhuzamosságra.

Dühöngő ifjúság : színmű három felvonásban = Look back in anger : a play in three acts / John Osborne ; [ford. Ottlik Géza]

Az előző két sorban részleges párhuzamosság hagyományos bibliográfiai leírására látunk egy példát, amikor a főcím és alcím párhuzamosan, angol nyelven is szerepel a dokumentumban, viszont a szerzőségi adatok (szerző és fordító) egynyelvűek, azaz nem párhuzamosak. Jól látszik, hogy a megegyező nyelvű címadatok egymás mellé kerültek, egynyelvű csoportokat képeznek.

Nézzük meg ugyanezt a példát USMARC formátumban:

```
245 10 ▲aDühöngő ifjúság :▲bszínmű három felvonásban = Look back in  
anger : a play in three acts /▲cJohn Osborne ; [ford. Ottlik  
Géza]
```

Azon túlmenően, hogy az egyezményes jelek a 245-ös adatmező almezőinek értékeivel keverednek, számos más problémát is láthatunk. A ▲b alcím, egyéb címadat almezőbe belezsúfolódik a magyar nyelvű alcímen kívül a párhuzamos főcím és a párhuzamos alcím. Az adatok keveredése mellett a párhuzamos főcím így elveszti visszakereshetőségét, ahhoz, hogy ezt a hiányt feloldjuk, a párhuzamos főcímet szerepeltetnünk kell még egyszer a bibliográfiai rekord 741-es adatmezőjében. Hasonló a gond a szerző és fordító visszakereshetősége körül az adatok keveredése miatt.

Következzen az előző példa HUNMARC formátumban:

```
245 10 $aDühöngő ifjúság$b színmű három felvonásban$ALook back in anger$Ba  
play in three acts$cJohn Osborne$e [ford. Ottlik Géza]
```

A változás a USMARC formátumú példához képest a bibliográfiai adatelemeknek megfelelő almező tagolás, valamint az egyezményes jelek közül csak a megengedett forrást jelölő szögletes zárójel maradt az adatmezőben. Az alcímek visszakereshetősége HUNMARC-ban sem megoldott. Hasonlóan a USMARC formátumhoz a \$b almezőbe nemcsak az alcímek, hanem az egyéb címadatok is

kerülhetnek, ugyanígy a \$B almezőben is a párhuzamos alcímek és egyéb címadatok tekintetében.

A HUNMARC szabályzat a párhuzamos adatok leírását nem teszi kötelezővé a 245-ös adatmezőben, csak akkor indokolt így eljárni, ha a szabványos bibliográfiai leírás megjelenítése nem lehetséges a párhuzamos főcím 741 mezőben történő leírásával. Ebben az adatmezőben azonban csak a párhuzamos címadatok megadása lehetséges, a párhuzamos szerzőségi adatok kifejezésére azonban nincs mód. Ha 245-ös adatmezőben a párhuzamos szerzőségi adatok leírása elmarad, akkor ezek az adatok a bibliográfiai rekordok, illetve az információ visszakeresés számára elvesztek.

BDML-ben a következőképpen alakul példánk szerkezete:

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE area1 SYSTEM "BDML_DTD_AREA1.dtd">
<area1>
  <newindent/>
  <char_position3./>
  <first_area>
    <farea>
      <titles>
        <titleproper>
          <title>Dühöngő ifjúság</title>
        </titleproper>
        <colon/>
        <subtitle>színmű három felvonásban</subtitle>
        <equals_sign/>
        <ptitleproper>
          <ptitle>Look back in anger</ptitle>
        </ptitleproper>
        <colon/>
        <psubtitle>a play in three acts</psubtitle>
      </titles>
      <diagonal_slash/>
      <statements_of_responsibility>
        <statement_of_responsibility>
          <statresp>John Osborne</statresp>
        </statement_of_responsibility>
        <semicolon/>
        <statement_of_responsibility>
          <statresp>
            <statresp.osource>
              <o_squarebracket/>
              <statresp.os>ford. Ottlik Géza</statresp.os>
              <c_squarebracket/>
            </statresp.osource>
          </statresp>
        </statement_of_responsibility>
      </statements_of_responsibility>
    </farea>
  </first_area>
</area1>
```

BDML-ben a természetes nyelvű jelölések használatával a bibliográfiai leírás/rekord értelmezése egyszerűbb, és nem feltétlenül szükséges a BDML DTD belső szerkezetének alapos ismerete. Ezzel szemben, az előzőekben már láttuk, hogy a USMARC és HUNMARC formátummal történő bibliográfiai feldolgozáshoz nem elegendő csupán a bibliográfiai leírás elveinek és szabályainak megfelelő ismerete és gyakorlata, ezen túlmenően ismerni kell e formátumok speciális felépítését, a különféle kódértékek, kódolt információk helyes használatát.

A BDML DTD kialakításánál elsődleges szempont volt, hogy a kialakítandó szerkezet tükrözze, visszaadja a már jól bevált, hagyományos bibliográfiai leírás struktúráját. Ha összehasonlítjuk a tárgyalt példa hagyományos és BDML-beli leírását, valamint az alkalmazott leírási elveket, nem találunk különbséget.

A következő példában szintén részleges párhuzamosság figyelhető meg, itt azonban már a szerző is párhuzamos adat. A további közreműködők egynyelvű szerzőségi adatok. Így a „csoportképzés” elvét kiterjesztjük a szerzőre is, míg az egynyelvű közreműködők adatait az adatcsoport végére „vetjük”.

A dokumentum 1. adatcsoportjának hagyományos bibliográfiai leírása:

A Budapesti Történeti Múzeum kiadványai és munkatársainak szakirodalmi tevékenysége 1976-1987 : bibliográfia / Schwarz Katalin = Veröffentlichungen des Budapester Historischen Museums und wissenschaftliche Tätigkeit der Mitarbeiter zwischen 1976-1987 : Bibliographie / Katalin Schwarz ; [szerk. Havassy Péter] ; [fotók Tihanyi Bence]

USMARC formátumban:

245 12 ▲aA Budapesti Történeti Múzeum kiadványai és munkatársainak szakirodalmi tevékenysége 1976-1987 :▲bbibliográfia /▲cSchwarz Katalin = Veröffentlichungen des Budapester Historischen Museums und wissenschaftliche Tätigkeit der Mitarbeiter zwischen 1976-1987 : Bibliographie / Katalin Schwarz ; [szerk. Havassy Péter] ; [fotók Tihanyi Bence]

Itt valamennyi párhuzamos adat és szerzőségi adat a ▲c almezőbe került.

## HUNMARC formátumban:

245 12 \$aA Budapesti Történeti Múzeum kiadványai és munkatársainak szakirodalmi tevékenysége 1976-1987\$b bibliográfia\$c Schwarz Katalin \$A Veröffentlichungen des Budapester Historischen Museums und wissenschaftliche Tätigkeit der Mitarbeiter zwischen 1976-1987\$B Bibliographie\$H Katalin Schwarz\$e [szerk. Havassy Péter]\$e [fotók Tihanyi Bence]

## HUNMARC-ban az adatok helyes tagolása és sorrendje megoldható.

## BDML-ben:

```
<?xml version="1.0"?>
<!DOCTYPE areal SYSTEM "BDML_DTD_AREA1.dtd">
<areal>
  <newindent/>
  <char_position3./>
  <first_area>
    <farea>
      <titles>
        <titleproper>
          <title>A Budapesti Történeti Múzeum kiadványai és
munkatársainak szakirodalmi tevékenysége 1976-1987</title>
          </titleproper>
          <colon/>
          <subtitle>bibliográfia</subtitle>
        </titles>
        <diagonal_slash/>
        <statements_of_responsibility>
          <statement_of_responsibility>
            <statresp>Schwarz Katalin</statresp>
          </statement_of_responsibility>
        </statements_of_responsibility>
      </farea>
      <equals_sign/>
      <farea>
        <titles>
          <ptitleproper>
            <ptitle>Veröffentlichungen des Budapester Historischen Museums
und wissenschaftliche Tätigkeit der Mitarbeiter zwischen 1976-1987</ptitle>
            </ptitleproper>
            <colon/>
            <psubtitle>Bibliographie</psubtitle>
          </titles>
          <diagonal_slash/>
          <statements_of_responsibility>
            <statement_of_responsibility>
              <statresp>Katalin Schwarz</statresp>
            </statement_of_responsibility>
            <semicolon/>
            <statement_of_responsibility>
              <statresp>
                <statresp.ospace>
                  <o_squarebracket/>
                  <statresp.os>szerk. Havassy Péter</statresp.os>
                  <c_squarebracket/>
                </statresp.ospace>
              </statresp>
            </statement_of_responsibility>
          </statements_of_responsibility>
        </farea>
      </areal>
    </first_area>
  </areal>
</BDML>
```

```

        </statresp>
    </statement_of_responsibility>
</semicolon/>
<statement_of_responsibility>
    <statresp>
        <statresp.osource>
            <o_squarebracket/>
            <statresp.os>fotók Tihanyi Bence</statresp.os>
            <c_squarebracket/>
        </statresp.osource>
    </statresp>
</statement_of_responsibility>
</statements_of_responsibility>
</farea>
</first_area>
</areal>

```

BDML-ben a párhuzamosságból eredő gondok nem jelentenek problémát.

A részleges és teljes párhuzamosság előfordulás variációit a bibliográfiai adatok vonatkozásában a <first\_area> és a <titles> elemekhez megadott struktúra kezeli le.

Gyűjteményes kiadásoknak nevezzük azokat a kiadványokat, amelyekben egynél több szellemi alkotás kap helyet. Ezek a művek kapcsolódhatnak egy szerzőhöz, ekkor beszélünk szerzői gyűjteményről, illetve különböző szerzők művei is kerülhetnek egy kiadványba meghatározott szempontok figyelembevételével, ekkor vegyes gyűjteményről van szó.

A következő példában közös főcím nélküli szerzői gyűjtemény 1. adatszoportjának leírása szerepel:

```

    Sakknovella ; Ámok ; Égő titok ; Érzések zűrzavara / Stefan Zweig ;
    [ford. Fónagy István [!Iván] et al.] ; [szerk. Babucs Éva]

```

USMARC formátumban:

```

245 10 ▲aSakknovella ; Ámok ; Égő titok ; Érzések zűrzavara /▲cStefan Zweig ;
    [ford. Fónagy István [!Iván] et al.] ; [szerk. Babucs Éva]

```

USMARC formátumban valamennyi, a gyűjteményt alkotó művek főcímei ugyanabba az \$a almezőbe kerülnek, így módon ez az \$a almező gyakorlatilag csak az első mű főcíméről szól, azt teszi visszakereshetővé, a második indikátor érték is ehhez

rendelődik. Tovább nehezítené a rekord elkészítését, ha az egyes mű főcímekekhez további címadatok is kapcsolódnának, például alcímek, egyéb címadatok vagy párhuzamos címadatok. A közöttük fennálló kapcsolatok helyes kifejezése USMARC formátumban nem lehetséges.

HUNMARC formátumban:

```
245 10 $aSakknovella
740 01 $aÁmok
740 01 $aÉgő titok
740 01 $aÉrzések zűrzavara$cStefan Zweig$e[ford. Fónagy István [!Iván] et
    al.]$e[szerk. Babucs Éva]
```

Mivel 245 és 740 (további cím és szerzőségi közlés) \$a almezői nem ismétlődhetnek egy bibliográfiai rekordon belül, és maga 245-ös adatmező sem ismétlődhet, ezért a szerzői gyűjteményt alkotó művek kifejezésére a 740 adatmező ismétlése ad megoldást. S bár valamennyi példában szereplő mű Stefan Zweig tollából való, a HUNMARC bibliográfiai rekordból ez a kapcsolat csak az utolsó mű esetén derül ki. A megjelenítési formátumoknál némi programozási munka segítségével ez a hiányosság kiküszöbölhető.

HUNMARC formátumban az egyes művek címadatai (főcím, alcím, egyéb címadat és ezek párhuzamos megfelelői) közötti kapcsolat kifejezése megoldható a bibliográfiai leírás 1. adatcsoportjának különböző adatmezőkre megbontott szerkezetéből következően.

BDML-ben:

```
<?xml version="1.0"?>
<!DOCTYPE area1 SYSTEM "BDML_DTD_AREA1.dtd">
<area1>
  <newindent/>
  <char_position3./>
  <first_area>
    <farea>
      <titles>
        <titleproper>
          <title>Sakknovella</title>
        </titleproper>
      </titles>
      <semicolon/>
      <titles>
        <titleproper>
```

```

        <title>Ámok</title>
      </titleproper>
    </titles>
  <semicolon/>
  <titles>
    <titleproper>
      <title>Égő titok</title>
    </titleproper>
  </titles>
  <semicolon/>
  <titles>
    <titleproper>
      <title>Érzések zűrzavara</title>
    </titleproper>
  </titles>
  <diagonal_slash/>
  <statements_of_responsibility>
    <statement_of_responsibility>
      <statresp>Stefan Zweig</statresp>
    </statement_of_responsibility>
    <semicolon/>
    <statement_of_responsibility>
      <statresp>
        <statresp.osource>
          <o_squarebracket/>
          <statresp.os>ford. Fónagy István [!Iván] et
al.</statresp.os>
        <c_squarebracket/>
        </statresp.osource>
      </statresp>
    </statement_of_responsibility>
    <semicolon/>
    <statement_of_responsibility>
      <statresp>
        <statresp.osource>
          <o_squarebracket/>
          <statresp.os>szerk. Babucs Éva</statresp.os>
          <c_squarebracket/>
        </statresp.osource>
      </statresp>
    </statement_of_responsibility>
  </statements_of_responsibility>
</farea>
</first_area>
</areal>

```

A USMARC és HUNMARC rekordoknál tárgyalt problémák BDML-ben nem jönnek elő, a probléma kezelését a <farea> elem teszi.

A következő példa közös főcím nélküli vegyes gyűjteményre vonatkozik. A példában szereplő dokumentum négy tanulmányt tartalmaz négy különböző szerzőtől. A kötetben szereplő második tanulmánynak van egy saját közreműködője, valamint az egész kötethez kapcsolódnak még ezen kívül közös közreműködők is.

A dokumentum 1. adatcsoportjának hagyományos bibliográfiai leírása:

Dió / írta ... Szentiványi Péter. Mandula / írta ... Pejovics Bogdán ; [... sajtó alá rend... Horn Ede]. Mogyoró / írta ... Horn Ede. Gesztenye / írta ... Szentiványi Péter ; [a kötetet szerk... Horn Ede] ; [... közrem... Körtvély Attila] ; [a rajzokat Nagy Zoltánné kész.]

#### USMARC formátumban:

```
245 10 ▲aDió /▲círta ... Szentiványi Péter. Mandula / írta ... Pejovics
Bogdán ; [... sajtó alá rend... Horn Ede]. Mogyoró / írta ... Horn Ede.
Gesztenye / írta ... Szentiványi Péter ; [a kötetet szerk... Horn Ede]
; [... közrem... Körtvély Attila] ; [a rajzokat Nagy Zoltánné kész.]
```

Az első tartalmazott mű szerzőjétől kezdődően minden további mű cím- és szerzőségi adatai, valamint a közös közreműködők is a ▲c almezőbe kerülnek.

#### HUNMARC formátumban:

```
245 10 $aDió$círta ... Szentiványi Péter
740 01 $aMandula$círta ... Pejovics Bogdán$e[... sajtó alá rend... Horn
Ede]$82
740 01 $aMogyoró$círta ... Horn Ede$83
740 01 $aGesztenye$círta ... Szentiványi Péter$e[a kötetet szerk... Horn
Ede]$e[... közrem... Körtvély Attila]$e[a rajzokat Nagy Zoltánné
kész.]$84
```

Hasonlóan a szerzői gyűjteménynél látott megoldáshoz, a második tartalmazott műtől kezdődően az egyes művek cím- és szerzőségi adatai 740 mezőkbe kerülnek. Az egész könyvhöz kapcsolódó közös szerzőségi adatok az utolsó 740 mezőben kifejezett mű adataihoz kerülnek.

#### BDML-ben:

```
<?xml version="1.0"?>
<!DOCTYPE area1 SYSTEM "BDML_DTD_AREA1.dtd">
<areal>
  <newindent/>
  <char_position3./>
  <first_area>
    <farea>
      <titles>
        <titleproper>
          <title>Dió</title>
        </titleproper>
      </titles>
      <diagonal_slash/>
      <statements_of_responsibility>
```

```

        <statement_of_responsibility>
          <statresp>írta ... Szentiványi Péter</statresp>
        </statement_of_responsibility>
      </statements_of_responsibility>
    </farea>
  </first_area>
<point/>
<first_area>
  <farea>
    <titles>
      <titleproper>
        <title>Mandula</title>
      </titleproper>
    </titles>
    <diagonal_slash/>
    <statements_of_responsibility>
      <statement_of_responsibility>
        <statresp>írta ... Pejovics Bogdán</statresp>
      </statement_of_responsibility>
      <semicolon/>
      <statement_of_responsibility>
        <statresp>
          <statresp.osource>
            <o_squarebracket/>
            <statresp.os>... sajtó alá rend... Horn
Ede</statresp.os>
          <c_squarebracket/>
        </statresp.osource>
      </statresp>
    </statement_of_responsibility>
  </statements_of_responsibility>
</farea>
</first_area>
<point/>
<first_area>
  <farea>
    <titles>
      <titleproper>
        <title>Mogyoró</title>
      </titleproper>
    </titles>
    <diagonal_slash/>
    <statements_of_responsibility>
      <statement_of_responsibility>
        <statresp>írta ... Horn Ede</statresp>
      </statement_of_responsibility>
    </statements_of_responsibility>
  </farea>
</first_area>
<point/>
<first_area>
  <farea>
    <titles>
      <titleproper>
        <title>Gesztenye</title>
      </titleproper>
    </titles>
    <diagonal_slash/>
    <statements_of_responsibility>
      <statement_of_responsibility>
        <statresp>írta ... Szentiványi Péter</statresp>
      </statement_of_responsibility>
      <semicolon/>
      <statement_of_responsibility>
        <statresp>

```

```

        <statresp.osource>
          <o_squarebracket/>
          <statresp.os>a kötetet szerk... Horn
Ede</statresp.os>
        <c_squarebracket/>
      </statresp.osource>
    </statresp>
  </statement_of_responsibility>
<semicolon></semicolon>
<statement_of_responsibility>
  <statresp>
    <statresp.osource>
      <o_squarebracket/>
      <statresp.os>... közrem... Körtvély
Attila</statresp.os>
    <c_squarebracket/>
  </statresp.osource>
</statresp>
</statement_of_responsibility>
<semicolon/>
<statement_of_responsibility>
  <statresp>
    <statresp.osource>
      <o_squarebracket/>
      <statresp.os>a rajzokat Nagy Zoltánné
kész.</statresp.os>
    <c_squarebracket/>
  </statresp.osource>
</statresp>
</statement_of_responsibility>
</statements_of_responsibility>
</farea>
</first_area>
</areal>

```

BDML-ben az <first\_area> elem ismételhősége és szerkezete adja a megoldást a problémára.

A következő példa, hasonlóan az előzőhöz, közös főcím nélküli vegyes gyűjteményről szól. Itt azonban megjelenik még egy, valamennyi tartalmazott műre vonatkozó közös alcím. Hagyományos bibliográfiai leírásban ennek kifejezése a következő:

Bánk bán / Katona József. Mózes / Madách Imre : drámák

USMARC formátumban:

245 10 ▲aBánk bán /▲cKatona József. Mózes / Madách Imre : drámák

HUNMARC formátumban:

245 10 \$a Bánk bán \$cKatona József  
740 01 \$aMózes\$cMadách Imre\$bdrámák\$82

BDML-ben:

