

Melléklet: Elektronikus könyvformátum leírása

Tartalom

<u>BEVEZETŐ.....</u>	<u>3</u>
<u>1. A DOCBOOK XML ELEKTRONIKUS DOKUMENTUMFORMÁTUM.....</u>	<u>4</u>
1.1. XML.....	4
1.2. DocBook.....	4
<u>2. MATEMATIKAI KIFEJEZÉSEK, EGYENLETEK.....</u>	<u>11</u>
<u>3. TOVÁBBI FORMAI KÖTÖTTSÉGEK.....</u>	<u>12</u>
<u>3. MARC ÉS DUBLIN CORE METAADATOK MEGADÁSA.....</u>	<u>13</u>
MARC.....	13
DUBLIN CORE.....	15
<u>6. MINTADOKUMENTUMOK.....</u>	<u>15</u>

Bevezető

Jó néhány, dokumentumokkal kapcsolatos feladat esetén, amikor nagy mennyiségű dokumentum egységes formátumban való elkészítése, az egyes dokumentumoknak ugyanazon hordozón, de eltérő megjelenésben történő publikálása (pl. navigációs elemekkel ellátott HTML formátum és nyomtatható HTML formátum), vagy a dokumentumok különböző médián vagy formátumban való publikálása a cél, a legkézenfekvőbb megoldás valamilyen újrahasznosítható dokumentumformátum használata.

E feladatok megvalósításához alapvetően szükséges az, hogy az elektronikus dokumentum készítésének folyamatát két részre válasszuk szét. Az első logikai réteg a dokumentum szerkezetét írja le, azaz hogyan tagolódik fejezetekre, bekezdésekre, stb., míg a második rétegnek elő kell tudnia állítani a megfelelő kimenetet ennek a struktúrának és a megjelenítő eszköznek, célformátumnak ismeretében.

Az így felépített dokumentumot a gépi feldolgozásakor azonos tartalommal lehet a különböző megjelenítő eszközökre publikálni, vagy fordítva, a dokumentum változásakor a közös forrásból egyszerűen lehet a megjelenítő réteg megváltoztatása nélkül újra publikálni. Ezt a gyakorlatban a dokumentumok valamely szabványos SGML vagy XML formátumban való kódolásával és azok transzformációival lehet a legegyszerűbben megoldani.

1.) Az elektronikus formában leadandó könyvek formátumául a Docbook XML-t választottuk. Az előírt formátum kialakításakor, illetve kiválasztásakor olyan, már létező megoldásokat vizsgáltunk meg, melyek a fentieknek megfelelnek, képesek az elektronikus dokumentumok metaadatainak megfelelő tárolására, kellően elterjedtek, a nagyobb szoftverfejlesztők támogatják és léteznek megfelelő minőségű transzformációs eszközök. További követelmény, hogy a formátum legyen szabadon módosítható és bővíthető, valamint nyílt szabvány, azaz ne kötődjön egyetlen nagy gyártóhoz sem. A DocBook XML formátum olyan általános dokumentumformátum, melyet a kilencvenes évek eleje óta használnak általános témájú dokumentumok és technikai leírások tárolására. A formátumot támogatják a Linux-os fejlesztői közösségek, valamint kereskedelmi és nyílt forrású szoftverek sora. Jelenleg az OASIS nonprofit konzorcium kezeli. Nyílt, jól dokumentált szabvány, élénk fejlesztői közösséggel.

<http://www.docbook.org/>

<http://www.oasis-open.org/about/>

2.) Az elektronikus könyvek metaadatait a MARCXML és Dublin Core szabványok szerint szükséges mellékelni.

<http://dublincore.org/>

<http://www.loc.gov/standards/marcxml/>

3.) Azon könyvek esetén, melyek matematikai formulákat tartalmaznak, az egyenletek megadása képként, szöveggént és MathML nyelven is szükséges.

4.) Szállítandó tartalom:

A leadott elektronikus könyv a következő, kötött könyvtárszerkezetben tárolt fájlokból kell álljon:

./book.xml - a tankönyv Docbook XML formátumban

./images/ - a tankönyv képanyagát tároló alkönyvtár *javasolt* elnevezése. A könyv XML forrásában a képek elérését relatív útvonallal szükséges megadni

./meta/dc.xml – a tankönyv metaadatai Dublin Core XML formátumban (a könyvtárnév és a fájlnev megtartása *kötelező*). A fájl ellenőrzéséhez (validálásához) használandó XML séma helye: <http://dublincore.org/schemas/xmls/qdc/2003/04/02/simpledc.xsd>.

./meta/marc.xml – a tankönyv metaadatai MARCXML formátumban (a könyvtárnév és a fájlnev megtartása *kötelező*). A fájl ellenőrzéséhez (validálásához) használandó XML séma helye: <http://www.loc.gov/standards/marcxml/schema/MARC21slim.xsd>.

./math/{*egyenlet id*}.math – *kötelezően* a ./math alkönyvtárban külön fájlokként tároljuk a tankönyv forrásában {*egyenled id*} id-vel jelölt formulákat (equation, inlineequation, informalequation). A fájlok ellenőrzéséhez (validálásához) használandó séma helye: <http://www.w3.org/Math/XMLSchema/mathml2/mathml2.xsd>.

5.) A fájlnevekben csak az angol ábécé kisbetűi, valamint a '-' (kötőjel) és '_' (aláhúzásjel) szerepelhetnek, nagybetűk, ékezetes karakterek vagy szóköz nem.