```
<?xml version="1.0"?>
<!DOCTYPE area1 SYSTEM "BDML_DTD_AREA1.dtd">
<area1>
  <newindent/>
  <char_position3./>
  <first_area>
    <farea>
      <titles>
        <titleproper>
          <title>Bánk bán</title>
        </titleproper>
      </titles>
      <diagonal_slash/>
      <statements_of_responsibility>
        <statement_of_responsibility>
          <statresp>Katona József</statresp>
        </statement_of_responsibility>
      </statements_of_responsibility>
    </farea>
  </first_area>
  <point/>
  <first_area>
    <farea>
      <titles>
        <titleproper>
          <title>Mózes</title>
        </titleproper>
      </titles>
      <diagonal_slash/>
      <statements_of_responsibility>
        <statement_of_responsibility>
          <statresp>Madách Imre</statresp>
        </statement_of_responsibility>
      </statements_of_responsibility>
    </farea>
  </first_area>
  <colon/>
  <subtitle>drámák</subtitle>
</area1>
```

BDML-ben az <area1> elem ad lehetőséget a közös főcím nélküli vegyes gyűjtemények esetén a könyv egészére vonatkozó közös alcímek és egyéb cím adatok szabványos kifejezésére.

#### 4.4.2. A leírás 2. adatcsoportja, az <area2> elem

A feldolgozott dokumentum kiadásjelzésének megadására szolgál az <area2> elem. Tartalmazza az adott kiadás és a dokumentum más kiadásaitól megkülönböztető adatokat.

A dokumentum 2. adatcsoportjának hagyományos bibliográfiai leírása:

. - 4. bőv. átdolg. kiad.

USMARC formátumban:

250   ▲a4. bőv. átdolg. kiad.

HUNMARC formátumban:

250   \$a4. bőv. átdolg. kiad.

BDML-ben:

```
<?xml version="1.0"?>
<!DOCTYPE area2 SYSTEM "BDML_DTD_AREA2.dtd">
<area2>
  <pointspacedashspace/>
  <edition_statement>4. bőv. átdolg. kiad.</edition_statement>
</area2>
```

A dokumentum 2. adatcsoportjának hagyományos bibliográfiai leírása, amennyiben a kiadásjelzés megengedett forrásból származik:

. - [New ed.]

USMARC formátumban:

250   ▲a [New ed.]

HUNMARC formátumban:

250     \$a[New ed.]

BDML-ben:

```
<?xml version="1.0"?>
<!DOCTYPE area2 SYSTEM "BDML_DTD_AREA2.dtd">
<area2>
  <pointspacedashspace/>
  <edition_statement>
    <edition_statement.osource>
      <o_squarebracket/>
      <edition_statement.os>New ed.</edition_statement.os>
      <c_squarebracket/>
    </edition_statement.osource>
  </edition_statement>
</area2>
```

A dokumentum 2. adatcsoportjának hagyományos bibliográfiai leírása, amennyiben az adott kiadáshoz olyan szerzőségi adat kapcsolható, amely kétséget kizáróan rá vonatkozik:

. - Díszkiad. / Reich Károly illusztrációival

USMARC formátumban:

250     ▲aDíszkiad. /▲bReich Károly illusztrációival

HUNMARC formátumban:

250     \$aDíszkiad.\$bReich Károly illusztrációival

## BDML-ben:

```
<?xml version="1.0"?>
<!DOCTYPE area2 SYSTEM "BDML_DTD_AREA2.dtd">
<area2>
  <pointspacedashspace/>
  <edition_statement>Díszkiad.</edition_statement>
  <diagonal_slash/>
  <statements_of_responsibility>
    <statement_of_responsibility>
      <statresp>Reich Károly illusztrációival</statresp>
    </statement_of_responsibility>
  </statements_of_responsibility>
</area2>
```

A következő arra példa, amikor a kiadásjelzés szerzőségi adatként leírt közlés megengedett forrásból származik.

A dokumentum 2. adatcsoportjának hagyományos bibliográfiai leírása:

. - Új kiad. / [lezárva 2003. aug. 30.]

## USMARC formátumban:

250   ▲aÚj kiad. /▲b[lezárva 2003. aug. 30.]

## HUNMARC formátumban:

250   \$aÚj kiad.\$b[lezárva 2003. aug. 30.]

## BDML-ben:

```
<?xml version="1.0"?>
<!DOCTYPE area2 SYSTEM "BDML_DTD_AREA2.dtd">
<area2>
  <pointspacedashspace/>
  <edition_statement>Új kiad.</edition_statement>
  <diagonal_slash/>
  <statements_of_responsibility>
    <statement_of_responsibility>
      <statresp>
        <statresp.osource>
          <o_squarebracket/>
          <statresp.os>lezárva 2003. aug. 30.</statresp.os>
          <c_squarebracket/>
        </statresp.osource>
      </statresp>
    </statement_of_responsibility>
  </statements_of_responsibility>
</area2>
```

```
</statements_of_responsibility>  
</area2>
```

#### 4.4.3. A leírás 3. adatcsoportja, az <area3> elem

A Megjelenés adatcsoportjának felel meg az <area3> elem. Tartalmazza a kiadóra, illetve a terjesztőre, valamint a nyomdára, ezek székhelyére (telephelyére) és a megjelenés évére vonatkozó adatok összességét.

Az első példa a megjelenési adatok egy egyszerű előfordulását mutatja be.

A dokumentum 3. adatcsoportjának hagyományos bibliográfiai leírása:

. - Budapest : Akadémiai K., 2002

USMARC formátumban:

260      ▲aBudapest :▲bAkadémiai K., ▲c2002

HUNMARC formátumban:

260 0 \$aBudapest\$bAkadémiai K.\$c2002

BDML-ben:

```
<?xml version="1.0"?>  
<!DOCTYPE area3 SYSTEM "BDML_DTD_AREA3.dtd">  
<area3>  
  <pointspacedashspace/>  
  <place_of_publ>Budapest</place_of_publ>  
  <colon/>  
  <name_of_publ>Akadémiai K.</name_of_publ>  
  <comma/>  
  <date_of_publ>2002</date_of_publ>  
</area3>
```

A második példában a dokumentum egy kiadónak két, egyenrangú székhelyét nevezi meg, ahol a székhelyek egyértelmű közlése érdekében szükséges a

helységnevek után megkülönböztetésre alkalmas kifejezés (ebben az esetben állam nevek) megadása.

A dokumentum 3. adatcsoportjának hagyományos bibliográfiai leírása:

. - Menlo Park [Calif.] ; Don Mills [Ont.] : Addison-Wesley, 1998

USMARC formátumban:

260     ▲aMenlo Park [Calif.] ;▲aDon Mills [Ont.] :▲bAddison-Wesley,▲c1998

HUNMARC formátumban:

260 0 \$aMenlo Park [Calif.]\$aDon Mills [Ont.]\$bAddison-Wesley\$c1998

BDML-ben:

```
<?xml version="1.0"?>
<!DOCTYPE area3 SYSTEM "BDML_DTD_AREA3.dtd">
<area3>
  <pointspacedashspace/>
  <place_of_publ>Menlo Park [Calif.]</place_of_publ>
  <semicolon/>
  <place_of_publ>Don Mills [Ont.]</place_of_publ>
  <colon/>
  <name_of_publ>Addison-Wesley</name_of_publ>
  <comma/>
  <date_of_publ>1998</date_of_publ>
</area3>
```

A következő példában a dokumentum megegyező székhellyel két kiadót nevez meg, valamint a bibliográfia (a katalógus vagy a dokumentációs szolgáltatás) szempontjai indokolhatják, hogy a megjelenés éve után feltüntessük például a copyright évét, illetve a nyomdára vonatkozó adatokat (nyomda székhelye és megnevezése).

A dokumentum 3. adatcsoportjának hagyományos bibliográfiai leírása:

. - Budapest : Szépirodalmi Kvk. : Magvető, 1998, cop. 1995 (Debrecen : Alföldi Ny.)

USMARC formátumban:

```
260      ▲aBudapest :▲bSzépirodalmi Kvk. :▲bMagvető,▲c1998▲e(Debrecen
        :▲fAlföldi Ny.)
```

HUNMARC formátumban:

```
260 0 $aBudapest$bSzépirodalmi Kvk.$bMagvető$c1998$eDebrecen$fAlföldi Ny.
```

USMARC és HUNMARC formátumban a megjelenés éve mellett a copyright éve, továbbá HUNMARC-ban a nyomtatás éve sem adható meg; csak a megjelenés éve hiányában (azt helyettesítendő) van erre mód. A USMARC legutolsó átdolgozása már lehetővé teszi a nyomtatás évének megadását az újonnan definiált ▲g almezőben.

BDML-ben:

```
<?xml version="1.0"?>
<!DOCTYPE area3 SYSTEM "BDML_DTD_AREA3.dtd">
<area3>
  <pointspacedashspace/>
  <place_of_publ>Budapest</place_of_publ>
  <colon/>
  <name_of_publ>Szépirodalmi Kvk.</name_of_publ>
  <colon/>
  <name_of_publ>Magvető</name_of_publ>
  <comma/>
  <date_of_publ>1998</date_of_publ>
  <comma/>
  <date_of_cop>1995</date_of_cop>
  <o_curvedbracket/>
  <place_of_print>Debrecen</place_of_print>
  <colon/>
  <name_of_print>Alföldi Ny.</name_of_print>
  <c_curvedbracket/>
</area3>
```

BDML-ben a copyright (<date\_of\_cop>) és a nyomtatás éve (<date\_of\_print>) megadásához külön XML-elem áll a rendelkezésre, amelyeket szabadon, szükség szerint használhatunk.

#### 4.4.4. A leírás 4. adatszoportja, az <area4> elem

Az <area4> elem a dokumentum fizikai kiterjedését, fizikai jellemzőit, terjedelmét írja le.

A következő példában egy meglehetősen bőséges adattartalmú 4. adatszoport hagyományos bibliográfiai leírását látjuk:

. - IX, 237, [3] p. : színes illetve, rajzok ; 29 cm + 1 kottafüz. (16 p. ; 16 cm)

USMARC formátumban:

```
300      ▲aIX, 237, [3] p. :▲bszínes illetve, rajzok ;▲c29 cm +▲e1 kottafüz.
        (16 p. ; 16 cm)
```

HUNMARC formátumban:

```
300      $aIX, 237, [3] p.$bszínes illetve, rajzok$c29 cm$e1 kottafüz.$h16
        p.$k16 cm
```

BDML-ben:

```
<!DOCTYPE area4 SYSTEM "BDML_DTD_AREA4.dtd">
<area4>
  <pointspacedashspace/>
  <extent>IX, 237, [3] p.</extent>
  <colon/>
  <illustration>színes illetve, rajzok</illustration>
  <semicolon/>
  <dimensions>29 cm</dimensions>
  <plus/>
  <accompanying_material>
    <accompmat>1 kottafüz.</accompmat>
    <o_curvedbracket/>
    <extent>16 p.</extent>
    <semicolon/>
    <dimensions>16 cm</dimensions>
    <c_curvedbracket/>
  </accompanying_material>
</area4>
```

#### 4.4.5. A leírás 5. adatcsoportja, az <area5> elem

Az <area5> elem a dokumentum sorozatba tartozásával kapcsolatos adatokat mutatja meg.

Az első példában a sorozat főcíme párhuzamosan is szerepel a dokumentumban.

A dokumentum 5. adatcsoportjának hagyományos bibliográfiai leírása:

. - (Pest-Budaer Kurir = Pest-budai hírmondó, ISSN 0238-7751 ; 2.)

USMARC formátumban:

```
440 0 ▲aPest-Budaer Kurir,▲xISSN 0238-7751 ;▲v2.
```

HUNMARC formátumban:

```
440 0 $aPest-Budaer Kurir$xISSN 0238-7751$v2.
```

A bibliográfiai leírás szabványa szerint a sorozati adatcsoportban a sorozat párhuzamos főcímét nem kötelező megadni, csak abban az esetben, ha a párhuzamos nyelv a magyar. Látható azonban, hogy a USMARC és HUNMARC formátumokban nincs adatmező a párhuzamos sorozati főcímek számára, így az előző példánk leírása adatvesztéssel párosult.

BDML-ben:

```
<?xml version="1.0"?>
<!DOCTYPE area5 SYSTEM "BDML_DTD_AREA5.dtd">
<area5>
  <pointspacedashspace/>
  <o_curvedbracket/>
  <series>
    <series_title>Pest-Budaer Kurir</series_title>
    <equals_sign/>
    <pseries_title>Pest-budai hírmondó</pseries_title>
    <comma/>
    <ISSN>ISSN 0238-7751</ISSN>
    <semicolon/>
    <series_numbering>2.</series_numbering>
  </series>
  <c_curvedbracket/>
</area5>
```

BDML-ben a <pseries\_title> elembe kerülnek a párhuzamos sorozati főcímek, és ehhez hasonlóan a párhuzamos sorozati alcímek, egyéb címadatok leírásban történő kifejezésére is külön-külön BDML-elemek készültek.

A második példában szereplő sorozat szerzőségi adattal rendelkezik.

A dokumentum 5. adatcsoportjának hagyományos bibliográfiai leírása:

. - (Társadalom és művelődéstörténeti tanulmányok / sorozatszerk. Glatz Ferenc, ISSN 0018-7251 ; 6.)

USMARC formátumban:

```
440 0 ▲aTársadalom és művelődéstörténeti tanulmányok,▲xISSN 0018-7251 ;▲v6.
```

HUNMARC formátumban:

```
440 0 $aTársadalom és művelődéstörténeti tanulmányok$xISSN 0018-7251$v6.
```

USMARC és HUNMARC formátumokban sorozathoz nem kapcsolható szerzőségi adat, csak abban az esetben, ha a sorozat generikus főcímmel rendelkezik, ekkor a \$z almező használható.

BDML-ben:

```
<?xml version="1.0"?>
<!DOCTYPE area5 SYSTEM "BDML_DTD_AREA5.dtd">
<area5>
  <pointspacedashspace/>
  <o_curvedbracket/>
  <series>
    <series_title>Társadalom és művelődéstörténeti
tanulmányok</series_title>
    <diagonal_slash/>
    <statements_of_responsibility>
      <statement_of_responsibility>
        <statresp>szorozatszerk. Glatz Ferenc</statresp>
      </statement_of_responsibility>
    </statements_of_responsibility>
    <comma/>
    <ISSN>ISSN 0018-7251</ISSN>
    <semicolon/>
    <series_numbering>6.</series_numbering>
  </series>
```

<c\_curvedbracket/>  
</area5>

BDML-ben szükség szerint és a bibliográfia (a katalógus vagy a dokumentációs szolgáltatás) szempontjai szerint közölhetünk sorozati szerzőségi adatokat, akár többet is a konkrét dokumentumtól és elvárásoktól függően. A szerzőségi adatok kezelésének elvei és szabályai a sorozati adatcsoportban, azaz az <area5> elemben teljes mértékben megegyeznek az első adatcsoportnál, azaz az <area1> elemnél már megismertekkel.

A következő példában szereplő sorozat összetett, azaz fő és alárendelt részből áll. Az alsorozathoz alsorozati megjelölő és alsorozati cím is kapcsolódik, valamint ISSN és alsorozati szám is.

A dokumentum 5. adatcsoportjának hagyományos bibliográfiai leírása:

. - (István Király Múzeum közleményei. D sorozat, Múzeumi vezetők, katalógusok, ISSN 0586-3759 ; 36.)

USMARC formátumban:

```
440 0 ▲aIstván Király Múzeum közleményei.▲nD sorozat,▲pMúzeumi vezetők,  
katalógusok,▲xISSN 0586-3759 ;▲v36.
```

HUNMARC formátumban:

```
440 0 $aIstván Király Múzeum közleményei$nD sorozat$pMúzeumi vezetők,  
katalógusok$xISSN 0586-3759$v36.
```

A USMARC és HUNMARC rekordokból nem derül ki, hogy az ISSN és az alsorozati szám az alsorozathoz, nem pedig a főorozathoz tartozik, hiszen ugyanabba az x és v almezőkbe kerülnek, mint az előző példában, ahol egyszerű sorozatról volt szó.

BDML-ben:

```
<?xml version="1.0"?>
<!DOCTYPE area5 SYSTEM "BDML_DTD_AREA5.dtd">
<area5>
  <pointspacedashspace/>
  <o_curvedbracket/>
  <series>
    <series_title>István Király Múzeum közleményei</series_title>
    <equals_sign/>
    <pseries_title>Bulletin du Musée Roi Saint-Etienne</pseries_title>
    <point/>
    <subseries_design>D. sorozat</subseries_design>
    <comma/>
    <subseries_title>Múzeumi vezetők, katalógusok</subseries_title>
    <comma/>
    <ISSN>ISSN 0586-3759</ISSN>
    <semicolon/>
    <subseries_numbering>36.</subseries_numbering>
  </series>
  <c_curvedbracket/>
</area5>
```

A BDML az alsorozati ISSN-hez és alsorozati számhoz saját, önálló elemeket ad meg.

A következő példában ismét összetett sorozat szerepel. Az előzőhöz képest ez annyiban más, hogy mind a fősorozat főcíme, mind pedig az alsorozat megjelölője szerepel a dokumentumban párhuzamosan, így alkalmazható a „csoportképzés” elve.

A dokumentum 5. adatcsoportjának hagyományos bibliográfiai leírása:

```
. - (WHO Regionális Iroda kiadványai. Európai sorozat = WHO regional
publications. European series ; 15.)
```

USMARC formátumban:

```
440 0 ▲aWHO Regionális Iroda kiadványai.▲nEurópai sorozat ;▲v15.
```

HUNMARC formátumban:

```
440 0 $aWHO Regionális Iroda kiadványai$nEurópai sorozat$v15.
```

A USMARC és HUNMARC rekordokból a korábban már említett okok miatt a párhuzamos adatok elmaradnak.

BDML-ben:

```
<?xml version="1.0"?>
<!DOCTYPE area5 SYSTEM "BDML_DTD_AREA5.dtd">
<area5>
  <pointspacedashspace/>
  <o_curvedbracket/>
  <series>
    <series_title>WHO Regionális Iroda kiadványai</series_title>
    <point/>
    <subseries_design>Európai sorozat</subseries_design>
    <equals_sign/>
    <pseries_title>WHO regional publications</pseries_title>
    <point/>
    <psubseries_design>European series</psubseries_design>
    <semicolon/>
    <subseries_numbering>15.</subseries_numbering>
  </series>
  <c_curvedbracket/>
</area5>
```

BDML-ben a csoportképzés és a párhuzamos adatok közlése megoldható, és a sorszám az alsorozathoz kapcsolódik.

Végül az utolsó sorozati példa olyan összetett sorozatról szól, ahol mind a főszorozat, mind pedig az alsorozat saját ISSN számmal rendelkezik.

A dokumentum 5. adatcsoportjának hagyományos bibliográfiai leírása:

```
. - (Kossuth Köri füzetek, ISSN 1215-964X. Kossuth Lajos emléktúrák, ISSN 1215-8631 ; 3.)
```

USMARC formátumban:

```
440 0 ▲aKossuth Köri füzetek.▲pKossuth Lajos emléktúrák,▲xISSN 1215-8631
;▲v3.
```

HUNMARC formátumban:

```
440 0 $aKossuth Köri füzetek$pKossuth Lajos emléktúrák$xISSN 1215-8631$w3.
```

A USMARC és HUNMARC rekordokban a fősorozati ISSN a 440 adatmezőben nem fejezhető ki. Annak érdekében, hogy ez az ISSN ne maradjon el, a 022 adatmezőben vehető fel.

BDML-ben:

```
<?xml version="1.0"?>
<!DOCTYPE area5 SYSTEM "BDML_DTD_AREA5.dtd">
<area5>
  <pointspacedashspace/>
  <o_curvedbracket/>
  <series>
    <series_title>Kossuth Köri füzetek</series_title>
    <comma/>
    <ISSN>ISSN 1215-964X</ISSN>
    <point/>
    <subseries_title>Kossuth Lajos emléktúrák</subseries_title>
    <comma/>
    <ISSN>ISSN 1215-8631</ISSN>
    <semicolon/>
    <subseries_numbering>3.</subseries_numbering>
  </series>
  <c_curvedbracket/>
</area5>
```

#### 4.4.6. A leírás 6. adatszoportja, az <area6> elem

Az <area6> elemben a feldolgozott dokumentumra vonatkozó megjegyzések adhatók meg, amelyek lehetnek a bibliográfiai leírás adatait kiegészítő, illetve módosító megjegyzések, vagy a leírt példány sajátosságaira, vagy a dokumentum tartalmára, illetve a dokumentum (és a benne foglalt mű, szellemi alkotás) és más kiadványok (és művek) közötti bibliográfiai kapcsolatokra vonatkozó megjegyzések.

## A dokumentum 6. adatcsoportjának hagyományos bibliográfiai leírása:

. - Eredeti cím: Toward a theory of instruction. - A borítófedélen a szerző neve tévesen

### USMARC formátumban:

```
240 10 ▲aToward a theory of instruction▲lmagyar
500 ▲aA borítófedélen a szerző neve tévesen
```

### HUNMARC formátumban:

```
240 10 $aToward a theory of instruction$imagyar
500 $aA borítófedélen a szerző neve tévesen
```

A MARC formátumokban a különféle megjegyzés-típusok kifejezésére különféle, 5XX hívójelű megjegyzés-adatmezők bőséges kínálata áll a rendelkezésre.

Az 500 adatmezőben lehet közölni azokat a dokumentumra vonatkozó megjegyzéseket, amelyek a valamely speciális ismerv alapján kialakított további megjegyzés-mezők (5XX) egyikébe sem sorolhatók be.

Az eredeti címhez kapcsolódó megjegyzés leírására MARC formátumban specifikus megjegyzés-mező nincs kialakítva, az 500 mezőben pedig „nem szokás” az eredeti címet megadni. Amennyiben az eredeti címből már egységesített címet alkottunk, mint besorolási adatot vehetjük fel a 240 adatmezőbe.

### BDML-ben:

```
<?xml version="1.0"?>
<!DOCTYPE area7 SYSTEM "BDML_DTD_AREA6.dtd">
<area6>
  <pointspacedashspace/>
  <note retrieve="y" type="orig_title">Toward a theory of instruction
  </note>
  <pointspacedashspace/>
  <note type="general">A borítófedélen a szerző neve tévesen</note>
</area6>
```

#### 4.4.7. A leírás 7. adatcsoportja, az <area7> elem

Az <area7> elemben a könyvek nemzetközi azonosító száma, illetve az egyéb terjesztési adatok kifejezése lehetséges. Az ISBN-t az MSZ ISO 2108 Információ és dokumentáció. Könyvek nemzetközi azonosító számozása szabvány alapján kell leírni.

A következő példa egy összetettebb terjesztési adatokkal rendelkező dokumentumról szól, ahol a dokumentum két árváltozatot közöl. Az egyes árváltozatokkal kapcsolatos kiegészítő adatok kerek zárójelben, a bibliográfia nyelvén követik azokat, mint az adott árra vonatkozó magyarázatok.

A dokumentum 7. adatcsoportjának hagyományos bibliográfiai leírása:

. - ISBN 963 05 0653 X : 2500,-Ft (kották nélkül) : 3700,-Ft (kottákkal)

USMARC formátumban:

```
020    ▲a963050653X :▲c2500.00Ft (kották nélkül)
020    ▲c3700.00Ft (kottákkal)
```

A MARC formátumokban az azonosító számokhoz kapcsolódó ISBN rövidítést a 020 adatmezőkben nem adjuk meg. USMARC formátumban az ISBN szám jelentéssel bíró számcsoportjainak tagolását nem őrizzük meg, a számjegyek folyamatosan, közvetlenül egymás mellé kerülnek. A 020 adatmező ▲c almezője nem ismételhető, ezért a második árváltozat megadásához egy újabb 020 adatmező alkalmazására van szükség. Ennek akadálya nincs, hiszen a 020 ismételhető adatmező.

HUNMARC formátumban:

```
020    $a963-05-0653-X$c2500,-Ft (kották nélkül)$d3700Ft (kottákkal)
```

A HUNMARC formátum az ISBN szám megadásához a jelentéssel bíró számcsoportok csoportelválasztó jeleként a kötőjelek alkalmazását ajánlja.

HUNMARC formátumban az ismételhető \$d almező ad lehetőséget a dokumentumhoz kapcsolódó különböző árváltozatok közlésére.

BDML-ben:

```
<?xml version="1.0"?>
<!DOCTYPE area7 SYSTEM "BDML_DTD_AREA7.dtd">
<area7>
  <pointspacedashspace/>
  <elements>
    <ISBN>ISBN 963 05 0653 X</ISBN>
    <binding>fűzött</binding>
    <colon/>
    <price>2500,-Ft</price>
    <o_curvedbracket/>
    <text>kották nélkül</text>
    <c_curvedbracket/>
    <colon/>
    <price>3700,-Ft</price>
    <o_curvedbracket/>
    <text>kottákkal</text>
    <c_curvedbracket/>
  </elements>
</area7>
```

Ha egy dokumentum több kiadó közös gondozásában jelent meg, és ezért több ISBN számmal rendelkezik, akkor ezek leírásban történő közlése kötelező. Minden egyes ISBN szám után, kerek zárójelbe téve, a megfelelő kiadó rövidített nevét is meg kell adni.

A dokumentum 7. adatcsoportjának hagyományos bibliográfiai leírása:

```
. - ISBN 963 15 0699 X (Szépirodalmi Kvk.) kötött : 1490,-Ft. ISBN 963 270
512 X (Magvető)
```

USMARC formátumban:

```
020   ▲a963150699X (Szépirodalmi Kvk.) (kötött) :▲c1400.00Ft
020   ▲a963270512X (Magvető)
```

USMARC formátumban az ISBN számhoz tartozó kiadó megnevezése az ▲a almezőben történik. A dokumentum kötésének típusára vonatkozó kifejezések megadásához önálló almező szintén nem áll rendelkezésre a 020 adatmezőben.

## HUNMARC formátumban:

020     \$a963-150-0699-X\$iSzépirodalmi Kvk.\$jKötött\$c1400,-Ft  
020     \$a963-270-512-X\$iMagvető

HUNMARC formátumban mind az ISBN számmal kapcsolatos kiadó nevének, mind pedig a kötés típusának kifejezésére önálló almezők állnak rendelkezésre.

## BDML-ben:

```
<?xml version="1.0"?>
<!DOCTYPE area7 SYSTEM "BDML_DTD_AREA7.dtd">
<area7>
  <pointspacedashspace/>
  <elements>
    <ISBN>ISBN 963 15 0699 X</ISBN>
    <o_curvedbracket/>
    <text>Szépirodalmi Kvk.</text>
    <c_curvedbracket/>
    <binding>kötött</binding>
    <colon/>
    <price>1490 Ft</price>
  </elements>
  <point/>
  <elements>
    <ISBN>ISBN 963 270 512 X</ISBN>
    <o_curvedbracket/>
    <text>Magvető</text>
    <c_curvedbracket/>
  </elements>
</area7>
```

## 4.5. Továbbfejlesztési lehetőségek

A BDML DTD jelenleg a könyv típusú dokumentumok bibliográfiai leírásának formalizálását tartalmazza. Az elkészült DTD továbbfejlesztése mindenképpen szükséges, a továbbfejlesztés lehetséges irányaként most a következők látszanak:

- A BDML DTD-ben kialakított struktúra bővítése, kiterjesztése annak érdekében, hogy további dokumentumtípusok (mint például időszaki kiadványok, kották, elektronikus dokumentumok) bibliográfiai leírása is elkészíthető legyen XML-alapokon.
- A BDML DTD-ben kialakított bibliográfiai leírási struktúra kiegészítése a besorolási adatokhoz, utalásokhoz szükséges további szerkezeti részekkel. Az egységesítés és az adatfelvitel megkönnyítése érdekében, a formátum egyes elemeihez (például személynevek, egységesített címek) kötött listák definiálása, kialakítása lenne szerencsés, amelyek igény szerint bővíthetők és átírhatók lennének.  
Valamint a későbbiekben a dokumentumok példány- és lelőhelyinformációinak kifejezésére szolgáló további XML-elemek bevezetése válik majd szükségessé.
- A besorolási adatokkal kiegészített formátum alkalmassá válhat arra, hogy ezek segítségével a bibliográfiai rekordok hypertext rendszerűen felfűzhetőek legyenek.
- Célszerű lenne a későbbiekben meghatározni, hogy a leírás bibliográfiai szintje szerint az egyszerűsített leírás készítéséhez mely BDML DTD elemekre van szükség.

- A BDML DTD kialakításánál szerzett tapasztalatok hasznosítására, és a további fejlesztésre két aktuálisan futó projektben is lehetőség nyílik.

Az egyik az EMIR (Elektronikus Magyar Irodalom) projekt, amely az Eötvös Loránd Tudományegyetem Bölcsészeti Informatika Önálló Program, a Veszprémi Egyetem Magyar Irodalomtudományi Tanszékének és a Debreceni Egyetem könyvtár tanszékének együttműködése. A projekt célja a kortárs magyar irodalom komplex és sokoldalú online feldolgozása és a nagyközönség számára elérhetővé tétele, egyben egy nagyobb léptékű, összefüggő, modulárisan építkező, közös hozzáférésű, elosztott rendszerre épülő szolgáltatás létrehozása, amely egy rendszerben kapcsolná össze a digitálisan feldolgozott katalógusrendszereket, folyóiratokat, kölcsönzési és vásárlási lehetőségeket.

A másik kezdeményezés a Debreceni Egyetem Zeneműtárához és a Debreceni Egyetem könyvtár tanszékéhez kapcsolódik. A cél ebben az esetben egy digitális zeneműtár létrehozása, amely nemcsak a hangzó dokumentumok bibliográfiai adatait tárolná elektronikus formában, hanem magukat a zenei dokumentumokat is. Az adatokat és a zenei dokumentumokat a felhasználók a könyvtári terminálok segítségével érhetik majd el, a könyvtár belső hálózatáról (erre a korlátozásra a szerzői jogi problémák megelőzése érdekében van szükség).

## **5. ÖSSZEGRZÉS**

Kétségtelen, hogy napjainkra a különféle dokumentumok és források feldolgozásához szabályok, szabványok, kézikönyvek, dokumentációk tekintélyes fegyvertára áll a rendelkezésünkre. A komplex, elektronikus környezet, természetesen jól felkészült szakembereket kíván, akik képesek eligazodni és lépést tartani a gyorsan változó körülményekkel és elvárásokkal.

Mindebből következik, hogy a dokumentumok, információk források körültekintő és alapos feldolgozása meglehetősen költséges feladat. Ez a költség a számítástechnika és a telekommunikációs vívmányok könyvtári alkalmazásával, a munkamegosztást célzó különféle könyvtári programokkal (közös, osztott katalogizálás), már létező bibliográfiai rekordok letöltésével és helyi igényeknek megfelelő módosításával ésszerűbbé tehető, és nem mellékes az sem, hogy ezeknek a rekordoknak a színvonala és a minősége ideális esetben növelhető. Napjainkban azonban még sok időt és energiát követel a különböző könyvtári adatformátumok közötti összhang megteremtése, az egyik adatformátumban szereplő adatok egy másik adatformátumba való konvertálása. Ezen túlmenően emberi beavatkozás szükséges annak eldöntésére is, hogy vajon történt-e adatvesztés az átalakítás során, illetve a rekordban szereplő adatok különféle javításaira. S ha még a weben is szeretnénk hozzáférhetővé tenni a katalógusunkat (lehetőleg minél kevesebb munkával) ?!

Az XML-ben való gondolkodás megoldást jelenthet mindezekre a nehézségekre, hiszen fő célja az internetes információcsere támogatása.

Látható, hogy az XML nem csupán dokumentumok egyszeri megadásában hasznos eszköz, hanem a későbbi, az internet fejlődésével lépést tartó alkalmazások és XML-bővítmények kerete is.

Az XML-technológia használatával a könyvtáraknak lehetőségük nyílik elmozdulni a nyílt forráskódú könyvtári szoftverek felé. Számos előnyt és lehetőséget hordoz magában ez a tendencia: ezek a könyvtári rendszerek olcsóbbak lehetnek, gyorsabban és érzékenyebben reagálhatnak a technikai újításokra és a felhasználói igényekre. Mivel ezeknek a szoftvereknek a forráskódja szabadon hozzáférhető, lehetőség nyílik arra, hogy a könyvtár saját küldetésének, vállalt feladatainak megfelelően szabja azt testre.

Napjainkban már léteznek ilyenek, igaz kísérleti jellegű, XML-technológián alapuló, nyílt forráskódú könyvtári rendszerek, mint például az Avanti, OSDLS/Pytheas, Koha, FSLP és az Open Book. Szükséges lenne integrálni a webopacban a helyi és távoli elérésű digitális tartalmakat a hagyományos hordozón megjelenőekkel.

A bibliográfiai adatfeldolgozás területén a létező szabványokban számos olyan jellemvonás áll a rendelkezésünkre, amelyek jól kihasználhatók az egyre modernebb és fejlettebb információtechnológiai környezetben, ha megfelelően továbbfejlesztjük ezeket.

## **IRODALOMJEGYZÉK**

- [1] Baker, T.: A grammar of Dublin Core. *D-Lib Magazine*, October, 2000. (<http://www.dlib.org/dlib/october00/baker/10baker.html>) dc-agents – Electronic discussion list. <http://www.mailbase.ac.uk/lists/dc-agents/>
- [2] Bakonyi G.: Szürke irodalom – kommunikációs modellen alapuló szövegstruktúrák. *Könyvtári figyelő*, úf., 10. = 46. évf. (2000) 4. sz. p., 603-611.
- [3] Bakonyi G.: *Tartalomszolgáltatás – egy folyóirat digitális feldolgozása*. <http://nws.iif.hu/ncd2000/docs/eloadas/29/index.htm>
- [4] Berke Barnabásné: Párizs 1977 – Koppenhága 1998. A nemzeti bibliográfiai ajánlások megújítása. *Könyv, könyvtár, könyvtáros*, 9. évf. (2000) 8. sz. p., 11-20.
- [5] Bradley, N.: *Az XML-kézikönyv*. Bicske, 2000, Szak.
- [6] Buckland, M.: *Redesigning library services*. A manifesto. Chicago, 1992, ALA
- [7] Budimir, G.: *MARC records and XML*. Infoteka, 5 (2004) 1-2., p. 121-136.
- [8] Burkart, M.: Dokumentations Sprachen. In: Buder, M. - Rehfeld, W. - Seeger, T. (hrsg.): *Grundlagen der praktischen Information und Dokumentation*. München, 1990, K. G. Saur.
- [9] Caplan, P.: *International Metadata Initiative*. [http://www.loc.gov/catdir/bibcontrol/caplan\\_paper.html](http://www.loc.gov/catdir/bibcontrol/caplan_paper.html)
- [10] Caplan, P.: *Metadata fundamentals for all librarians*. Chicago, 2003, ALA.
- [11] *Document Object Model Core*. <http://www.w3.org/TR/2004/REC-DOM-Level-3-Core-20040407/>
- [12] Drótos L.: eleMEK. *Tudományos és műszaki tájékoztatás*, 52. évf. (2005.) 2. sz., p. 55-58.

- [13] *Dublin Core Metadata Initiative*. <http://purl.org/dc/>
- [14] *Elektronikus Periodika Archívum és Adatbázis*. <http://epa.oszk.hu/>
- [15] *Extensible Markup Language (XML) 1.0*. 2. ed. - <http://www.w3.org/TR/REC-xml>
- [16] *Extensible Stylesheet Language (XSL)*, Version 1.0. W3C Recommendation, 2001. <http://www.w3c.org/TR/xsl>
- [17] *Family of ISBDs* : Publication list. International Federation of Library Associations and Institutions, 2003. <http://www.ifla.org/VI/3/nd1/isbdlist.htm>
- [18] Fiander, D. J.: Applying XML to the Bibliographic Description. *Cataloging & Classification Quarterly*, Vol. 33(2) 2001, p. 17-28.
- [19] Hagler, R.: *The bibliographic record and information technology*. 3. ed. Chicago-London, ALA-Ottawa, CLA, 1997.
- [20] Horváth T.: A feldolgozás egységei és a feltárás célja. *Könyvtárosok kézikönyve. 2. Feltárás és visszakeresés*. Szerk. Horváth Tibor, Papp István. Budapest, 2001, Osiris, p. 15-34.
- [21] *HUNMARC*. A bibliográfiai rekordok adatcsere formátuma. Budapest, 2002, OSZK. [http://www.oszk.hu/hun/szakmai/hunmarc/hunmarc\\_index\\_hu.htm](http://www.oszk.hu/hun/szakmai/hunmarc/hunmarc_index_hu.htm)
- [22] IFLA study on functional requirements for bibliographic records. *International Cataloguing and Bibliographic Control*, vol. 24. (1995) no. 3., p. 43-50. <http://www.ifla.org/VII/s13/frbr/frbr.pdf>
- [23] *Indecs Framework Ltd*. <http://www.indecs.org/>
- [24] *Information Retrieval (Z39.50)* : Application Service Definition and Protocol Specification. National Information Standards Organization, 2002. <http://www.niso.org/standards/resources/Z39-50-200x.pdf>
- [25] *International Standard ISO 8879 Information Processing – Text and Office Systems - Standardized Generalized Markup Language (SGML)*, First Edition. International Organization for Standardization, 1996.

- [26] *ISBD (G)*. General international standard bibliographic description. Annotated tex. London, 1977, IFLA Int. Office for UBC. <http://www.ifla.org/VII/s13/pubs/isbdg.htm>
- [27] *ISBD (M)*. International standard bibliographic description for monographic publications. 1. standard ed. London, 1974, IFLA Committee on Cataloguing. [http://www.ifla.org/VII/s13/pubs/isbd\\_m0602.pdf](http://www.ifla.org/VII/s13/pubs/isbd_m0602.pdf)
- [28] ISBD(ER). International Standard Bibliographic Description for Electronic Resources. <http://www.ifla.org/VII/s13/pubs/isbd.htm>
- [29] *ISO 2709:1996*. Information and documentation - Format for information exchange. TC 46/SC 4.
- [30] KSZ/2 Bibliográfiai leírás. Elektronikus dokumentumok. Elektronikus kiad. 2005. [http://www.ki.oszk.hu/szabalyzatok/eldok\\_ksz2.pdf](http://www.ki.oszk.hu/szabalyzatok/eldok_ksz2.pdf)
- [31] *Magyar Elektronikus Könyvtár*. <http://www.mek.oszk.hu>
- [32] *Mandatory data elements for internationally shared resource authority records*. Report of the IFLA UBCIM Working Group on Minimal Level Authority Records and the ISADN. 1998.
- [33] *MARC Standards*. Network Development and MARC Standard Office, Library of Congress, 09/16/2003. <http://lcweb.loc.gov/marc/>
- [34] *MARCXML*. Library of Congress. <http://www.loc.gov/standards/marcxml/>
- [35] *Metadata (UKOLN website)*. <http://www.ukoln.ac.uk/metadata/>
- [36] *Metadata and Resource Description*. <http://www.w3.org/Metadata>
- [37] *Metadata Encoding & Transmission Standard*. <http://www.loc.gov/standards/mets/>
- [38] *Metadata Object Description Schema (MODS)*. Library of Congress, 11/25/2003. <http://www.loc.gov/standards/mods/v3/mods-3-0.xsd>
- [39] Miller, D. R.: *XML and MARC*: a choice or replacement? Chicago, 2000, ALA., 12 p. <http://elane.stanford.edu/laneauth/ALChicago2000.html>
- [40] *MSZ 3424/1*. Bibliográfiai leírás. Könyvek

- [41] Network Development and MARC Standards Office.  
<http://www.loc.gov/marc/ndmso.html>
- [42] *ONIX for Books*. EDItEUR. <http://www.editeur.org/onix.html>
- [43] Poprády G.: Könyvtári trendek. *Könyvtárosok kézikönyve. 5. Segédletek*. Szerk. Horváth Tibor, Papp István. Budapest, 2003, Osiris., p. 13-62.
- [44] Rácz Á.: A kiadványok bibliográfiai számbavétele; leíró katalogizálás. *Könyvtárosok kézikönyve. 2. Feltárás és visszakeresés*. Szerk. Horváth Tibor, Papp István. Budapest, 2001, Osiris., p. 187-296.
- [45] Renkecz A.: Elektronikus időszaki kiadványok nyilvántartása és archiválása. *Tudományos és műszaki tájékoztatás*, 52. évf. (2005.) 2. sz., p. 59-69.
- [46] *Resource Description Framework (RDF)*. <http://www.mozilla.org/rdf/doc/>
- [47] *A syntax for Dublin Core metadata*.  
<http://www.dlib.org/dlib/february00/02contents.html>
- [48] Tennant, R.: Building a new bibliographic infrastructure. *Digital Libraries, Library Journal*. 1/15/2004, p. 38.
- [49] Tennant, R.: MARC must die. *Library Journal* 2002, 10.  
<http://libraryjournal.reviewsnews.com>
- [50] Tennant, R.: Metadata leadership. *Digital Libraries, Library Journal*. 8/15/2004
- [51] *UNICODE Home Page*. <http://www.unicode.org>
- [52] *UNIMARC Manual* : Bibliographic Format 1994. International Federation of Library Associations and Institutions. <http://www.ifla.org/VI/3/p1996-1/sec-uni.htm>
- [53] *UNIMARC Manual* : Bibliographic Format 1994, APPENDIX H : cataloguing Rules and Formats Codes. International Federation of Library Associations and Institutions. <http://www.ifla.org/VI/3/p1996-1/appx-h.htm>
- [54] *USEMARCON Plus* – The Universal MARC Record Converter.  
<http://www.bl.uk/services/bibliographic/usemarcon.html>

- [55] *XML Path Language*. W3C Recommendation, 1999.  
<http://www.w3.org/TR/xpath>
- [56] *XSL Transformations (XSLT)*, Version 1.0. W3C Recommendation, 1999.  
<http://www.w3c.org/TR/xslt>
- [57] Vajda E.: A besorolási adatok egységesítése. Fogalmak tisztázása és terminológiai javaslat. *Tudományos és műszaki tájékoztatás*, 37 évf. (1990) 5. sz., p. 186-190.
- [58] Lagoze, C.: The Warwick Framework : A Container Architecture for Diverse Sets of Metadata. *D-Lib Magazine*, July/August 1996.  
<http://www.dlib.org/dlib/july96/lagoze/07lagoze.html>
- [59] *Web Services Activity*. W3C Recommendation, 2002-2003.  
<http://www.w3.org/2002/ws/>
- [60] *Working Group on Functional Requirements and Numbering of Authority Records (FRANAR)*, 25 July 2005. <http://www.ifla.org/VII/d4/wg-franar.htm>



## **FÜGGELÉK**

### **Glosszárium**

AACR	Anglo-American Cataloguing Rules
BDML	Bibliographic Description Markup Language
CCF	Common Communication Format
CEN	European Committee for Standardization
CERL	Committee of European Research Librarians
DC	Dublin Core
DCME	Dublin Core Metadata Elements
DCMES	Dublin Core Metadata Element Set
DCMI	Dublin Core Metadata Initiative
DOM	Document Object Model
DTD	Document Type Definition
EDI	Electronic Data Interchange
FRBR	Functional Requirements for Bibliographic Records
FRANAR	Functional Requirements and Numbering of Authority Records
GARE	Guidelines for Authority and References Entries
IFLA	International Federation of Library Associations and Institutions
ISBD	International Standard Bibliographic Description
ISBN	International Standard Book Number
ISADN	International Standard of Authority Data Number
HTML	HyperText Markup Language
MARC	Machine Readable Cataloguing
MCF	Meta Content Framework
MODS	Metadata Object Description Schema
OAI	Open Archives Initiative
OCLC	Online Computer Library Center
OPAC	Online Public Access Catalogue

RDF	Resource Description Framework
SAX	Simple API for XML
SGML	Standardized Generalized Markup Language
SVG	Scalable Vector Graphics
UBCIM	Universal Bibliographic Control and International MARC
UKOLN	U.K. Office for Library and Information Networking
UNISIST	United Nations Information System in Science and Technology
URN	Uniform Resource Name
USEMARCON	User Controlled Generic MARC Converter
XML	Extensible Markup Language
XFRML	XML Financial and Reporting Markup Language
W3C	World Wide Web Consortium

## 1. melléklet: BDML\_DTD\_AREA1.dtd

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- Internal entities -->

<!ENTITY % repeated_first_area "(point, first_area)*, (colon, subtitle)*,
(colon, othertitle)*">

<!-- Parallel titles and parallel statements of responsibility -->
<!-- When a title page bears one or more parallel titles and/or parallel
statements of other title information and parallel statements of
responsibility, but one statement of responsibility is in only one
language and/or script, this statement of responsibility, if given, is
given after the last parallel statements of responsibility transcribed. --
>

<!ENTITY % parallel_farea "(equals_sign, farea)*, (semicolon, statresp)*">

<!ENTITY % repeated_titles "(semicolon, titles)*, (colon, subtitle)*,
(colon, othertitle)*">

<!ENTITY % opt_statements_of_responsibility "(diagonal_slash,
statements_of_responsibility)?">

<!ENTITY % pothertitles "(colon, othertitle), (equals_sign,
pothertitle)*">

<!ENTITY % psubtitles "((colon, subtitle), (equals_sign, psubtitle)*)*">

<!ENTITY % ptitlesubtitle "(equals_sign, ptitleproper, (colon,
psubtitle)+)">

<!ENTITY % titlesubtitleothertitle "titleproper, (colon, subtitle)+,
(colon, othertitle)+">

<!ENTITY % ptitlesubtitleothertitle "(((equals_sign, ptitleproper, (colon,
psubtitle)+, (colon, pothertitle)+) | ((ptitleproper, (colon,
psubtitle)*, (colon, pothertitle)*)))">

<!ENTITY % titles_var1 "(titleproper, (equals_sign, ptitleproper)*,
(colon, subtitle)*, (colon, othertitle)*">

<!ENTITY % titles_var2 "(titleproper, (equals_sign, ptitleproper)*,
(colon, subtitle)+, %pothertitles;)">

<!ENTITY % titles_var3 "(titleproper, %psubtitles;, (colon,
othertitle)*">

<!-- When title page bears one or more parallel titles and other title
information in more than one language and/or script, each statement of
other title information, if given, follows the title to which it is
linguistically related. -->
<!ENTITY % titles_var4 "(titleproper, (colon, subtitle)+,
%ptitlesubtitle;, (colon, othertitle)*">
```

```

<!ENTITY % titles_var5 "(%titlesubtitleothertitle;,
%titlesubtitleothertitle;)">

<!ENTITY % titles_var6 "(%titlesubtitleothertitle;)">

<!ENTITY % repeated_statements_of_responsibility "(semicolon,
statement_of_responsibility)*">

<!-- Publication can contain individual works by different authors, unless
a linking word or phrase is given in the publication -->
<!ELEMENT area1
(newindent,char_position3.,first_area,%repeated_first_area;)>

<!-- Parallel titles and parallel statements of responsibility -->
<!ELEMENT first_area (farea,%parallel_farea;)>

<!-- Publication can contain individual works by the same author, unless a
linking word or phrase is given in the publication -->
<!ELEMENT farea
(titles,%repeated_titles;,%opt_statements_of_responsibility;)>

<!ELEMENT titles (%titles_var1; | %titles_var2; | %titles_var3; |
%titles_var4; | %titles_var5; | %titles_var6;)>

<!ELEMENT titleproper (title,(alternative_title)*)>
<!ELEMENT ptitleproper (ptitle,(palternative_title)*)>

<!ELEMENT title (#PCDATA)>
<!ATTLIST title retrieve (y|n) "y"
caption CDATA #FIXED "főcím" >
<!ELEMENT ptitle (#PCDATA)>
<!ATTLIST ptitle retrieve (y|n) "y"
caption CDATA #FIXED "párhuzamos főcím">
<!ELEMENT alternative_title (#PCDATA)>
<!ATTLIST alternative_title retrieve (y|n) "y"
caption CDATA #FIXED "vagylagos cím">
<!ELEMENT palternative_title (#PCDATA)>
<!ATTLIST palternative_title retrieve (y|n) "y"
caption CDATA #FIXED "párhuzamos vagylagos cím">
<!ELEMENT subtitle (#PCDATA)>
<!ATTLIST subtitle retrieve (y|n) "n"
caption CDATA #FIXED "alcím">
<!ELEMENT psubtitle (#PCDATA)>
<!ATTLIST psubtitle retrieve (y|n) "n"
caption CDATA #FIXED "párhuzamos alcím">

<!ELEMENT othertitle (#PCDATA|othertitle.osource)*>
<!ATTLIST othertitle retrieve (y|n) "n"
caption CDATA #FIXED "egyéb címadat">
<!ELEMENT othertitle.osource
(o_squarebracket,othertitle.os,c_squarebracket)>
<!ELEMENT othertitle.os (#PCDATA)>
<!ATTLIST othertitle.os retrieve (y|n) "n"
caption CDATA #FIXED "egyéb címadat megengedett
forrásból">

<!ELEMENT pothertitle (#PCDATA|pothertitle.osource)*>
<!ATTLIST pothertitle retrieve (y|n) "n"
caption CDATA #FIXED "párhuzamos egyéb címadat">
<!ELEMENT pothertitle.osource
(o_squarebracket,pothertitle.os,c_squarebracket)>
<!ELEMENT pothertitle.os (#PCDATA)>
<!ATTLIST pothertitle.os retrieve (y|n) "n"
caption CDATA #FIXED "párhuzamos egyéb címadat
megengedett forrásból">

```

```

<!-- Definition of Statements of Responsibility -->

<!ELEMENT statements_of_responsibility (statement_of_responsibility,
%repeated_statements_of_responsibility;)>
<!ELEMENT statement_of_responsibility (statresp, (equals_sign,pstatresp)*)>

<!ELEMENT statresp (#PCDATA|statresp.osource)*>
<!ATTLIST statresp caption CDATA #FIXED "szerzőségi adat">
<!ELEMENT statresp.osource (o_squarebracket,statresp.os,c_squarebracket)>
<!ELEMENT statresp.os (#PCDATA)>
<!ATTLIST statresp.os caption CDATA #FIXED "szerzőségi adat megengedett
forrásból">

<!ELEMENT pstatresp (#PCDATA|pstatresp.osource)*>
<!ATTLIST pstatresp caption CDATA #FIXED "párhuzamos szerzőségi adat">
<!ELEMENT pstatresp.osource
(o_squarebracket,pstatresp.os,c_squarebracket)>
<!ELEMENT pstatresp.os (#PCDATA)>
<!ATTLIST pstatresp.os caption CDATA #FIXED "párhuzamos szerzőségi adat
megengedett forrásból">

<!-- Definition of Punctuation Symbols -->

<!ELEMENT c_curvedbracket EMPTY>
<!ELEMENT char_position3. EMPTY>
<!ELEMENT colon EMPTY>
<!ELEMENT comma EMPTY>
<!ELEMENT c_squarebracket EMPTY>
<!ELEMENT diagonal_slash EMPTY>
<!ELEMENT newindent EMPTY>
<!ELEMENT o_curvedbracket EMPTY>
<!ELEMENT o_squarebracket EMPTY>
<!ELEMENT plus EMPTY>
<!ELEMENT point EMPTY>
<!ELEMENT pointspacedashspace EMPTY>
<!ELEMENT equals_sign EMPTY>
<!ELEMENT semicolon EMPTY>

```

## 2. melléklet: BDML\_DTD\_AREA2.dtd

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- Definition of Edition Area -->

<!ENTITY % opt_statements_of_responsibility "(diagonal_slash,
statements_of_responsibility)?">
<!ENTITY % repeated_statements_of_responsibility "(semicolon,
statement_of_responsibility)*">

<!-- Definition of Statements of Responsibility -->

<!ELEMENT statements_of_responsibility (statement_of_responsibility,
%repeated_statements_of_responsibility;)>
<!ELEMENT statement_of_responsibility (statresp, (equals_sign, pstatresp)*)>

<!ELEMENT statresp (#PCDATA|statresp.osource)*>
<!ATTLIST statresp caption CDATA #FIXED "szerzőségi adat">
<!ELEMENT statresp.osource (o_squarebracket, statresp.os, c_squarebracket)>
<!ELEMENT statresp.os (#PCDATA)>
<!ATTLIST statresp.os caption CDATA #FIXED "szerzőségi adat megengedett
forrásból">

<!ELEMENT pstatresp (#PCDATA|pstatresp.osource)*>
<!ATTLIST pstatresp caption CDATA #FIXED "párhuzamos szerzőségi adat">
<!ELEMENT pstatresp.osource
(o_squarebracket, pstatresp.os, c_squarebracket)>
<!ELEMENT pstatresp.os (#PCDATA)>
<!ATTLIST pstatresp.os caption CDATA #FIXED "párhuzamos szerzőségi adat
megengedett forrásból">

<!ELEMENT area2 (pointspacedashspace, edition_statement,
%opt_statements_of_responsibility;)>
<!ELEMENT edition_statement (#PCDATA|edition_statement.osource)*>
<!ATTLIST edition_statement caption CDATA #FIXED "kiadásjelzés">
<!ELEMENT edition_statement.osource
(o_squarebracket, edition_statement.os, c_squarebracket)>
<!ELEMENT edition_statement.os (#PCDATA)>
<!ATTLIST edition_statement.os caption CDATA #FIXED "kiadásjelzés
megengedett forrásból">

<!-- Definition of Punctuation Symbols -->

<!ELEMENT c_curvedbracket EMPTY>
<!ELEMENT char_position3. EMPTY>
<!ELEMENT colon EMPTY>
<!ELEMENT comma EMPTY>
<!ELEMENT c_squarebracket EMPTY>
<!ELEMENT diagonal_slash EMPTY>
<!ELEMENT newindent EMPTY>
<!ELEMENT o_curvedbracket EMPTY>
<!ELEMENT o_squarebracket EMPTY>
<!ELEMENT plus EMPTY>
<!ELEMENT point EMPTY>
<!ELEMENT pointspacedashspace EMPTY>
<!ELEMENT equals_sign EMPTY>
<!ELEMENT semicolon EMPTY>
```

### 3. melléklet: BDML\_DTD\_AREA3.dtd

```
<?xml version="1.0" encoding="UTF-8"?>

<!ENTITY % firstpublisher "place_of_publ, (semicolon, place_of_publ)?,
(semicolon, place_of_publ)?, (colon, name_of_publ)">
<!ENTITY % otherpublishers "((semicolon, place_of_publ)?, (semicolon,
place_of_publ)?, (semicolon, place_of_publ)?, (colon, name_of_publ))*">
<!ENTITY % firstprinter "place_of_print, (semicolon, place_of_print)?,
(semicolon, place_of_print)?, (colon, name_of_print)">
<!ENTITY % otherprinters "((semicolon, place_of_print)?, (semicolon,
place_of_print)?, (semicolon, place_of_print)?, (colon, name_of_print))*">

<!ENTITY % publisher "%firstpublisher;, %otherpublishers;">
<!ENTITY % dates "comma, (date_of_publ | notGregorian_cal), (comma,
date_of_cop)?, (comma, date_of_print)?">
<!ENTITY % printer "(o_curvedbracket, %firstprinter;, %otherprinters;,
c_curvedbracket)?">

<!-- Definition of Publication, Distribution Area -->

<!ELEMENT area3 (pointspacedashspace, %publisher;, %dates;, %printer;)>

<!ELEMENT place_of_publ (#PCDATA|place_of_publ.osource)*>
<!ATTLIST place_of_publ retrieve (y|n) "n"
caption CDATA #FIXED "megjelenés helye">
<!ELEMENT place_of_publ.osource
(o_squarebracket, place_of_publ.os, c_squarebracket)>
<!ELEMENT place_of_publ.os (#PCDATA)>
<!ATTLIST place_of_publ.os retrieve (y|n) "n"
caption CDATA #FIXED "megjelenés helye megengedett
forrásból">