1. A DocBook XML elektronikus dokumentumformátum

1.1. XML

Az XML, vagyis *eXtensible Markup Language* jelentése: bővíthető jelölőnyelv. A World Wide Web Consortium XML 1.0 ajánlása olyan szintaxist ad meg, melyet betartva különböző jelölő nyelvek (mint például az XHTML) hozhatóak létre. Ezeknek a nyelveknek a segítségével szabványos módon megoldható a különböző alkalmazások, platformok közötti adatcsere, felhasználva már meglévő feldolgozó programokat, ami nagy mértékben lecsökkenti a fejlesztés időtartamát és költségét. A meglévő adatok szolgáltatása pedig egyszerűbbé válik, egy közös adatforrásból, de több formátumban is.

<http://www.w3.org/XML/>
<http://www.xmlinfo.hu/XML/xmlalapok.html>
<http://www.xmlinfo.hu/XML/miazaxml.html>
http://www.w3c.hu/forditasok/XML_10_pontban.html

Az XML – a HTML nyelvtől eltérően – a tag-ek elnevezésében a kis- és nagybetűk használatára érzékeny jelölőnyelv, tehát a <pelda> a <Pelda> és a <PELDA> tag-ek egymástól különbözőek. A DocBook XML verziójában a tag-eket mindig kisbetűs írásmóddal használjuk. A tag-ekhez rendelt tulajdonságokat mindig idézőjelek közé kell zárunk. Erre használhatjuk az egyszeres (') és dupla (") idézőjelet, azonban a nyomdai idézőjelet („ és ”) nem.

1.2. DocBook

Az elektronikus dokumentumok tárolására olyan saját XML formátumot használunk, amely több szabványos és elterjedt formátumra épül. A dokumentum felépítését tekintve a DocBook XML szabványt követi, ezt ajánlásunkban egyrészt szűkítjük (lásd 2. pont), illetve kiegészítjük olyan elemekkel, amelyeket a dokumentumra vonatkozó metainformációk tárolására használunk majd.

A DocBook XML elemkészlet

A DocBook dokumentumtípust több mint tíz évvel ezelőtt a HaL Computer Systems és az O'Reilly & Associates hozta létre, később fejlesztésében részt vett a Novell, a Digital, a Hewlett Packard, a Sun, az SCO, és mások. Nyílt szabvány, jelenleg egy nonprofit konzorcium, az OASIS (Organization for the Advancement of Structured Information Standards) kezeli.¹

Kik használják?

Mivel a DocBook dokumentumok felépítése követi a könyvek szerkezetét, és nagyszámú elemmel rendelkezik a tartalom elrendezéséhez, jelentős létszámú szerzői közösség használja, továbbá támogatja számos kereskedelmi forgalomban lévő és szabad szoftver. A DocBook formátumot széles körben használják a fent említett számítástechnikai cégeken kívül a szabad szoftvereket fejlesztő közösségek pl. a Linux Documentation Project, a FreeBSD Documentation Project, a GNOME Documentation Project, a KDE Documentation Project, a Caldera Systems, a Mandrakesoft, a Red Hat és a SuSE, több millió oldalnyi dokumentáció, elektronikus könyv és más kiadvány tárolására.

Előnyei

A DocBook formátum minden olyan előnnyel rendelkezik, amivel az XML dokumentumok általában: logikus szerkezeti felépítésű, gép által értelmezhető, feldolgozható, valamint platform- és nyelv-független, testre szabható, és bővíthető.

A DocBook formátumot létező eszközök, kereskedelmi és szabad szoftverek már támogatják és ismerik, például az OpenOffice.org irodai szoftvercsomag, az emacs, Quanta, Altova XMLSpy és más XML szerkesztő programok.

A formátum fejlesztése folyamatos, élénk fejlesztői közösség működik közre a kidolgozásában, további alakításában. A DocBook viszonylag hosszabb ideje szerepel a palettán – így kiforrott, kipróbált formátumnak tekinthető.

A már meglévő és testre szabható szoftvereszközökkel a DocBook forrásfájlból konvertálhatunk (X)HTML, HTMLhelp, MIF (Adobe FrameMaker), PDF, PostScript, RTF (Microsoft), TeX, TXT, XHTML formátumokba, azaz így egy fájlból tudjuk kiszolgálni a különböző formátumú, de azonos tartalmú dokumentumok iránti igényeket.

¹ Az OASIS támogatói: <http://www.oasis-open.org/about/>.

Az alábbiakban röviden bemutatjuk egy DocBook dokumentum felépítését, a benne használt legfontosabb elemeket. Az interneten angol nyelven természetesen elérhető a [teljes dokumentáció](http://docbook.org/tdg/en/) (<http://docbook.org/tdg/en/>).