<!ELEMENT name_of_publ (#PCDATA|name_of_publ.osource)*>
<!ATTLIST name_of_publ retrieve (y|n) "n"
caption CDATA #FIXED "kiadó neve">
<!ELEMENT name_of_publ.osource
(o_squarebracket, name_of_publ.os, c_squarebracket)>
<!ELEMENT name_of_publ.os (#PCDATA)>
<!ATTLIST name_of_publ.os retrieve (y|n) "n"
caption CDATA #FIXED "kiadó neve megengedett forrásból">

<!ELEMENT date_of_publ (#PCDATA|date_of_publ.osource)*>
<!ATTLIST date_of_publ caption CDATA #FIXED "megjelenés éve">
<!ELEMENT date_of_publ.osource
(o_squarebracket, date_of_publ.os, c_squarebracket)>
<!ELEMENT date_of_publ.os (#PCDATA)>
<!ATTLIST date_of_publ.os caption CDATA #FIXED "megjelenés éve megengedett
forrásból">

<!ELEMENT notGregorian_cal
(notGregorian_date, comma, o_squarebracket, date, c_squarebracket)>

<!ELEMENT notGregorian_date (#PCDATA)>
<!ATTLIST notGregorian_date caption CDATA #FIXED "megjelenés éve eltérő
időszámításban">
<!ELEMENT date (#PCDATA)>
<!ATTLIST date caption CDATA #FIXED "megjelenés éve időszámításunk szerint
megengedett forrásból">
```

```

<!ELEMENT date_of_cop (#PCDATA|date_of_cop.osource)*>
<!ATTLIST date_of_cop caption CDATA #FIXED "copyright éve">
<!ELEMENT date_of_cop.osource
(o_squarebracket,date_of_cop.os,c_squarebracket)>
<!ELEMENT date_of_cop.os (#PCDATA)>
<!ATTLIST date_of_cop.os caption CDATA #FIXED "copyright éve megengedett
forrásból">

<!ELEMENT date_of_print (#PCDATA|date_of_print.osource)*>
<!ATTLIST date_of_print caption CDATA #FIXED "nyomtatás éve">
<!ELEMENT date_of_print.osource
(o_squarebracket,date_of_print.os,c_squarebracket)>
<!ELEMENT date_of_print.os (#PCDATA)>
<!ATTLIST date_of_print.os caption CDATA #FIXED "nyomtatás éve megengedett
forrásból">

<!ELEMENT place_of_print (#PCDATA|place_of_print.osource)*>
<!ATTLIST place_of_print retrieve (y|n) "n"
caption CDATA #FIXED "nyomda székhelye">
<!ELEMENT place_of_print.osource
(o_squarebracket,place_of_print.os,c_squarebracket)>
<!ELEMENT place_of_print.os (#PCDATA)>
<!ATTLIST place_of_print.os retrieve (y|n) "n"
caption CDATA #FIXED "nyomda székhelye megengedett
forrásból">

<!ELEMENT name_of_print (#PCDATA|name_of_print.osource)*>
<!ATTLIST name_of_print retrieve (y|n) "n"
caption CDATA #FIXED "nyomda neve">
<!ELEMENT name_of_print.osource
(o_squarebracket,name_of_print.os,c_squarebracket)>
<!ELEMENT name_of_print.os (#PCDATA)>
<!ATTLIST name_of_print.os retrieve (y|n) "n"
caption CDATA #FIXED "nyomda neve megengedett
forrásból">

<!-- Definition of Punctuation Symbols -->

<!ELEMENT c_curvedbracket EMPTY>
<!ELEMENT char_position3. EMPTY>
<!ELEMENT colon EMPTY>
<!ELEMENT comma EMPTY>
<!ELEMENT c_squarebracket EMPTY>
<!ELEMENT diagonal_slash EMPTY>
<!ELEMENT newindent EMPTY>
<!ELEMENT o_curvedbracket EMPTY>
<!ELEMENT o_squarebracket EMPTY>
<!ELEMENT plus EMPTY>
<!ELEMENT point EMPTY>
<!ELEMENT pointspacedashspace EMPTY>
<!ELEMENT equals_sign EMPTY>
<!ELEMENT semicolon EMPTY>

```

#### 4. melléklet: BDML\_DTD\_AREA4.dtd

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- Definition of Physical Description Area -->

<!ELEMENT area4
(pointspacedashspace,extent,(colon,illustration)?,semicolon,dimensions,(pl
us,accompanying_material)?)>

<!ELEMENT extent (#PCDATA)>
<!ATTLIST extent caption CDATA #FIXED "terjedelem">
<!ELEMENT illustration (#PCDATA)>
<!ATTLIST illustration caption CDATA #FIXED "illusztráció">
<!ELEMENT dimensions (#PCDATA)>
<!ATTLIST dimensions caption CDATA #FIXED "méret">

<!ELEMENT accompanying_material
(accompmat,(o_curvedbracket,extent,(colon,illustration)?,(semicolon,dimens
ions)?,c_curvedbracket)?)>

<!ELEMENT accompmat (#PCDATA)>
<!ATTLIST accompmat caption CDATA #FIXED "melléklet">

<!-- Definition of Punctuation Symbols -->

<!ELEMENT c_curvedbracket EMPTY>
<!ELEMENT char_position3. EMPTY>
<!ELEMENT colon EMPTY>
<!ELEMENT comma EMPTY>
<!ELEMENT c_squarebracket EMPTY>
<!ELEMENT diagonal_slash EMPTY>
<!ELEMENT newindent EMPTY>
<!ELEMENT o_curvedbracket EMPTY>
<!ELEMENT o_squarebracket EMPTY>
<!ELEMENT plus EMPTY>
<!ELEMENT point EMPTY>
<!ELEMENT pointspacedashspace EMPTY>
<!ELEMENT equals_sign EMPTY>
<!ELEMENT semicolon EMPTY>
```

## 5. melléklet: BDML\_DTD\_AREA5.dtd

```
<?xml version="1.0" encoding="UTF-8"?>

<!ENTITY % opt_statements_of_responsibility "(diagonal_slash,
statements_of_responsibility)?">
<!ENTITY % parallel_series_titles "(series_title, point, (subseries_design
| ((subseries_design, comma)?, subseries_title)), (equals_sign,
pseries_title, point, (psubseries_design | ((psubseries_design, comma)?,
psubseries_title)))+">

<!ENTITY % series_titles "((series_title, (equals_sign, pseries_title)*,
(colon, series_subtitle)*, (colon, series_othertitle)* |
%parallel_series_titles;)">
<!ENTITY % subseries_titles "(subseries_title, (equals_sign,
psubseries_title)*, (colon, subseries_subtitle)*, (colon,
subseries_othertitle)*)">
<!ENTITY % simple_series "(%series_titles;,
%opt_statements_of_responsibility;, (comma, ISSN)?, (semicolon,
series_numbering)?)">

<!ENTITY % mainseries "%simple_series;">

<!ENTITY % subseries "(((point, (subseries_design | %subseries_titles; |
(subseries_design, comma, %subseries_titles;)),
%opt_statements_of_responsibility;)*, (comma, ISSN)?, (semicolon,
subseries_numbering)?)*)">

<!ENTITY % totalparallel_complex_series "((%parallel_series_titles;,
(comma, ISSN)?, (semicolon, subseries_numbering)?)*)">

<!ENTITY % complex_series "(%mainseries;, (%subseries; |
%totalparallel_complex_series;))">

<!ENTITY % repeated_statements_of_responsibility "(semicolon,
statement_of_responsibility)*">

<!ELEMENT statements_of_responsibility (statement_of_responsibility,
%repeated_statements_of_responsibility;)>
<!ELEMENT statement_of_responsibility (statresp, (equals_sign, pstatresp)*)>

<!ELEMENT statresp (#PCDATA|statresp.osource)*>
<!ATTLIST statresp caption CDATA #FIXED "szerzőségi adat">
<!ELEMENT statresp.osource (o_squarebracket, statresp.os, c_squarebracket)>
<!ELEMENT statresp.os (#PCDATA)>
<!ATTLIST statresp.os caption CDATA #FIXED "szerzőségi adat megengedett
forrásból">

<!ELEMENT pstatresp (#PCDATA|pstatresp.osource)*>
<!ATTLIST pstatresp caption CDATA #FIXED "párhuzamos szerzőségi adat">
<!ELEMENT pstatresp.osource
(o_squarebracket, pstatresp.os, c_squarebracket)>
<!ELEMENT pstatresp.os (#PCDATA)>
<!ATTLIST pstatresp.os caption CDATA #FIXED "párhuzamos szerzőségi adat
megengedett forrásból">

<!-- Definition of Series Area -->
```

```

<!ELEMENT area5
(pointspacedashspace,o_curvedbracket,series,c_curvedbracket,(point,o_curve
dbracket,series,c_curvedbracket)*>
<!ELEMENT series (%simple_series; | %complex_series;)>

<!ELEMENT series_title (#PCDATA|series_title.osource)*>
<!ATTLIST series_title retrieve (y|n) "y"
        caption CDATA #FIXED "sorozati főcím">
<!ELEMENT series_title.osource
(o_squarebracket,series_title.os,c_squarebracket)>
<!ELEMENT series_title.os (#PCDATA)>
<!ATTLIST series_title.os retrieve (y|n) "y"
        caption CDATA #FIXED "sorozati főcím megengedett
forrásból">

<!ELEMENT pseries_title (#PCDATA|pseries_title.osource)*>
<!ATTLIST pseries_title retrieve (y|n) "y"
        caption CDATA #FIXED "párhuzamos sorozati főcím">
<!ELEMENT pseries_title.osource
(o_squarebracket,pseries_title.os,c_squarebracket)>
<!ELEMENT pseries_title.os (#PCDATA)>
<!ATTLIST pseries_title.os retrieve (y|n) "y"
        caption CDATA #FIXED "párhuzamos sorozati főcím
megengedett forrásból">

<!ELEMENT series_subtitle (#PCDATA|series_subtitle.osource)*>
<!ATTLIST series_subtitle retrieve (y|n) "y"
        caption CDATA #FIXED "sorozati alcím">
<!ELEMENT series_subtitle.osource
(o_squarebracket,series_subtitle.os,c_squarebracket)>
<!ELEMENT series_subtitle.os (#PCDATA)>
<!ATTLIST series_subtitle.os retrieve (y|n) "y"
        caption CDATA #FIXED "sorozati alcím megengedett
forrásból">

<!ELEMENT series_othertitle (#PCDATA|series_othertitle.osource)*>
<!ATTLIST series_othertitle retrieve (y|n) "n"
        caption CDATA #FIXED "sorozati egyéb címadat">
<!ELEMENT series_othertitle.osource
(o_squarebracket,series_othertitle.os,c_squarebracket)>
<!ELEMENT series_othertitle.os (#PCDATA)>
<!ATTLIST series_othertitle.os retrieve (y|n) "n"
        caption CDATA #FIXED "sorozati egyéb címadat
megengedett forrásból">

<!ELEMENT ISSN (#PCDATA|ISSN.osource)*>
<!ATTLIST ISSN caption CDATA #FIXED "ISSN">
<!ELEMENT ISSN.osource (o_squarebracket,ISSN.os,c_squarebracket)>
<!ELEMENT ISSN.os (#PCDATA)>
<!ATTLIST ISSN.os caption CDATA #FIXED "ISSN megengedett forrásból">

<!ELEMENT series_numbering (#PCDATA|series_numbering.osource)*>
<!ATTLIST series_numbering caption CDATA #FIXED "sorozati szám">
<!ELEMENT series_numbering.osource
(o_squarebracket,series_numbering.os,c_squarebracket)>
<!ELEMENT series_numbering.os (#PCDATA)>
<!ATTLIST series_numbering.os caption CDATA #FIXED "sorozati szám
megengedett forrásból">

<!ELEMENT subseries_design (#PCDATA|subseries_design.osource)*>
<!ATTLIST subseries_design caption CDATA #FIXED "alsorozati megjelölő">
<!ELEMENT subseries_design.osource
(o_squarebracket,subseries_design.os,c_squarebracket)>
<!ELEMENT subseries_design.os (#PCDATA)>

```

```

<!ATTLIST subseries_design.os caption CDATA #FIXED "alsorozati megjelölő
megengedett forrásból">

<!ELEMENT subseries_numbering (#PCDATA|subseries_numbering.osource)*>
<!ATTLIST subseries_numbering caption CDATA #FIXED "alsorozati szám">
<!ELEMENT subseries_numbering.osource
(o_squarebracket,subseries_numbering.os,c_squarebracket)>
<!ELEMENT subseries_numbering.os (#PCDATA)>
<!ATTLIST subseries_numbering.os caption CDATA #FIXED "alsorozati szám
megengedett forrásból">

<!ELEMENT subseries_title (#PCDATA|subseries_title.osource)*>
<!ATTLIST subseries_title retrieve (y|n) "y"
caption CDATA #FIXED "alsorozati cím">
<!ELEMENT subseries_title.osource
(o_squarebracket,subseries_title.os,c_squarebracket)>
<!ELEMENT subseries_title.os (#PCDATA)>
<!ATTLIST subseries_title.os retrieve (y|n) "y"
caption CDATA #FIXED "alsorozati cím megengedett
forrásból">

<!ELEMENT psubseries_title (#PCDATA|psubseries_title.osource)*>
<!ATTLIST psubseries_title retrieve (y|n) "y"
caption CDATA #FIXED "párhuzamos al sorozati cím">
<!ELEMENT psubseries_title.osource
(o_squarebracket,psubseries_title.os,c_squarebracket)>
<!ELEMENT psubseries_title.os (#PCDATA)>
<!ATTLIST psubseries_title.os retrieve (y|n) "y"
caption CDATA #FIXED "párhuzamos al sorozati cím
megengedett forrásból">

<!ELEMENT subseries_subtitle (#PCDATA|subseries_subtitle.osource)*>
<!ATTLIST subseries_subtitle retrieve (y|n) "y"
caption CDATA #FIXED "alsorozati alcím">
<!ELEMENT subseries_subtitle.osource
(o_squarebracket,subseries_subtitle.os,c_squarebracket)>
<!ELEMENT subseries_subtitle.os (#PCDATA)>
<!ATTLIST subseries_subtitle.os retrieve (y|n) "y"
caption CDATA #FIXED "alsorozati alcím megengedett
forrásból">

<!ELEMENT subseries_othertitle (#PCDATA|subseries_othertitle.osource)*>
<!ATTLIST subseries_othertitle retrieve (y|n) "n"
caption CDATA #FIXED "alsorozati egyéb címadat">
<!ELEMENT subseries_othertitle.osource
(o_squarebracket,subseries_othertitle.os,c_squarebracket)>
<!ELEMENT subseries_othertitle.os (#PCDATA)>
<!ATTLIST subseries_othertitle.os retrieve (y|n) "n"
caption CDATA #FIXED "alsorozati egyéb címadat
megengedett forrásból">

<!ELEMENT psubseries_design (#PCDATA|psubseries_design.osource)*>
<!ATTLIST psubseries_design caption CDATA #FIXED "párhuzamos al sorozati
megjelölő">
<!ELEMENT psubseries_design.osource
(o_squarebracket,psubseries_design.os,c_squarebracket)>
<!ELEMENT psubseries_design.os (#PCDATA)>
<!ATTLIST psubseries_design.os caption CDATA #FIXED "párhuzamos al sorozati
megjelölő megengedett forrásból">

<!-- Definition of Punctuation Symbols -->

<!ELEMENT c_curvedbracket EMPTY>
<!ELEMENT char_position3. EMPTY>

```

```
<!ELEMENT colon EMPTY>
<!ELEMENT comma EMPTY>
<!ELEMENT c_squarebracket EMPTY>
<!ELEMENT diagonal_slash EMPTY>
<!ELEMENT newindent EMPTY>
<!ELEMENT o_curvedbracket EMPTY>
<!ELEMENT o_squarebracket EMPTY>
<!ELEMENT plus EMPTY>
<!ELEMENT point EMPTY>
<!ELEMENT pointspacedashspace EMPTY>
<!ELEMENT equals_sign EMPTY>
<!ELEMENT semicolon EMPTY>
```

## 6. melléklet: BDML\_DTD\_AREA6.dtd

```
<?xml version="1.0" encoding="UTF-8"?>

<!-- Definition of Note Area -->

<!ELEMENT area6 ((pointspacedashspace, note, (pointspacedashspace, note)*
| ((newindent, char_position3.), note, ((newindent, char_position3.),
note)*))>

<!ELEMENT note (#PCDATA)>
<ATTLIST note retrieve (y|n) "n"
           type (general | variant_title | bibl_hist | orig_title |
orig_publ | contents | bibliogr)
           caption CDATA #FIXED "megjegyzés">

<!-- Definition of Punctuation Symbols -->

<!ELEMENT c_curvedbracket EMPTY>
<!ELEMENT char_position3. EMPTY>
<!ELEMENT colon EMPTY>
<!ELEMENT comma EMPTY>
<!ELEMENT c_squarebracket EMPTY>
<!ELEMENT diagonal_slash EMPTY>
<!ELEMENT newindent EMPTY>
<!ELEMENT o_curvedbracket EMPTY>
<!ELEMENT o_squarebracket EMPTY>
<!ELEMENT plus EMPTY>
<!ELEMENT point EMPTY>
<!ELEMENT pointspacedashspace EMPTY>
<!ELEMENT equals_sign EMPTY>
<!ELEMENT semicolon EMPTY>
```

## 7. melléklet: BDML\_DTD\_AREA7.dtd

```
<?xml version="1.0" encoding="UTF-8"?>

<!ENTITY % qualification "(o_curvedbracket, text, c_curvedbracket)?">

<!ENTITY % isISBN "(ISBN, %qualification;, (binding | (colon, price,
%qualification;) | (binding, colon, price, %qualification;))*">

<!ENTITY % notISBN "(binding | (price, %qualification;) | (binding, colon,
price, %qualification;))+>

<!-- Definition of Standard Number and Terms of Availability Area -->

<!ELEMENT area7
((pointspacedashspace,elements,(point,elements)*|((newindent,char_positio
n3.),elements,((newindent,char_position3.),elements)*))>

<!ELEMENT elements (%isISBN; | %notISBN;)>

<!ELEMENT text (#PCDATA)>
<!ATTLIST text caption CDATA #FIXED "magyarázat">

<!ELEMENT ISBN (#PCDATA)>
<!ATTLIST ISBN retrieve (y|n) "y"
caption CDATA #FIXED "ISBN">

<!ELEMENT binding (#PCDATA)>
<!ATTLIST binding caption CDATA #FIXED "kötés">
<!ELEMENT price (#PCDATA)>
<!ATTLIST price caption CDATA #FIXED "ár">

<!-- Definition of Punctuation Symbols -->

<!ELEMENT c_curvedbracket EMPTY>
<!ELEMENT char_position3. EMPTY>
<!ELEMENT colon EMPTY>
<!ELEMENT comma EMPTY>
<!ELEMENT c_squarebracket EMPTY>
<!ELEMENT diagonal_slash EMPTY>
<!ELEMENT newindent EMPTY>
<!ELEMENT o_curvedbracket EMPTY>
<!ELEMENT o_squarebracket EMPTY>
<!ELEMENT plus EMPTY>
<!ELEMENT point EMPTY>
<!ELEMENT pointspacedashspace EMPTY>
<!ELEMENT equals_sign EMPTY>
<!ELEMENT semicolon EMPTY>
```