A DocBook XML felépítése

1.) Fő elemek

A DocBook elektronikus könyvek mellett alkalmas más dokumentumok, például cikkek, különálló fejezetek, weboldalak, szójegyzékek, vagy számítógépes prezentációk leképezésére is. Az általunk elvárt DocBook dokumentum fő eleme a `<book>` (<http://www.docbook.org/tdg/en/html/book.html>) elem. Ezen belül pedig – a példa szerint – elkülönül egymástól a könyvre vonatkozó metainformációk halmaza, valamint a könyv szerkezetének megfelelően tagolt tartalom.

```
<book lang="hu">
  <bookinfo> ... </bookinfo>
  <preface>
    <title>Előszó</title> ...
  </preface>
  <chapter> ... </chapter>
  <chapter> ... </chapter>
  <chapter> ... </chapter>
  <bibliography> ... </bibliography>
  <appendix> ... </appendix>
  <index> ... </index>
</book>
```

`<bookinfo>` - (<http://www.docbook.org/tdg/en/html/bookinfo.html>) Az elektronikus könyvre vonatkozó metainformációk

`<preface>` - (<http://www.docbook.org/tdg/en/html/preface.html>) Előszó

`<chapter>` - (<http://www.docbook.org/tdg/en/html/chapter.html>) A könyv fejezetei

`<bibliography>` - (<http://www.docbook.org/tdg/en/html/bibliography.html>) Bibliográfia

`<appendix>` - (<http://www.docbook.org/tdg/en/html/appendix.html>) Függelék

`<index>` - (<http://www.docbook.org/tdg/en/html/index.html>) Tárgymutató

2.) Metainformációk

A könyvre vonatkozó metainformációkat a `<bookinfo>` elem tárolja az alábbi példa szerint, hasonlóképpen a HTML oldalak `<head>` eleméhez, azonban itt jóval több információt, megfelelően tagolva adhatunk meg.

```
<bookinfo>
  <title>A könyv címe</title>
  <author>
    <surname>Postak</surname>
    <firstname>Ábel</firstname>
  </author>
  <copyright>
    <year>2004</year>
    <holder>Postak Ábel</holder>
  </copyright>
</bookinfo>
```

A DocBook formátum nagyszámú elemkészletéből az elektronikus könyv megfelelő leírásához a fenti egyszerű példánál több elemet kell használnunk.

`<title>` - (<http://www.docbook.org/tdg/en/html/title.html>) A könyv címe

`<authorgroup>` - (<http://www.docbook.org/tdg/en/html/authorgroup.html>) Ha a könyv több szerző műve, az `<author>` (<http://www.docbook.org/tdg/en/html/author.html>), azaz szerző elemeket csoportba foglaljuk.

`<author>` - A könyv szerzője. Tartalmazza a szerző vezetéknévét `<surname>`, keresztnévét `<firstname>`, munkahelyét, kiadóját vagy más kapcsolódását az `<affiliation>` tag-ben, (<http://www.docbook.org/tdg/en/html/affiliation.html>), valamint a címét `<address>` (<http://www.docbook.org/tdg/en/html/address.html>). A címet értelemszerűen tagoljuk az `<street>`, `<city>`, `<country>`, `<postcode>`, `<email>` tag-ek használatával.

`<editor>` - (<http://www.docbook.org/tdg/en/html/editor.html>) Szerkesztő.

`<copyright>` - (<http://www.docbook.org/tdg/en/html/copyright.html>) Az elektronikus könyvre vonatkozó szerzői jogi információk: kiadás éve `<year>`, és a jogtulajdonos `<holder>` (<http://www.docbook.org/tdg/en/html/holder.html>).

`<edition>` - (<http://www.docbook.org/tdg/en/html/edition.html>) Kiadás

<pubdate> - (<http://www.docbook.org/tdg/en/html/pubdate.html>) A kiadás éve

<legalnotice> - (<http://www.docbook.org/tdg/en/html/legalnotice.html>) A szerzői jogokra vonatkozó hosszabb magyarázó szöveg.

<biblioid> - (<http://www.docbook.org/tdg/en/html/biblioid.html>) Azonosító, pl. ISBN, ISSN, vagy internet URL. A korábbi DocBook verziókban előforduló, de az új változatban elhagyásra kerülő <isbn> és <issn> tag-eket NE használjuk.

<keywordset> - (<http://www.docbook.org/tdg/en/html/keywordset.html>) Kulcsszavak gyűjtőeleme, minden kulcsszó <keyword> elemmel jelölve.

<abstract> - (<http://www.docbook.org/tdg/en/html/abstract.html>) Rövid összefoglalás, kivonat.

<publisher> - (<http://www.docbook.org/tdg/en/html/publisher.html>) A Kiadó neve. Az <address> (<http://www.docbook.org/tdg/en/html/address.html>) elemen belül a kiadó címe.

<revhistory> - (<http://www.docbook.org/tdg/en/html/revhistory.html>) Ha a dokumentumban változtatások történtek, a változások bejegyzése.

<printhistory> - (<http://www.docbook.org/tdg/en/html/printhistory.html>) Az eredeti könyv nyomtatásban való megjelenéseire vonatkozó információk

3.) Szöveg, szövegformázás, listák

a.) Struktúra

Az elektronikus könyv szövege fejezetekre oszlik fel, ezt jelöljük a <chapter> tag-gel (<http://www.docbook.org/tdg/en/html/chapter.html>), ezen belül pedig öt szintre tagolhatjuk, <sect1> ... <sect5> (<http://www.docbook.org/tdg/en/html/sect1.html>). A szakaszokon belüli, bekezdésszintű szövegességet <para> jelöli.

```
<chapter>
  <title>A Lundi alapelvek</title>

  <sect1>
    <title>A kezdetek</title>
    <para>Az e-kultúra előretörésével világossá vált,
      hogy az európai kulturális műhelyek, a könyvtárak,
      egyetemek, más tartalomszolgáltatók esetleg nem fognak
      tudni megbirkózni a még zömmel papír alapú tudásbázisok
      átalakításával úgy, hogy azt kompatibilissé lehessen tenni
      az Unió felhasználói számára.</para>
  </sect1>
</chapter>
```

Az egyes fejezetekről és szakaszokról többféle metainformációt (kulcsszavak, nyelv, mottó) tárolhatunk (ezek leírása az internetes dokumentációban megtalálható), azonban a legfontosabb a cím <title> (<http://www.docbook.org/tdg/en/html/title.html>) megadása, mert a HTML formátumba való konvertáláskor ez jelenik meg <h1>...<h6> címsorokként. Több nyelvű dokumentum esetén a magasabb szinten megadott lang="hu" (vagy lang="en", stb) tulajdonságot felülbíráhatjuk. Fontos ügyelni a szakaszok megfelelő egymásba ágyazására, azaz egy <sect1>-en belül ne következzen rögtön <sect3>, hanem jelenjen meg a <sect2> elem is, még ha azon belül az alacsonyabb szintű szakaszon kívül más tartalmat nem helyezünk el. Ennek a hierarchiának a pontos betartása az automatikus tartalomjegyzék-generálás miatt fontos lesz. Ha olyan hierarchiát szeretnénk kialakítani, ami az öt szintű, számozott szakaszokkal nem megoldható, ill. annál mélyebb szintig szükséges, a <section> (<http://www.docbook.org/tdg/en/html/section.html>) elemet használhatjuk, és tetszőleges számban önmagába ágyazhatjuk.

Bekezdés megjelölésére általában a <para> (<http://www.docbook.org/tdg/en/html/para.html>) elem szolgál. Emellett két változatát is felhasználhatjuk: a <simpara> (<http://www.docbook.org/tdg/en/html/simpara.html>) azzal a megkötéssel használható, hogy nem tartalmazhat további bekezdésszintű (blokként megjelenített) elemeket, pl. listákat. A <formalpara> (<http://www.docbook.org/tdg/en/html/formalpara.html>) olyan bekezdést jelöl, melynek kötelezően címet <title> is adunk.

```
<para>Bekezdés szövege</para>

<formalpara>
  <title>Ez a bekezdés címe</title>
  <para>A bekezdés szövege. A bekezdés szövege. A bekezdés szövege. </para>
</formalpara>
```

A bekezdés szintjén érdemes megkülönböztetnünk a blokként és a szövegben folyamatosan (inline) megjelenített elemeket. A blokkelemek általánosságban a bekezdésszintű elemek, például listák, táblázatok, a szövegből kiemelt képek, ábrák, képletek, idézetek, stb. A legtöbb blokk elem tartalmazhat további blokk elemeket és folyamatos szöveget. Az inline elemek általában sortörés nélkül, a szövegben jelennek meg, jelenlétüket betűstílus-váltás mutathatja, de legtöbbszor ezeket nem is különböztetjük meg vizuálisan. Ezek az elemek tartalmazhatnak szöveget, illetve adatokat (nevet, címet, keresztneveket, formázott vagy alsó/felső indexben elhelyezett szöveget, stb.).

b.) Szövegformázás

Dőlt és félkövér betűstílust az `<emphasis>` (<http://www.docbook.org/tdg/en/html/emphasis.html>) tag használatával állíthatunk be a szöveg részeire. A paraméter nélkül megadott `<emphasis>` a legtöbb esetben dőlt betűs írásmódot jelent, míg a félkövér megjelenítést az `<emphasis role="strong">` használatával érhetjük el. Általában törekedjünk arra, hogy a szöveg formázása önmagában ne hordozzon információt (tekintettel a vakokra és gyengénlátókra). A szövegen belül alsó és felső indexbe kerülő szakaszokat (pl. H_2O vagy $a^2+b^2=c^2$) értelemszerűen a `<subscript>` alsó index, illetve `<superscript>` felső index tag-ekkel jelölhetjük.

c.) Felsorolások

Felsorolások, listák készítésekor használhatunk számozatlan vagy sorszámozott listát.

Példa egyszerű felsorolásra a `<simplelist>` (<http://docbook.org/tdg/en/html/simplelist.html>) elem használatával:

```
<simplelist>
  <member>lista eleme</member>
  <member>lista eleme</member>
  <member>lista eleme</member>
</simplelist>
```

A megjelenítéskor ez az egyszerű lista viselkedhet pl. mondaton belüli felsorolásként, de ábrázolható sortöréssel is, minden tagja külön-külön.

Példa számozatlan, a sorok elején általában körrel vagy más grafikus szimbólummal jelölt listára:

```
<itemizedlist>
  <listitem><para>lista eleme</para></listitem>
  <listitem><para>lista eleme</para></listitem>
  <listitem><para>lista eleme</para></listitem>
</itemizedlist>
```

A felsorolás objektumok egymásba ágyazhatóak, tehát egy `<listitem>`-en belül létrehozhatunk újabb, alacsonyabb szintű felsorolást (vagy akár sorszámozott listát is), ezt a megjelenítő programok és a konverterek képesek értelmezni. A lista stílusa testre szabható az `<itemizedlist>` elem „mark” tulajdonságával.