## 8. melléklet: A teljes BDML DTD

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- XML BIBLIOGRAPHIC DESCRIPTION DTD
      Bibliographic Description Markup Language (BDML)
      AUTHOR: Medveczki M. (marianna@inf.unideb.hu)
      DOCUMENT TYPE: MONOGRAPHIC PUBLICATIONS (based on MSZ 3424/1)
      VERSION: 2.5 (2005.08.15.) -->

<!-- DECLARATIONS OF ENTITIES -->

<!--External entities -->

<!ENTITY % ISOnum PUBLIC "ISO 8879:1986//ENTITIES Numeric and Special
Characters//EN" "/ENT/ISOnum.ent">
<!ENTITY % ISolat1 PUBLIC "ISO 8879:1986//ENTITIES Added Latin 1//EN"
"/ENT/ISolat1.ent">
<!ENTITY % ISolat2 PUBLIC "ISO/IEC 8859/2//ENTITIES Added Latin 2//EN"
"/ENT/ISolat2.ent">
<!ENTITY % ISOgrk1 PUBLIC "ISO 8879:1986//ENTITIES Greek Letters//EN"
"/ENT/ISOgrk1.ent">
<!ENTITY % ISOpub PUBLIC "ISO 8879:1986//ENTITIES Publishing//EN"
"/EGYE/ISOpub.ent">
<!ENTITY % ISodia PUBLIC "ISO 8879:1986//ENTITIES Diacritical Markss//EN"
"/ENT/ISodia.ent">

%ISOnum;
%ISolat1;
%ISolat2;
%ISOgrk1;
%ISOpub;
%ISodia;

<!-- Internal entities -->

<!ENTITY % repeated_first_area "(point, first_area)*, (colon, subtitle)*,
(colon, othertitle)*">

<!-- Parallel titles and parallel statements of responsibility -->
<!-- When a title page bears one or more parallel titles and/or parallel
statements of other title information and parallel statements of
responsibility, but one statement of responsibility is in only one
language and/or script, this statement of responsibility, if given, is
given after the last parallel statements of responsibility transcribed. --
>

<!ENTITY % parallel_farea "(equals_sign, farea)*, (semicolon, statresp)*">

<!ENTITY % repeated_titles "(semicolon, titles)*, (colon, subtitle)*,
(colon, othertitle)*">

<!ENTITY % opt_statements_of_responsibility "(diagonal_slash,
statements_of_responsibility)?">
```

```

<!ENTITY % pothertitles "((colon, othertitle), (equals_sign,
pothertitle)*)*">

<!ENTITY % psubtitles "((colon, subtitle), (equals_sign, psubtitle)*)*">

<!ENTITY % ptitlesubtitle "(equals_sign, ptitleproper, (colon,
psubtitle)+)">

<!ENTITY % titlesubtitleothertitle "titleproper, (colon, subtitle)+,
(colon, othertitle)+">

<!ENTITY % ptitlesubtitleothertitle "(((equals_sign, ptitleproper, (colon,
psubtitle)+, (colon, pothertitle)+) | ((ptitleproper, (colon,
psubtitle)*, (colon, pothertitle)*)+))">

<!ENTITY % titles_var1 "(titleproper, (equals_sign, ptitleproper)*,
(colon, subtitle)*, (colon, othertitle)*)">

<!ENTITY % firstpublisher "place_of_publ, (semicolon, place_of_publ)?,
(semicolon, place_of_publ)?, (colon, name_of_publ)">
<!ENTITY % otherpublishers "((semicolon, place_of_publ)?, (semicolon,
place_of_publ)?, (semicolon, place_of_publ)?, (colon, name_of_publ))*">
<!ENTITY % firstprinter "place_of_print, (semicolon, place_of_print)?,
(semicolon, place_of_print)?, (colon, name_of_print)">
<!ENTITY % otherprinters "((semicolon, place_of_print)?, (semicolon,
place_of_print)?, (semicolon, place_of_print)?, (colon, name_of_print))*">

<!ENTITY % parallel_series_titles "(series_title, point, (subseries_design
| ((subseries_design, comma)?, subseries_title)), (equals_sign,
pseries_title, point, (psubseries_design | ((psubseries_design, comma)?,
psubseries_title)))">

<!ENTITY % series_titles "((series_title, (equals_sign, pseries_title)*,
(colon, series_subtitle)*, (colon, series_othertitle)* |
%parallel_series_titles;)">

<!ENTITY % subseries_titles "(subseries_title, (equals_sign,
psubseries_title)*, (colon, subseries_subtitle)*, (colon,
subseries_othertitle)*)">

<!ENTITY % simple_series "(%series_titles;,
%opt_statements_of_responsibility;, (comma, ISSN)?, (semicolon,
series_numbering)?)">

<!ENTITY % mainseries "%simple_series;">

<!ENTITY % subseries "(((point, (subseries_design | %subseries_titles; |
(subseries_design, comma, %subseries_titles;)),
%opt_statements_of_responsibility;)*, (comma, ISSN)?, (semicolon,
subseries_numbering)?)*">

<!ENTITY % totalparallel_complex_series "((%parallel_series_titles;,
(comma, ISSN)?, (semicolon, subseries_numbering)?)*">

<!ENTITY % complex_series "(%mainseries;, (%subseries; |
%totalparallel_complex_series;))">

<!ENTITY % qualification "(o_curvedbracket, text, c_curvedbracket)?">

<!ENTITY % titles_var2 "(titleproper, (equals_sign, ptitleproper)*,
(colon, subtitle)+, %pothertitles;)">

```

```

<!ENTITY % titles_var3 "(titleproper, %psubtitles;, (colon,
othertitle)*)">

<!-- When title page bears one or more parallel titles and other title
information in more than one language and/or script, each statement of
other title information, if given, follows the title to which it is
linguistically related. -->
<!ENTITY % titles_var4 "(titleproper, (colon, subtitle)+,
%ptitlesubtitle;, (colon, othertitle)*)">

<!ENTITY % titles_var5 "(%titlesubtitleothertitle;,
%ptitlesubtitleothertitle;)">

<!ENTITY % titles_var6 "(%titlesubtitleothertitle;)">

<!ENTITY % repeated_statements_of_responsibility "(semicolon,
statement_of_responsibility)*">

<!ENTITY % publisher "%firstpublisher;, %otherpublishers;">
<!ENTITY % dates "comma, (date_of_publ | notGregorian_cal), (comma,
date_of_cop)?, (comma, date_of_print)?">
<!ENTITY % printer "(o_curvedbracket, %firstprinter;, %otherprinters;,
c_curvedbracket)?">

<!ENTITY % isISBN "(ISBN, %qualification;, (binding | (colon, price,
%qualification;) | (binding, colon, price, %qualification;))*">

<!ENTITY % notISBN "(binding | (price, %qualification;) | (binding, colon,
price, %qualification;))+>

<!-- MAIN STRUCTURE -->

<!-- Bibliographic description is document element -->
<!ELEMENT mp_bibl_desc (single_level|multi_level|component_parts)>
<!ATTLIST mp_bibl_desc
    id ID #REQUIRED
    autopsy (y|n) "y">

<!ELEMENT single_level (single|single_volume)>

<!ELEMENT single (area1,(area2)?,area3,area4,(area5)?,area6,area7)>

<!-- Definition of Title and Statement of Responsibility Area -->

<!-- Publication can contain individual works by different authors, unless
a linking word or phrase is given in the publication -->
<!ELEMENT area1
(newindent, char_position3.,first_area,%repeated_first_area;)>

<!-- Parallel titles and parallel statements of responsibility -->
<!ELEMENT first_area (farea,%parallel_farea;)>

<!-- Publication can contain individual works by the same author, unless a
linking word or phrase is given in the publication -->

```

```

<!ELEMENT farea
(titles,%repeated_titles;,%opt_statements_of_responsibility;)>

<!ELEMENT titles (%titles_var1; | %titles_var2; | %titles_var3; |
%titles_var4; | %titles_var5; | %titles_var6;)>

<!ELEMENT titleproper (title,(alternative_title)*)>
<!ELEMENT ptitleproper (ptitle,(palternative_title)*)>

<!ELEMENT title (#PCDATA)>
<!ATTLIST title retrieve (y|n) "y"
caption CDATA #FIXED "főcím" >
<!ELEMENT ptitle (#PCDATA)>
<!ATTLIST ptitle retrieve (y|n) "y"
caption CDATA #FIXED "párhuzamos főcím">
<!ELEMENT alternative_title (#PCDATA)>
<!ATTLIST alternative_title retrieve (y|n) "y"
caption CDATA #FIXED "vagylagos cím">
<!ELEMENT palternative_title (#PCDATA)>
<!ATTLIST palternative_title retrieve (y|n) "y"
caption CDATA #FIXED "párhuzamos vagylagos cím">
<!ELEMENT subtitle (#PCDATA)>
<!ATTLIST subtitle retrieve (y|n) "n"
caption CDATA #FIXED "alcím">
<!ELEMENT psubtitle (#PCDATA)>
<!ATTLIST psubtitle retrieve (y|n) "n"
caption CDATA #FIXED "párhuzamos alcím">

<!ELEMENT othertitle (#PCDATA|othertitle.osource)*>
<!ATTLIST othertitle retrieve (y|n) "n"
caption CDATA #FIXED "egyéb címadat">
<!ELEMENT othertitle.osource
(o_squarebracket,othertitle.os,c_squarebracket)>
<!ELEMENT othertitle.os (#PCDATA)>
<!ATTLIST othertitle.os retrieve (y|n) "n"
caption CDATA #FIXED "egyéb címadat megengedett
forrásból">

<!ELEMENT pothertitle (#PCDATA|pothertitle.osource)*>
<!ATTLIST pothertitle retrieve (y|n) "n"
caption CDATA #FIXED "párhuzamos egyéb címadat">
<!ELEMENT pothertitle.osource
(o_squarebracket,pothertitle.os,c_squarebracket)>
<!ELEMENT pothertitle.os (#PCDATA)>
<!ATTLIST pothertitle.os retrieve (y|n) "n"
caption CDATA #FIXED "párhuzamos egyéb címadat
megengedett forrásból">

<!-- Definition of Statements of Responsibility -->

<!ELEMENT statements_of_responsibility (statement_of_responsibility,
%repeated_statements_of_responsibility;)>
<!ELEMENT statement_of_responsibility (statresp,(equals_sign,pstatresp)*)>

<!ELEMENT statresp (#PCDATA|statresp.osource)*>
<!ATTLIST statresp caption CDATA #FIXED "szerzőségi adat">
<!ELEMENT statresp.osource (o_squarebracket,statresp.os,c_squarebracket)>
<!ELEMENT statresp.os (#PCDATA)>
<!ATTLIST statresp.os caption CDATA #FIXED "szerzőségi adat megengedett
forrásból">

<!ELEMENT pstatresp (#PCDATA|pstatresp.osource)*>
<!ATTLIST pstatresp caption CDATA #FIXED "párhuzamos szerzőségi adat">

```

```

<!ELEMENT pstatresp.osource
(o_squarebracket,pstatresp.os,c_squarebracket)>
<!ELEMENT pstatresp.os (#PCDATA)>
<!ATTLIST pstatresp.os caption CDATA #FIXED "párhuzamos szerzőségi adat
megengedett forrásból">

<!-- Definition of Edition Area -->

<!ELEMENT area2 (pointspacedashspace,edition_statement,
%opt_statements_of_responsibility;)>
<!ELEMENT edition_statement (#PCDATA|edition_statement.osource)*>
<!ATTLIST edition_statement caption CDATA #FIXED "kiadásjelzés">
<!ELEMENT edition_statement.osource
(o_squarebracket,edition_statement.os,c_squarebracket)>
<!ELEMENT edition_statement.os (#PCDATA)>
<!ATTLIST edition_statement.os caption CDATA #FIXED "kiadásjelzés
megengedett forrásból">

<!-- Definition of Publication, Distribution Area -->

<!ELEMENT area3 (pointspacedashspace,%publisher;, %dates;, %printer;)>

<!ELEMENT place_of_publ (#PCDATA|place_of_publ.osource)*>
<!ATTLIST place_of_publ retrieve (y|n) "n"
caption CDATA #FIXED "megjelenés helye">
<!ELEMENT place_of_publ.osource
(o_squarebracket,place_of_publ.os,c_squarebracket)>
<!ELEMENT place_of_publ.os (#PCDATA)>
<!ATTLIST place_of_publ.os retrieve (y|n) "n"
caption CDATA #FIXED "megjelenés helye megengedett
forrásból">

<!ELEMENT name_of_publ (#PCDATA|name_of_publ.osource)*>
<!ATTLIST name_of_publ retrieve (y|n) "n"
caption CDATA #FIXED "kiadó neve">
<!ELEMENT name_of_publ.osource
(o_squarebracket,name_of_publ.os,c_squarebracket)>
<!ELEMENT name_of_publ.os (#PCDATA)>
<!ATTLIST name_of_publ.os retrieve (y|n) "n"
caption CDATA #FIXED "kiadó neve megengedett forrásból">

<!ELEMENT date_of_publ (#PCDATA|date_of_publ.osource)*>
<!ATTLIST date_of_publ caption CDATA #FIXED "megjelenés éve">
<!ELEMENT date_of_publ.osource
(o_squarebracket,date_of_publ.os,c_squarebracket)>
<!ELEMENT date_of_publ.os (#PCDATA)>
<!ATTLIST date_of_publ.os caption CDATA #FIXED "megjelenés éve megengedett
forrásból">

<!ELEMENT notGregorian_cal
(notGregorian_date,comma,o_squarebracket,date,c_squarebracket)>

<!ELEMENT notGregorian_date (#PCDATA)>
<!ATTLIST notGregorian_date caption CDATA #FIXED "megjelenés éve eltérő
időszámításban">
<!ELEMENT date (#PCDATA)>
<!ATTLIST date caption CDATA #FIXED "megjelenés éve időszámításunk szerint
megengedett forrásból">

<!ELEMENT date_of_cop (#PCDATA|date_of_cop.osource)*>
<!ATTLIST date_of_cop caption CDATA #FIXED "copyright éve">

```

```

<!ELEMENT date_of_cop.osource
(o_squarebracket,date_of_cop.os,c_squarebracket)>
<!ELEMENT date_of_cop.os (#PCDATA)>
<!ATTLIST date_of_cop.os caption CDATA #FIXED "copyright éve megengedett
forrásból">

<!ELEMENT date_of_print (#PCDATA|date_of_print.osource)*>
<!ATTLIST date_of_print caption CDATA #FIXED "nyomtatás éve">
<!ELEMENT date_of_print.osource
(o_squarebracket,date_of_print.os,c_squarebracket)>
<!ELEMENT date_of_print.os (#PCDATA)>
<!ATTLIST date_of_print.os caption CDATA #FIXED "nyomtatás éve megengedett
forrásból">

<!ELEMENT place_of_print (#PCDATA|place_of_print.osource)*>
<!ATTLIST place_of_print retrieve (y|n) "n"
caption CDATA #FIXED "nyomda székhelye">
<!ELEMENT place_of_print.osource
(o_squarebracket,place_of_print.os,c_squarebracket)>
<!ELEMENT place_of_print.os (#PCDATA)>
<!ATTLIST place_of_print.os retrieve (y|n) "n"
caption CDATA #FIXED "nyomda székhelye megengedett
forrásból">

<!ELEMENT name_of_print (#PCDATA|name_of_print.osource)*>
<!ATTLIST name_of_print retrieve (y|n) "n"
caption CDATA #FIXED "nyomda neve">
<!ELEMENT name_of_print.osource
(o_squarebracket,name_of_print.os,c_squarebracket)>
<!ELEMENT name_of_print.os (#PCDATA)>
<!ATTLIST name_of_print.os retrieve (y|n) "n"
caption CDATA #FIXED "nyomda neve megengedett
forrásból">

<!-- Definition of Physical Description Area -->

<!ELEMENT area4
(pointspacedashspace,extent,(colon,illustration)?,semicolon,dimensions,(pl
us,accompanying_material)?)>

<!ELEMENT extent (#PCDATA)>
<!ATTLIST extent caption CDATA #FIXED "terjedelem">
<!ELEMENT illustration (#PCDATA)>
<!ATTLIST illustration caption CDATA #FIXED "illusztráció">
<!ELEMENT dimensions (#PCDATA)>
<!ATTLIST dimensions caption CDATA #FIXED "méret">

<!ELEMENT accompanying_material
(accmpmat,(o_curvedbracket,extent,(colon,illustration)?,(semicolon,dimens
ions)?,c_curvedbracket)?)>

<!ELEMENT accmpmat (#PCDATA)>
<!ATTLIST accmpmat caption CDATA #FIXED "melléklet">

<!-- Definition of Series Area -->

<!ELEMENT area5
(pointspacedashspace,o_curvedbracket,series,c_curvedbracket,(point,o_curve
dbracket,series,c_curvedbracket)*)>
<!ELEMENT series (%simple_series; | %complex_series;)>

<!ELEMENT series_title (#PCDATA|series_title.osource)*>
<!ATTLIST series_title retrieve (y|n) "y"

```

```

        caption CDATA #FIXED "sorozati főcím">
<!ELEMENT series_title.osource
(o_squarebracket,series_title.os,c_squarebracket)>
<!ELEMENT series_title.os (#PCDATA)>
<!ATTLIST series_title.os retrieve (y|n) "y"
        caption CDATA #FIXED "sorozati főcím megengedett
forrásból">

<!ELEMENT pseries_title (#PCDATA|pseries_title.osource)*>
<!ATTLIST pseries_title retrieve (y|n) "y"
        caption CDATA #FIXED "párhuzamos sorozati főcím">
<!ELEMENT pseries_title.osource
(o_squarebracket,pseries_title.os,c_squarebracket)>
<!ELEMENT pseries_title.os (#PCDATA)>
<!ATTLIST pseries_title.os retrieve (y|n) "y"
        caption CDATA #FIXED "párhuzamos sorozati főcím
megengedett forrásból">

<!ELEMENT series_subtitle (#PCDATA|series_subtitle.osource)*>
<!ATTLIST series_subtitle retrieve (y|n) "y"
        caption CDATA #FIXED "sorozati alcím">
<!ELEMENT series_subtitle.osource
(o_squarebracket,series_subtitle.os,c_squarebracket)>
<!ELEMENT series_subtitle.os (#PCDATA)>
<!ATTLIST series_subtitle.os retrieve (y|n) "y"
        caption CDATA #FIXED "sorozati alcím megengedett
forrásból">

<!ELEMENT series_othertitle (#PCDATA|series_othertitle.osource)*>
<!ATTLIST series_othertitle retrieve (y|n) "n"
        caption CDATA #FIXED "sorozati egyéb címadat">
<!ELEMENT series_othertitle.osource
(o_squarebracket,series_othertitle.os,c_squarebracket)>
<!ELEMENT series_othertitle.os (#PCDATA)>
<!ATTLIST series_othertitle.os retrieve (y|n) "n"
        caption CDATA #FIXED "sorozati egyéb címadat
megengedett forrásból">