Példa számozott listára:

```
<orderedlist numeration="arabic">
  <listitem><para>sorszámozott listaelem</para></listitem>
  <listitem><para>sorszámozott listaelem</para></listitem>
  <listitem><para>sorszámozott listaelem</para></listitem>
  <listitem><para>sorszámozott listaelem</para></listitem>
</orderedlist>
```

Az `<orderedlist>` elem „numeration” tulajdonságával megadhatjuk a sorszámozás stílusát (arab számok, kis/nagybetűs római számok, kis/nagybetűs abc), illetve azt, hogy egy előző sorszámozott listát folytatunk-e.

d.) Idézet

Amennyiben a szövegünkben idézet szerepel, azt a `<blockquote>` tag-gel jelölhetjük.

e.) Sortöréssel formázott szöveg

Azokban az esetekben, amikor szükséges megtartanunk egy szöveg formázását (például egy vers sorait), az ilyen módon megjelenítendő szöveget a `<literallayout>` elembe tároljuk.

```
<blockquote>
  <attribution>Radnóti Miklós: <citetitle>Hetedik Ecloga/citetitle</attribution>
  <literallayout>
    Látod-e, esteledik s a szögesdróttal beszegett, vad
    tölgykerítés, barak oly lebegő, felszívja az este.
    Rabságunk keretét elereszti a lassu tekintet
    és csak az ész, csak az ész, az tudja, a drót feszülését.
    Látod-e drága, a képzelet itt, az is így szabadul csak,
    megtöretett testünket az álom, a szép szabadító
    oldja fel és a fogolytábor hazaindul ilyenkor.
  </literallayout>
</blockquote>
```

4.) Hiperhivatkozások

A dokumentumban az `<mlink>` tag-gel helyezhetünk el hiperhivatkozásokat, például:

```
<mlink url="http://www.hik.hu/">A HIK honlapja</mlink>
```

A dokumentumon belüli hivatkozásokat a `<link>` elemmel hozhatjuk létre.

```
<sect1>
  <para>Ebben a mondatban van egy <link linkend='nextsect'>link</link>,
  ami a dokumentumon belül egy meghatározott és elnevezett szakaszra
  mutat.</para>
  <sect2 id='nextsect'><title>Alfejezet</title>
  <para>Ez a szakasz az előző bekezdésben szereplő link
  célja.</para>
</sect2>
</sect1>
```

5.) Táblázatok

Cím és fejléc nélküli egyszerű táblázatot az `<informaltable>` elemmel hozhatunk létre (<http://www.docbook.org/tdg/en/html/informaltable.html>). A `<tgroup>` elem egy táblázat valamely logikai egységét foglalja magába. Az általunk használt egyszerűbb táblák általában egy `<tgroup>`-ból állnak. A `<tgroup>` elem „cols” tulajdonságával állíthatjuk be az oszlopok számát, a táblázat sorait a `<row>` elemmel jelöljük, ezen belül az egyes cellákat pedig az `<entry>` tag-gel. Az egyes oszlopok tulajdonságait (elnevezés, sorszámozás, szöveg igazítása, stb.) a `<colspec>` elemmel állíthatjuk be.

```
<informaltable>
  <tgroup cols='2'>
    <tbody>
      <row><entry>1</entry><entry>1</entry></row>
      <row><entry>2</entry><entry>4</entry></row>
      <row><entry>3</entry><entry>9</entry></row>
    </tbody>
  </tgroup>
</informaltable>
```

A bonyolultabb táblázatok, melyek összevont cellákat, táblafejet, címet tartalmaznak, a `<table>` elemmel (<http://www.docbook.org/tdg/en/html/table.html>) hozhatjuk létre.

```
<table frame='all'><title>Sample Table</title>
<tgroup cols='5' align='left' colsep='1' rowsep='1'>
  <colspec colname='c1' />
  <colspec colname='c2' />
  <colspec colname='c3' />
  <colspec colnum='5' colname='c5' />
  <thead>
    <row>
      <entry namest="c1" nameend="c2" align="center">Vízszintes
      cellaösszevonás</entry>
      <entry>a3</entry>
      <entry>a4</entry>
      <entry>a5</entry>
    </row>
  </thead>
  <tfoot>
    <row>
      <entry>f1</entry>
      <entry>f2</entry>
      <entry>f3</entry>
      <entry>f4</entry>
      <entry>f5</entry>
    </row>
  </tfoot>
  <tbody>
    <row>
      <entry>b1</entry>
      <entry>b2</entry>
      <entry>b3</entry>
      <entry>b4</entry>
      <entry morerows='1' valign='middle'><para>Függőleges
      cellaösszevonás</para></entry>
    </row>
    <row>
      <entry>c1</entry>
      <entry namest="c2" nameend="c3" align="center" morerows='1'
      valign='bottom'>Több sort és oszlopot átfogó cella</entry>
      <entry>c4</entry>
    </row>
    <row>
      <entry>d1</entry>
      <entry>d4</entry>
    </row>
  </tbody>
</table>
```

```
<entry>d5</entry>
</row>
</tbody>
</tgroup>
</table>
```

A fenti példából az alábbi táblaszerkezet jön létre:

Vízszintes cellaösszevonás				
f1	f2	f3	f4	f5
b1	b2	b3	b4	
c1	Több sort és oszlopot átfogó cella		c4	Függőleges cellaösszevonás
d1			d4	d5

6.) Médiaobjektumok

A szövegbe beillesztett médiaobjektumokat (videó, kép, hang) az `<inlinemediaobject>` elemmel jelöljük (<http://www.docbook.org/tdg/en/html/inlinemediaobject.html>). Az alábbi példában:

```
<para>Einstein leghíresebb egyenlete,
<inlineequation id="eq0001">
<inlinemediaobject>
<imageobject>
<imagedata fileref="figures/emc2.png"/>
</imageobject>
<textobject>
<phrase>E=mc<superscript>2</superscript</phrase>
</textobject>
</inlinemediaobject>
</inlineequation>, kifejezi az anyag és energia közötti
kapcsolatot.</para>
```

látható, hogy a médiaobjektum több, alternatív formában tartalmazhatja ugyanazt az információt. Erre szükség is van: pl. egy `<imageobject>` (<http://www.docbook.org/tdg/en/html/imageobject.html>) grafikus objektumot a HTML formába való konvertáláskor így el tudunk látni megfelelő ALT attribútummal. A médiaobjektumok használatakor fontos szem előtt tartani azt, hogy egy adott formátumot esetleg a célközönség/célmédia nem képes megjeleníteni, ezért mozgókép mellett használjunk egy (elegendő információt tartalmazó) állóképet, tekintettel a nyomtatható formátumokba való konvertálásra, illetve állókép mellett mindig adjunk meg elegendő szöveges információt, tekintettel a vakok és gyengénlátók igényeire, az általuk használt felolvasó eszközökre. A DocBook következő verzióiból kimaradó `<inlinegraphic>` és `<graphic>` elemek helyett (ezek például nem tárolnak alternatív szöveges információt a grafikus objektumról) használjuk az `<inlinemediaobject>` vagy `<mediaobject>` elemeket.

7.) Bibliográfia

Az elektronikus dokumentumban szerepelhet bibliográfia. Ezt a `<bibliography>` elemmel hozhatjuk létre, amire két módszert használhatunk. Az alábbi példában az egyszerűbbet mutatjuk be, de emellett van lehetőség kötöttebb, adatbázisszerű („raw”) formázásra is.

```
<bibliography>
<title>Bibliográfia</title>
<bibliomixed>
<bibliomset>
<title>(Lund-) Action Plan on coordination of digitisation programmes
and Policies, Digital Heritage & Cultural Content</title>
<pubdate>2003</pubdate>.,
http://www.cordis.lu/ist/ka3/digicult/lund_ap_browse.htm
</bibliomset>
</bibliomixed>
<bibliomixed>
<bibliomset>
<title>http://digitalizalas.lap.hu/</title>,
<pubdate>1999-2003</pubdate>., Szerk.:
<editor><surname>Kedves</surname><firstname>Csaba</firstname></editor>
</bibliomset>
</bibliomixed>
</bibliography>
```

8.) Lábjegyzetek

Az alábbi példa bemutatja, hogy miként lehet lábjegyzetet elhelyezni a szövegben a `<footnote>` (<http://www.docbook.org/tdg/en/html/footnote.html>) elem használatával.

```
<para>Ezeknek a kihívásoknak a megnyugtató megoldására-, az Internet kultúráközvetítő szerepére apellálva az Európa Tanács 2000-ben a feirai ülésén elfogadták azt az eEuropa 2002 akciótervet <footnote><para><ulink url="http://europa.eu.int/information_society/eeurope/action_plan/index_en.htm">http://europa.eu.int/information_society/eeurope/action_plan/index_en.htm</ulink></para></footnote>, melynek egyik következménye a 2001-es lundi tanácskozás.
```

A lábjegyzetek számozására nem kell figyelmet fordítani, az a céldokumentumban (HTML, PDF, stb.) automatikusan, a lábjegyzetek sorrendjének megfelelően jön majd létre.

9.) Előszó, Ajánlás, Függelék, Kolofon

Az elektronikus dokumentum ezen (szabadon választható ill. elhagyható) speciális részei a fejezet <chapter> elemhez hasonlóan viselkednek.

<dedication> - (<http://www.docbook.org/tdg/en/html/dedication.html>) Ajánlás. A dokumentum elején jelenik meg.

<preface> - (<http://www.docbook.org/tdg/en/html/preface.html>) Előszó. A dokumentum elején jelenik meg.

<appendix> - (<http://www.docbook.org/tdg/en/html/appendix.html>) Függelék, utószó, mely a dokumentum végén jelenik meg.

<colophon> - (<http://www.docbook.org/tdg/en/html/colophon.html>) Kolofon.

10.) Több nyelvű szövegek

Abban az esetben, ha a szövegünk több nyelvű, mindig használjuk a lang="" tulajdonságot az érintett szakaszban, pl. <sect2 lang="jp">. Az UTF-8 kódolás használatával nem csak latin karaktereket használhatunk, hanem a teljes közel-keleti és ázsiai írásjegy-készletet. [Mintadokumentumunkban](#) az arab, héber, orosz, görög, japán, koreai és thai (vegyes nyelvű) szövegeket mutatjuk be példaként.