<!ELEMENT ISSN (#PCDATA|ISSN.osource)*>
<!ATTLIST ISSN caption CDATA #FIXED "ISSN">
<!ELEMENT ISSN.osource (o_squarebracket,ISSN.os,c_squarebracket)>
<!ELEMENT ISSN.os (#PCDATA)>
<!ATTLIST ISSN.os caption CDATA #FIXED "ISSN megengedett forrásból">

<!ELEMENT series_numbering (#PCDATA|series_numbering.osource)*>
<!ATTLIST series_numbering caption CDATA #FIXED "sorozati szám">
<!ELEMENT series_numbering.osource
(o_squarebracket,series_numbering.os,c_squarebracket)>
<!ELEMENT series_numbering.os (#PCDATA)>
<!ATTLIST series_numbering.os caption CDATA #FIXED "sorozati szám
megengedett forrásból">

<!ELEMENT subseries_design (#PCDATA|subseries_design.osource)*>
<!ATTLIST subseries_design caption CDATA #FIXED "alsorozati megjelölő">
<!ELEMENT subseries_design.osource
(o_squarebracket,subseries_design.os,c_squarebracket)>
<!ELEMENT subseries_design.os (#PCDATA)>
<!ATTLIST subseries_design.os caption CDATA #FIXED "alsorozati megjelölő
megengedett forrásból">

<!ELEMENT subseries_numbering (#PCDATA|subseries_numbering.osource)*>
<!ATTLIST subseries_numbering caption CDATA #FIXED "alsorozati szám">
<!ELEMENT subseries_numbering.osource
(o_squarebracket,subseries_numbering.os,c_squarebracket)>
<!ELEMENT subseries_numbering.os (#PCDATA)>

```

```

<!ATTLIST subseries_numbering.os caption CDATA #FIXED "alsorozati szám
megengedett forrásból">

<!ELEMENT subseries_title (#PCDATA|subseries_title.osource)*>
<!ATTLIST subseries_title retrieve (y|n) "y"
          caption CDATA #FIXED "alsorozati cím">
<!ELEMENT subseries_title.osource
(o_squarebracket,subseries_title.os,c_squarebracket)>
<!ELEMENT subseries_title.os (#PCDATA)>
<!ATTLIST subseries_title.os retrieve (y|n) "y"
          caption CDATA #FIXED "alsorozati cím megengedett
forrásból">

<!ELEMENT psubseries_title (#PCDATA|psubseries_title.osource)*>
<!ATTLIST psubseries_title retrieve (y|n) "y"
          caption CDATA #FIXED "párhuzamos al sorozati cím">
<!ELEMENT psubseries_title.osource
(o_squarebracket,psubseries_title.os,c_squarebracket)>
<!ELEMENT psubseries_title.os (#PCDATA)>
<!ATTLIST psubseries_title.os retrieve (y|n) "y"
          caption CDATA #FIXED "párhuzamos al sorozati cím
megengedett forrásból">

<!ELEMENT subseries_subtitle (#PCDATA|subseries_subtitle.osource)*>
<!ATTLIST subseries_subtitle retrieve (y|n) "y"
          caption CDATA #FIXED "alsorozati alcím">
<!ELEMENT subseries_subtitle.osource
(o_squarebracket,subseries_subtitle.os,c_squarebracket)>
<!ELEMENT subseries_subtitle.os (#PCDATA)>
<!ATTLIST subseries_subtitle.os retrieve (y|n) "y"
          caption CDATA #FIXED "alsorozati alcím megengedett
forrásból">

<!ELEMENT subseries_othertitle (#PCDATA|subseries_othertitle.osource)*>
<!ATTLIST subseries_othertitle retrieve (y|n) "n"
          caption CDATA #FIXED "alsorozati egyéb címadat">
<!ELEMENT subseries_othertitle.osource
(o_squarebracket,subseries_othertitle.os,c_squarebracket)>
<!ELEMENT subseries_othertitle.os (#PCDATA)>
<!ATTLIST subseries_othertitle.os retrieve (y|n) "n"
          caption CDATA #FIXED "alsorozati egyéb címadat
megengedett forrásból">

<!ELEMENT psubseries_design (#PCDATA|psubseries_design.osource)*>
<!ATTLIST psubseries_design caption CDATA #FIXED "párhuzamos al sorozati
megjelölő">
<!ELEMENT psubseries_design.osource
(o_squarebracket,psubseries_design.os,c_squarebracket)>
<!ELEMENT psubseries_design.os (#PCDATA)>
<!ATTLIST psubseries_design.os caption CDATA #FIXED "párhuzamos al sorozati
megjelölő megengedett forrásból">

<!-- Definition of Note Area -->

<!ELEMENT area6
((pointspacedashspace,note,(pointspacedashspace,note)*)|((newindent,char_p
osition3.),note,((newindent,char_position3.),note)*))>

<!ELEMENT note (#PCDATA)>
<!ATTLIST note retrieve (y|n) "n"
          type (general | variant_title | bibl_hist | orig_title |
orig_publ | contents | bibliogr)
          caption CDATA #FIXED "megjegyzés">

```

```

<!-- Definition of Standard Number and Terms of Availability Area -->

<!ELEMENT area7
((pointspacedashspace,elements,(point,elements)*|((newindent,char_positio
n3.),elements,((newindent,char_position3.),elements)*))>

<!ELEMENT elements (%isISBN; | %notISBN;)>

<!ELEMENT text (#PCDATA)>
<ATTLIST text caption CDATA #FIXED "magyarázat">

<!ELEMENT ISBN (#PCDATA)>
<ATTLIST ISBN retrieve (y|n) "y"
caption CDATA #FIXED "ISBN">

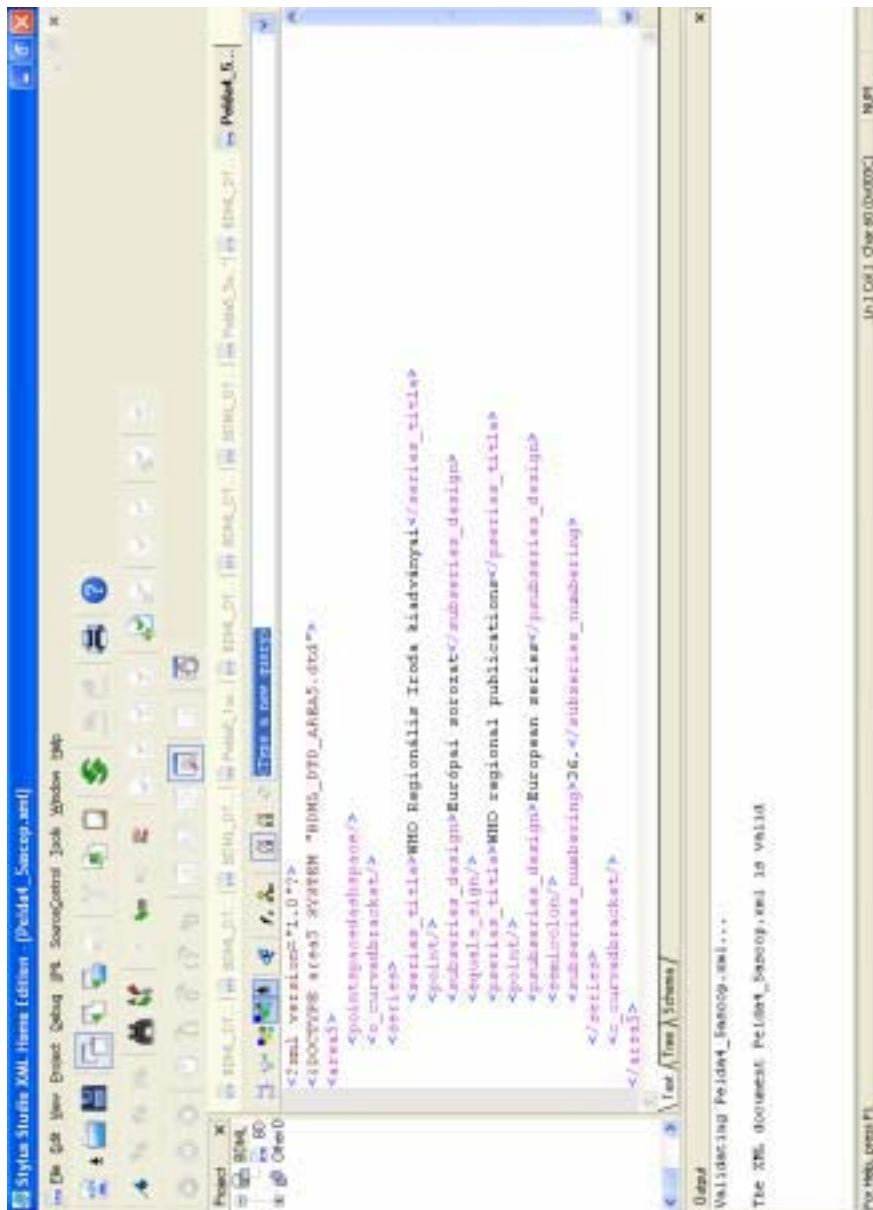
<!ELEMENT binding (#PCDATA)>
<ATTLIST binding caption CDATA #FIXED "kötés">
<!ELEMENT price (#PCDATA)>
<ATTLIST price caption CDATA #FIXED "ár">

<!-- Definition of Punctuation Symbols -->

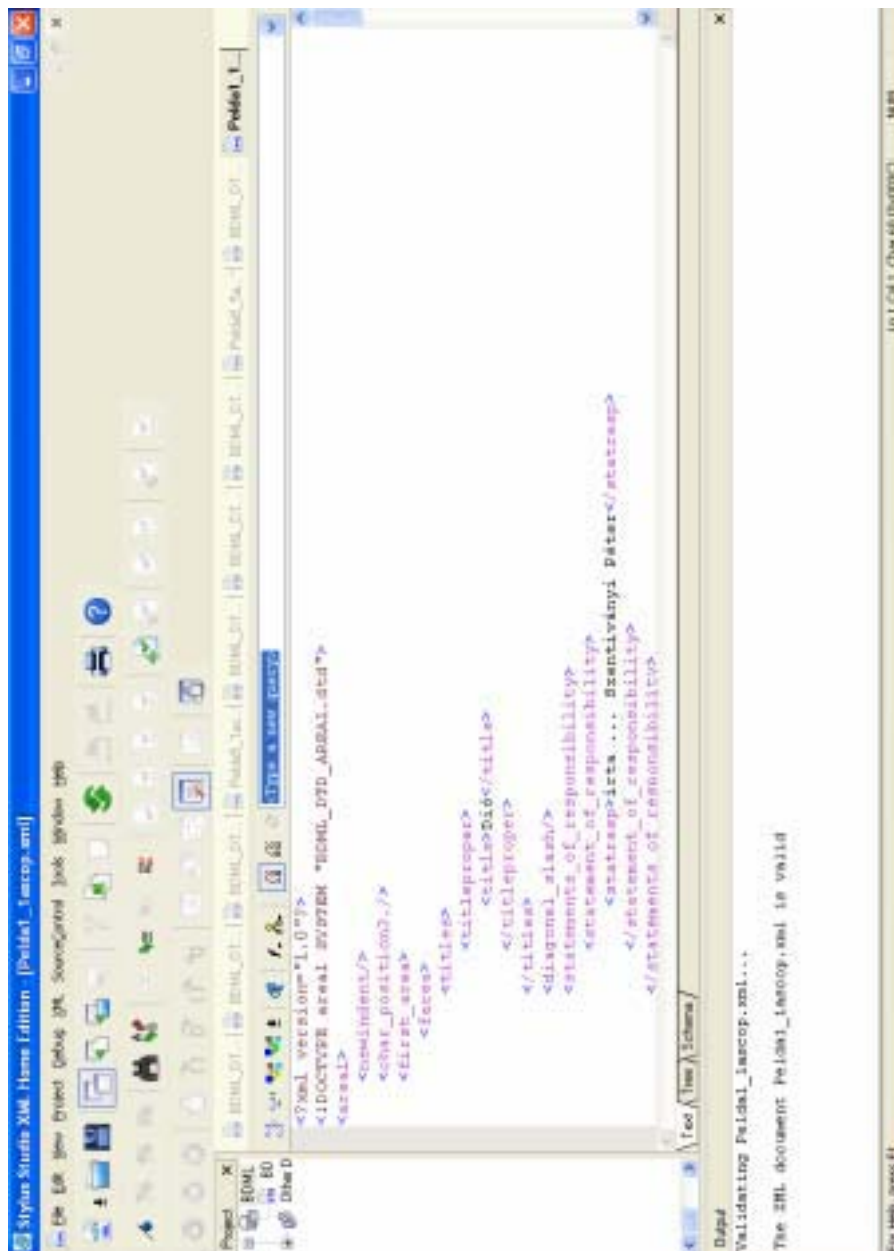
<!ELEMENT c_curvedbracket EMPTY>
<!ELEMENT char_position3. EMPTY>
<!ELEMENT colon EMPTY>
<!ELEMENT comma EMPTY>
<!ELEMENT c_squarebracket EMPTY>
<!ELEMENT diagonal_slash EMPTY>
<!ELEMENT newindent EMPTY>
<!ELEMENT o_curvedbracket EMPTY>
<!ELEMENT o_squarebracket EMPTY>
<!ELEMENT plus EMPTY>
<!ELEMENT point EMPTY>
<!ELEMENT pointspacedashspace EMPTY>
<!ELEMENT equals_sign EMPTY>
<!ELEMENT semicolon EMPTY>

```

## 9. melléklet: XML-dokumentum érvényességének ellenőrzése Stylus Studio 6 segítségével



## 10. melléklet: XML-dokumentum érvényességének ellenőrzése Stylus Studio 6 segítségével



## 11. melléklet: BDML DTD készítése Stylus Studio 6 –ban



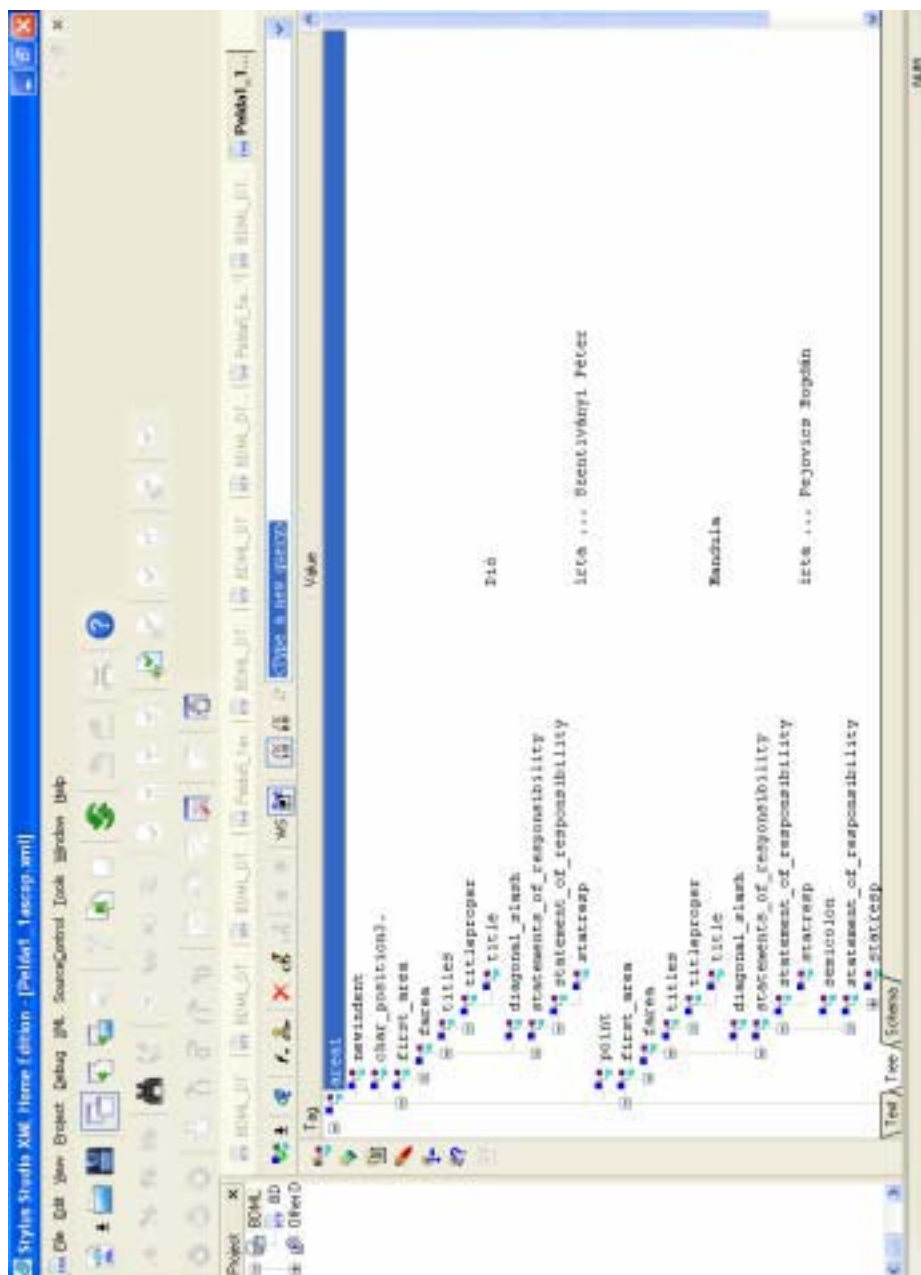
```
<?xml version="1.0" encoding="UTF-8"?>
<!-- XML EMBROIDERED DESCRIPTION DTD
Bibliographic Description Markup Language (BDML)
Author: Mervyn D. (mervyn@inf.unidb.ca)
Document Type: MONOGRAPH PUBLICATIONS (based on ISO 8879/1)
Version: 2.5 (2005.08.15.) -->

<!-- DECLARATIONS OF ENTITIES -->
<!-- External entities -->

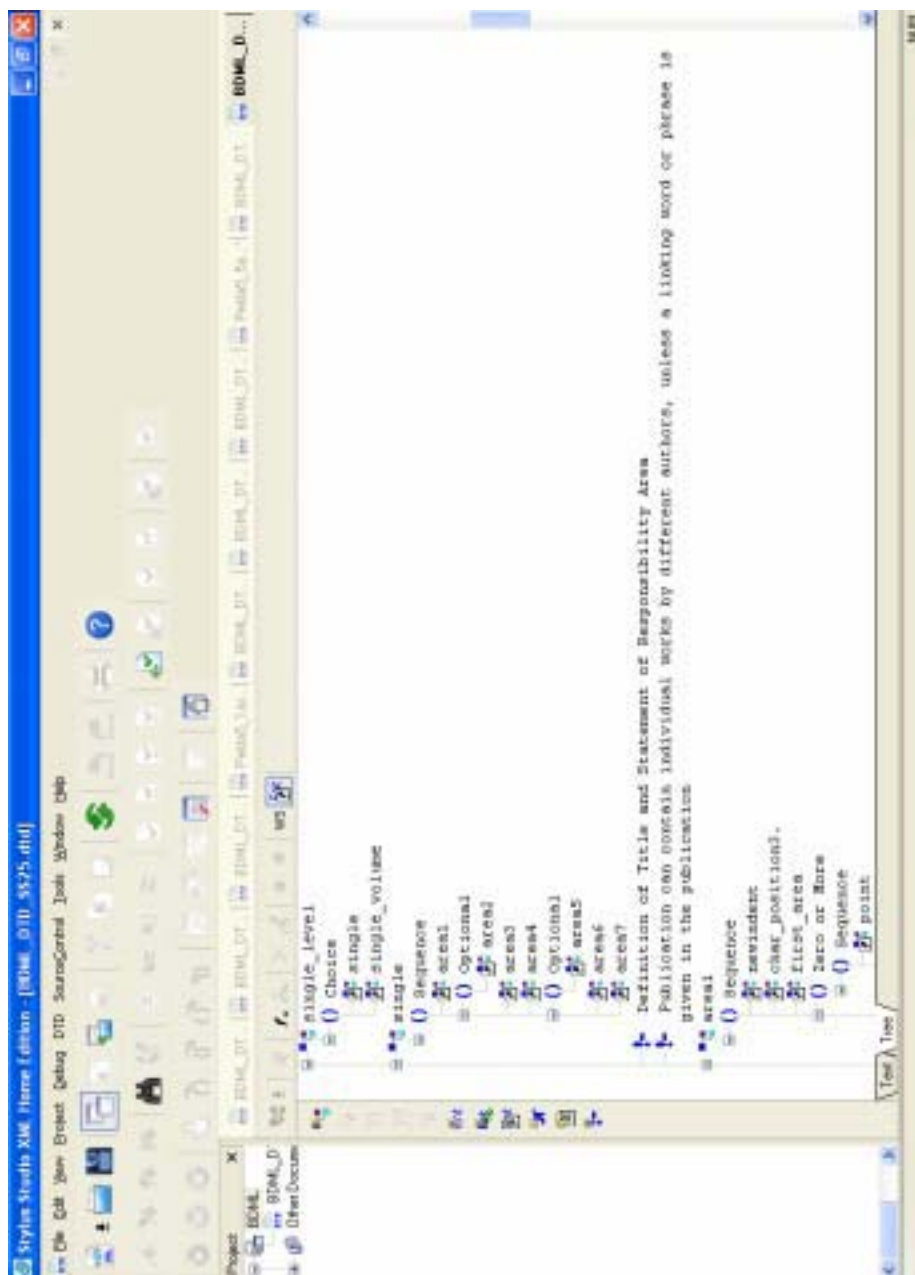
<ISO8879117 & ISO8879:1986//ENTITIES Numeric and Special Characters//EN" */ENT/ISO8879:1986.etc
<ISO8879118 & ISO8879:1986//ENTITIES Added Latin 1//EN" */ENT/ISOlat1.ent">
<ISO8879119 & ISO8879:1986//ENTITIES Added Latin 2//EN" */ENT/ISOlat2.ent">
<ISO8879120 & ISO8879:1986//ENTITIES Greek Letters//EN" */ENT/ISOgrki.ent">
<ISO8879121 & ISO8879:1986//ENTITIES Publishing//EN" */ENVE/ISOpub.ent">
<ISO8879122 & ISO8879:1986//ENTITIES Diacritical Marks//EN" */ENT/ISOdia.ent">

ISO8879:1986.etc
ISOlat1
ISOlat2
ISOgrki
ISOpub
ISOdia;
```

## 12. melléklet: XML-dokumentum hierarchikus megjelenítése Stylus Studio 6-ban



### 13. melléklet: BDML DTD-részlet grafikus megjelenítése Stylus Studio 6-ban





## MAGYAR NYELVŰ ÖSSZEFOGLALÁS

Ma a világon elterjedt és használt integrált könyvtári rendszerek többsége a MARC (Machine Readable Cataloguing) formátumot támogatja, amelynek kialakítása még jóval a web megjelenése előtt, az 1960-as években kezdődött meg a könyvtári szoftverek egy zártabb világa számára.

Szükségesnek látszik az, hogy olyan új bibliográfiai adtaformátumokról és ezek kialakítási lehetőségeiről gondolkodjunk, amelyek képesek alkalmazkodni az új körülményekhez, illetve képesek a webes technológiák nyílt szabványai és nyílt forráskódú szoftverei által nyújtott új lehetőségeket minél jobban kiaknázni.

Dolgozatom indításaként röviden áttekintem a számítógépes bibliográfiai adatformátumok kialakulásának és fejlődésének legfontosabb állomásait, összefoglalom a nemzetközi és hazai tendenciákat ezen a területen.

A dolgozat második részében egy új, saját fejlesztésű, XML-alapú jelölő nyelv bemutatása következik a számítógépes bibliográfiai adatfeldolgozás számára, amely a BDML (Bibliographic Description Markup Language) nevet kapta. Indításként felvázolom az XML alkalmazásának azon lehetőségeit a könyvtári dokumentumok formai feldolgozásában, amelyek - véleményem szerint - pozitív előrelépést jelenthetnek ezen a területen. Majd a következő fejezetekben a BDML tervezéséről, fejlesztési elveiről szólok. Ezt követően konkrét, bibliográfiai rekordokat tartalmazó példák segítségével kívánom bemutatni a BDML nyelvet leíró DTD (Document Type Definition) legfontosabbnak ítélt részeit. A tárgyalás során felhasznált bibliográfiai leírási példák kifejezetten úgy kerültek összeválogatásra, hogy előforduljanak ezekben azok a leírási nehézségek, amelyek kezelése a jelenleg, magyar könyvtárakban használt, számítógépes bibliográfiai adatformátumok számára problémát okoz. Ilyen módon a BDML struktúrájának bemutatása egyben egy összehasonlító vizsgálat is, ahol az egyes bibliográfiai leírási problémák kezelhetőségét megnézem a USMARC és a HUNMARC formátumokban (amely két formátum alkalmazása a hazai könyvtárakban a legelterjedtebb), illetve bemutatom ezekre a BDML nyújtotta lehetőségeket és pozitív válaszokat is.

Végezetül, a dolgozat zárásaként a BDML nyelv néhány továbbfejlesztési lehetőségét veszem sorra, amelyek a további kutatásaimat fogják meghatározni ebben a témakörben.

A BDML DTD kialakításánál szerzett tapasztalatok hasznosítására, és a további fejlesztésre két aktuálisan futó projektben is lehetőség nyílik. Az egyik az EMIR (Elektronikus Magyar Irodalom) projekt, amely az Eötvös Loránd Tudományegyetem Bölcsészeti Informatika Önálló Program, a Veszprémi Egyetem Magyar Irodalomtudományi Tanszékének és a Debreceni Egyetem könyvtár tanszékének együttműködése. A projekt célja a kortárs magyar irodalom komplex és sokoldalú online feldolgozása és a nagyközönség számára elérhetővé tétele, egyben egy nagyobb léptékű, összefüggő, modulárisan építkező, közös hozzáférésű, elosztott rendszerre épülő szolgáltatás létrehozása, amely egy rendszerben kapcsolná össze a digitálisan feldolgozott katalógusrendszereket, folyóiratokat, kölcsönzési és vásárlási lehetőségeket.

A másik kezdeményezés a Debreceni Egyetem Zeneműtárához és a Debreceni Egyetem könyvtár tanszékéhez kapcsolódik. A cél ebben az esetben egy digitális zeneműtár létrehozása, amely nemcsak a hangzó dokumentumok bibliográfiai adatait tárolná elektronikus formában, hanem magukat a zenei dokumentumokat is.

A dolgozat rövid összeggzéssel zárul.

## **ANGOL NYELVŰ ÖSSZEFOGLALÁS**

The majority of integrated library systems currently used in the library field support MARC (Machine Readable Cataloguing) format the development of which began before appearance of the world wide web in 1960s for a more closed scope of library software. It is important to deal with new bibliographical data formats and their development possibilities, which are capable to adapt to new conditions, and to best exploit new possibilities provided by open standards and open source softwares of web-based technologies.

In introductory part of my dissertation I gave a short presentation of creation and evolution of bibliographical data formats, and I summarized international and national tendencies in this field.

In the second part of my dissertation a new, own-developed, XML-based markup language for bibliographic data processing is included. This is the BDML (Bibliographic Description Markup Language). I discussed those application possibilities of XML in processing of library materials which can affect this field. In the next chapters I gave a summary of design and developing principles of BDML. The most important parts of BDML DTD were presented with the help of examples included in the exact bibliographical records. Applied examples were packed to contain those processing difficulties which induce problems to computer-related bibliographical data formats currently used in Hungarian libraries. The presentation of BDML structure is a comparative examination too. Handling of bibliographical processing problems in USMARC and in HUNMARC formats were compared (they are two best prevalent formats in Hungary) and possibilities and solutions of BDML to these problems were indicated.

At the end of dissertation some development possibilities of BDML were presented which determine my research in this field.

Experiences of development of BDML DTD are usable and structure of BDML DTD developable in two topical projects. One of them is an EMIR (Electronic Hungarian Literature)-project. This is cooperation of Human Informatics Separate Program of Eötvös Lóránd University and Department of Hungarian Literature of University of Veszprém and Department of Library and Information

Science of University of Debrecen. The goal of this project is to create complex and multi-faceted online processing of contemporary Hungarian literature, accessible for everyone and to develop modularized, distributed service linked online catalogue systems, electronic journals, lending and acquisition possibilities.

The other project is linked to the Music Library of University of Debrecen and to the Department of Library and Information Science of University of Debrecen. The goal is to develop a digital music library which stores in electronic form the bibliographical data of audio documents and audio documents too.

Dissertation ends with a short summary.

## PUBLIKÁCIÓS JEGYZÉK

### Hivatkozás:

A retrospektív konverzió problémái a Kossuth Lajos Tudományegyetem Könyvtárában == Könyv és könyvtár, 18. (1997.), p. 163-188.

**Hiv.:** Berke Barnabásné: A Magyar Nemzeti Bibliográfia retrospektív adatbázisa : gondolatok a retrospektív konverzió tervezéséhez és megvalósításához == Könyvtári Figyelő, Úf., 8.= 44. évf. 2. sz. (1998.), p. 229-238.

Az XML : új perspektívák a könyvtár-informatikában ==Tudományos és Műszaki Tájékoztatás, 51. évf. 2 (2004), p. 61-71.

**Rec.:** Bíró Szabolcs: XML bibliográfiai leírások készítésére – Debrecenből. Salgáné Medveczki Marianna: Az XML: új perspektívák a könyvtár-informatikában c. írásának referátuma. == KIT Hírlevél, 2004/10., márc.10.

**Rec.:**MATISZ Hírlevél 2004, 3. [http://www.matisz.hu/MAKHIR/2004/hirlevel2004\\_03.html#kit](http://www.matisz.hu/MAKHIR/2004/hirlevel2004_03.html#kit)

Our electronic era and bibliographic informations : Computer-related bibliographic data formats, metadata formats and BDML == "Librarianship in the Information Age", The 13<sup>th</sup> BOBCATSSS Symposium, 31 Jan – 2 Febr 2005. Budapest : ELTE, 2005., p. 211-231.

**Rec.:** 13. BOBCATSSS – 2005, Budapest / Balázs S. N. [et al.] == Könyv, könyvtár, könyvtáros, 14.évf. 6.sz. (2005), p. 33-42.

### OTDK

1997. **Kiemelt első helyezés a XXIII. Országos Tudományos Diákköri Konferencián** a Könyvtár-Informatikai Szekcióban. A dolgozat és előadás címe: "A retrospektív konverzió problémái a KLTE Könyvtárában"

### Referált publikációk:

A retrospektív konverzió problémái a Kossuth Lajos Tudományegyetem Könyvtárában == Könyv és könyvtár, 18. (1997.), p. 163-188.

MAKSZAB 50102/97 A, MANCI 29066

Szürke irodalom szolgáltatás az Interneten : az európai SIGLE projekt és adatbázis == Könyvtári Figyelő, Úf., 10.= 46. évf. 4. sz. (2000.), p. 586-602.

LISA 8467 (2001. nr 8), MAKSZAB 55960/00 A, MANCI 39649

Az XML : új perspektívák a könyvtár-informatikában ==Tudományos és Műszaki Tájékoztatás, 51. évf. 2 (2004), p. 61-71.

LISA 294406, MAKSZAB 61020

A BDML jelölő nyelv fejlesztési elvei == Természettudományi Közlemények. Nyíregyháza : Nyíregyházi Főiskola, 2004., p. 251-256.

A Neumann-galaxis == Könyvtári Figyelő, Úf., 15.= 51. évf. 2. sz. (2005.), p. 275-285.

Metadata formats and the new markup language from library field == Teaching Mathematics and Computer Science, Vol. 3. Issues 1. (2005), p. 85-102.

### **Konferencia kiadványában megjelent cikkek:**

Remote access electronic documents in library systems == 2nd International Conference „Multimedia in Business 2000”. Czestochowa : Technical University of Czestochowa, 2000., p. 453-462. Megjelent CD-ROM-on is.

Az elektronikus dokumentumok megjelenése a könyvtári rendszerben : gondolatok az informatikus könyvtárosképzés és –továbbképzés új kihívásairól a bibliográfiai adatfeldolgozás területén == Informatika a felsőoktatásban 2002. Debrecen : Debreceni Egyetem, 2002., p. 152-158.

XML jelölőnyelv alkalmazása könyvtári területen == A Magyar Tudomány Napja. Nyíregyházi Főiskola, 2003. nov. 11. A konferencia kiadvány megjelenés alatt.

Our electronic era and bibliographic informations : Computer-related bibliographic data formats, metadata formats and BDML == "Librarianship in the Information Age", The 13<sup>th</sup> BOBCATSSS Symposium, 31 Jan – 2 Febr 2005. Budapest : ELTE, 2005., p. 211-231.

Iszály B., Salgáné Medveczki M.: Elektronikus zeneműtár és metaadat-rendszere == Informatika a felsőoktatásban 2005. Debrecen : Debreceni Egyetem, 2005., p. 113. A teljes anyag CD-ROM-on.

Bibliographic information in our days == EFITA/WCCA 2005 Joint Conference, Information systems and databases session. Proceedings. Vila Real, Portugal, 2005., p. 456-463.

### **Konferencia előadások:**

A retrospektív konverzió problémái a KLTE Könyvtárában == Kovács Máté Emlékkülés. KLTE, 1996.

Les pratiques de campement en Hongrie == CEMEA (Centres d'Entrainement Aux Méthodes d'Education Active) konferencia. Budapest, 1996. ápr. 19.

Szürke irodalom szolgáltatás a weben == "Szürke irodalom" - a jelenségről és jövőjéről az elektronikus környezetben A Magyar Könyvtárosok Egyesülete Társadalomtudományi Szekciójának szakmai konferenciája a szürke irodalomról. Országos Széchényi Könyvtár, 2000. nov. 14.

[http://www.oszk.hu/hun/publ/konferencia/konf2000/konf2000\\_szurkirod\\_hu.htm](http://www.oszk.hu/hun/publ/konferencia/konf2000/konf2000_szurkirod_hu.htm)

Remote access electronic documents in library systems == 2nd International Conference „Multimedia in Business 2000”. Czestochowa : Technical University of Czestochowa, 2000.

Elektronikus dokumentumok online bibliográfiai feldolgozása a hazai és külföldi gyakorlatban == Elektronikus dokumentumok bibliográfiai számbavétele, a Magyar Könyvtárosok Egyesületének szakmai konferenciája. Országos Széchényi Könyvtár, 2001. okt. 18.

[http://www.oszk.hu/hun/publ/konferencia/konf2001/konf2001\\_mkeszamba\\_hu.htm](http://www.oszk.hu/hun/publ/konferencia/konf2001/konf2001_mkeszamba_hu.htm)

Az elektronikus dokumentumok megjelenése a könyvtári rendszerben : gondolatok az informatikus könyvtárosképzés és –továbbképzés új kihívásairól a bibliográfiai adatfeldolgozás területén == Informatika a felsőoktatásban 2002. aug. 30.

Az iskolai könyvtári állomány feldolgozásának új lehetőségei == Tudásalapú társadalom – információs műveltség. Országos Pedagógiai Könyvtár és Múzeum Nemzetközi Konferenciája, 2003. okt. 16-17.

XML jelölőnyelv alkalmazása könyvtári területen == A Magyar Tudomány Napja. Nyíregyházi Főiskola, 2003. nov. 11.

A Neumann-galaxis == "2. Gyires Béla Informatikai Nap", Debreceni Egyetem Informatikai Intézet, 2004. máj. 14.

Our electronic era and bibliographic informations : Computer-related bibliographic data formats, metadata formats and BDML == "Librarianship in the Information Age", The 13<sup>th</sup> BOBCATSSS Symposium, 31 Jan – 2 Febr 2005.

Bibliographic information in our days == EFITA/WCCA 2005 Joint Conference, Information systems and databases session. Vila Real, Portugal, 25 July 2005.

### **Szakmaspecifikus alkotások:**

Online bibliográfiai feldolgozás c. **akkreditált továbbképzési program**. 2001

BDML (Bibliographic Description Markup Language) formátum fejlesztése 2003-tól

Részvétel az EMIR (Elektronikus Magyar Irodalom) projektben 2005-től. <http://www.emir.hu>

### **Szoftverfejlesztés**

#### **IntSys Vállalatirányítási Rendszer**

Raktár- és Készletnyilvántartó Modul  
Termelésirányítási és –optimalizáló Modul  
Számlázó és Pénzügyi Modul  
Bér- és Munkaügyi Modul  
Befektetett- és Tárgyieszköz Modul  
Gépjárműnyilvántartó és Menetlevél-kezelő Modul

#### **Kereskedelmi programcsomag**

#### **Növénytermesztési Technológia Tervező Rendszer**



EGY ÚJ JELÖLŐ NYELV (METAADAT-RENDSZER) KIDOLGOZÁSA A  
SZÁMÍTÓGÉPES BIBLIOGRÁFIAI ADATFELDOLGOZÁSHOZ

Értekezés a doktori (Ph.D.) fokozat megszerzése érdekében  
az Informatika tudományágban

Írta: Salgáné Medveczki Marianna okleveles informatikus könyvtáros – informatika  
tanár

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

Témavezető: Dr. Szabó József

A doktori szigorlati bizottság:

elnök: Dr. Sztrik János .....

tagok: Dr. Kormos János .....

Dr. Csőke Lajos .....

A doktori szigorlat időpontja: 2005. november 30.

Az értekezés bírálói:

Dr. Eszenyiné Borbély Mária .....

Dr. Fazekas Gábor .....

A bírálóbizottság:

elnök: Dr. Daróczy Zoltán .....

tagok: Dr. Bényei Miklós .....

Dr. Kovács Emőd .....

Dr. Kozma László .....

Dr. Virágos Márta .....

Az értekezés védésének időpontja: 200... ..