Írásirány: Az alapértelmezett írásirány a balról jobbra történő írásmód. Abban az esetben, ha vegyes írásirányú dokumentumokat hozunk létre, az alapértelmezett írásirányt felülbíráhatjuk az adott elem (pl. <setc1>...<sect5>, <section>, <para>, stb.) dir="rtl" tulajdonságával. A fejlett megjelenítő programok (és elsősorban a webes megjelenítést végző böngészőprogramok) megfelelően tudják kezelni a jobbról balra író nyelveket, így lehetséges kevert, a nyugati ábécét is használó, kétirányú szövegek megjelenítése. Ez a szolgáltatás (kétirányú Unicode) az Internet Explorer 5-ös verziójától érhető el, valamint a Gecko motort használó böngészőkben (Mozilla, Netscape 7), valamint az Opera 7.2 (vagy újabb) böngészőkben.

2. Matematikai kifejezések, egyenletek

Ha a szövegünkben matematikai kifejezéseket szeretnénk tárolni, a DocBook-ban a képleteket tárolhatjuk képként és megadhatunk alternatív szöveget (ahogy az előző példában láttuk), valamint MathML-t is használunk a képlet leírására. A MathML (Mathematical Markup Language, matematikai jelölőnyelv) a W3C ajánlása.

A Docbook XML dokumentumban a következő tag-eket használhatjuk:

`<equation>` (<http://www.docbook.org/tdg/en/html/equation.html>) blokként formázott egyenlet,
`<inlineequation>` (<http://www.docbook.org/tdg/en/html/inlineequation.html>) szövegen belüli egyenlet. Ha grafikus objektumot tartalmaz, azt az `<inlinemediaobject>` használatával adjuk meg.
`<informalequation>` (<http://www.docbook.org/tdg/en/html/informalequation.html>): kevésbé részletezett formában megadott (általában cím nélküli) egyenlet.

A szállítandó elektronikus dokumentumban a képletek használatakor mindhárom formában kötelező megadni ugyanazt az információt: MathML-ben, képfájlban és szöveggként is. Az egyenletek MathML verzióit külön fájlban, a könyv XML forrásához képest relatív útvonalon a `math/` alkönyvtárban tároljuk. A fájlnevek képzésének szabálya: egyenlet id-je + „.math”.

Leírása: <http://www.w3.org/TR/MathML2/>

Séma: <http://www.w3.org/Math/XMLSchema/mathml2/mathml2.xsd>

Példa:

```
<?xml version="1.0" encoding="UTF-8"?>
<math xmlns="http://www.w3.org/1998/Math/MathML"
xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.w3.org/1998/Math/MathML
http://www.w3.org/Math/XMLSchema/mathml2/mathml2.xsd">
<semantics>
  <mrow>
    <mi>x</mi>
    <mo>=</mo>
    <mfrac>
      <mrow>
        <mo>-</mo>
        <mi>b</mi>
        <mo>±</mo>
        <msqrt>
          <mrow>
            <msup>
              <mrow>
                <mi>b</mi>
              </mrow>
              <mrow>
                <mn>2</mn>
              </mrow>
            </msup>
            <mo>-</mo>
            <mn>4</mn>
            <mi>a</mi>
            <mi>c</mi>
          </mrow>
        </msqrt>
      </mrow>
    </mfrac>
  </mrow>
  <annotation>x=-b±sqrt(b^2-4ac) over 2a</annotation>
</semantics>
</math>
```

3. További formai kötöttségek.

Ahhoz, hogy az elektronikus könyv egy közös forrásból több célformátumba konvertálható legyen, fontos az XML dokumentum struktúráját, a DocBook XML formátumban tárolt szakasz szerkezetére vonatkozó előírásokat pontosan megtartanunk. A hierarchikus elemek (különösen a <sect1>...<sect5> szakaszok) megfelelő egymásba ágyazása teszi lehetővé például a pontos tartalomjegyzék-generálást. Az XML dokumentum felépítésének követnie **kell** a dokumentum felépítését.

Ahol lehetséges, adjunk meg minél több metainformációt (pl. személynevek jelölése, címek és összefoglalók, URL-ek, stb.) a szöveg egyes részeiről. Vegyük figyelembe a különböző médiumok eltérő adottságait: egy videoobjektumot megjeleníthetünk a képernyőn, de a nyomtatható verzió elkészítéséhez szükséges egy alternatív állókép megadása, valamint elégséges szöveges információt kell nyújtanunk a grafikus objektumokat megjeleníteni nem tudó, kizárólag szöveges formátumok kezelésére képes klienseknek (felolvasógép).

A szállított fájlokat a bevezető fejezetben leírt könyvtárstruktúrában kell elrendezni, a fájlnevezési szabályok betartásával.

Az XML fájlok kódolása **UTF-8** kell legyen. A névvel ellátott entitások (például ´ ê õ illetve û) az á, é, ő, ú és más hasonló, nem-ASCII karakterek jelölésére **nem** használhatóak. Használjunk UTF-8 képes szerkesztő eszközöket. Olyan esetekben, amikor nem rendelkezünk megfelelő beviteli eszközzel (pl. távol-keleti írásjegyek), használhatjuk a megfelelő Unicode karakterkódot, pl. a â vagy â jelölheti az â karaktert, 欢 vagy 欢 a 𐀂 karaktert, stb.

<http://www.fileformat.info/info/unicode/block/index.htm>

<http://www.unicode.org/charts/>

3. MARC és Dublin Core metaadatok megadása

MARC

A MARC elsősorban könyvtári szabvány, a dokumentumok rendszerezésében van kiemelkedő szerepe. A rekordfejet (000), az 001-es mezőt, 005, 008-as-, valamint a nem vastagon szedett mezőket a HIK tölti ki, illetve automatikusan generálja abban az esetben, ha a dokumentum a saját szerverén helyezkedik el. Abban az esetben, amikor nem, akkor ezek a mezők kitöltetlenek maradnak. A Kiadóknak a **vastagon szedett mezőket** kell kitölteniük kötelezően. A kötelezően kitöltendő mezőkhöz a 'Megjegyzés' rovatban segítséget nyújtunk, illetve a következő webhelyen található részletes leírás: <http://www.mokka.hu/hb1.htm#tart>.

A használni kívánt mezők száma	Almezők	Megjegyzés
000	LDR, vagy az ún.: Rekordfej	
001	kódszám	
005	A rekorddal történő utolsó művelet dátuma	
008	Meghatározott jellemzők	
020	\$aISBN \$cÁr	Akkor, ha van eredeti, nyomtatott mű.
035	\$a	
040	\$a \$b	
041	\$ahun	http://www.mokka.hu/hb13.htm
044	\$aHU	http://www.mokka.hu/hb14.htm
080	\$aETO	
092	\$aSzakjelzet \$bCutter	
100	Indikátor:1 \$aNév \$d(mettől meddig él-t)	
130	Egyeségesített cím főtétele	
240(10)	\$aEredeti nyelvű cím \$i(Magyar, finn, észt fordítás)	Csak akkor kötelező, ha fordításról van szó.
245	Indikátor: 10 \$aCím, \$bAlcím, \$cSzerkesztő, Főszerkesztő stb., \$h	
250	\$aPl.:2. jav., bőv. kiad.	
260	\$aHely \$bKiadó \$cÉv \$eNyomda	Akkor, ha van eredeti, nyomtatott mű. Ha nincs, csak \$a és \$c kötelező.
300	\$aOldalszám \$bIll. \$cCm	
490	Indikátor:1 \$a,\$x,\$y,\$v	
500	\$aCím a képernyőről	
538	\$aHozzáférés honnan Pl.www stb. többszörözhető mező!	
503	\$gKötetadat	
504	\$aBibliográfia	
583	\$xNaplózás	egyedi
650	Indikátor:4 \$aTárgyszó, \$Z, \$x,	
653	\$	
700	Indikátor: 1 \$aTeljes név , minden 245\$c-nek megfelelően egy-egy név, \$d(mettől meddig élt) \$c(foglalkozás) \$4(szerk, közreműködő stb.)	\$a kötelező.
710	\$aTestületi név, \$b	Testületi szerzőség esetén kötelező. (Pl. Magyar Tudományos Akadémia)

740	\$a,\$c	
830	\$aEgységesített sorozati cím	
850	\$a	
856	\$uHttp, \$zFtp, \$3, \$2	URL
949	\$aSzakjelzet+Cutter együtt \$n, \$z \$l, \$t	
999	Egyéb adat	

A MARC adatok tárolására a MARCXML szabvány elemkészletét használjuk fel.

Leírása: <http://www.loc.gov/standards/marcxml/>

Séma: <http://www.loc.gov/standards/marcxml/schema/MARC21slim.xsd>

<collection> - a MARCXML fájl gyökéreleme.

<record> - MARC rekord.

<leader> - Rekordfej (leader). Értékét a HIK adja meg.

<controlfield> - A 001, 005 és 008 MARC mezők automatikusan generált mezők, értéküket a HIK adja meg.

<datafield> - Adatmezők. A mező sorszámát a „tag”, indikátorait pedig az „ind1” és „ind2” tulajdonság adja meg.

<subfield> - MARC almezők. Az almező kódját a „code” tulajdonság tárolja.

Példa:

```
<?xml version="1.0" encoding="UTF-8"?>
<collection xmlns="http://www.loc.gov/MARC21/slim"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.loc.gov/MARC21/slim
http://www.loc.gov/standards/marcxml/schema/MARC21slim.xsd">
  <record>
    <leader>00364nmm 220024100 4500</leader>
    <controlfield tag="001">HIK00115848</controlfield>
    <controlfield tag="005">20040218164232.0</controlfield>
    <controlfield tag="008">040218s2004 hu n ua hun d</controlfield>
    <datafield tag="020" ind1=" " ind2=" ">
      <subfield code="a"></subfield>
      <subfield code="b"></subfield>
    </datafield>
    <datafield tag="035" ind1=" " ind2=" ">
      <subfield code="a"></subfield>
    </datafield>
    <datafield tag="100" ind1="1" ind2=" ">
      <subfield code="a">Postak Ábel</subfield>
      <subfield code="d">1980-</subfield>
    </datafield>
  </record>
</collection>
```

Dublin Core

Az elektronikus könyvre vonatkozó Dublin Core metainformációkat az alábbiak szerint kell kitölteni:

1. **Title** (cím), *kötelező*
2. **Creator** (alkotó), *kötelező*
3. **Subject and Keywords** (téma és kulcsszavak), *kötelező*
4. **Description** (leírás), *kötelező*
5. **Publisher** (kiadó), *kötelező*
6. **Contributor** (közreműködő), *kötelező*
7. **Date** (dátum), *kötelező*
Formátuma: <http://www.w3.org/TR/NOTE-datetime>
8. **Type** (dokumentumtípus), *kötelező*
Értéke: **Text** (<http://dublincore.org/documents/dcmi-type-vocabulary/>)
9. **Format** (fájl formátum), *kötelezően text/xml*
10. **Resource Identifier** (azonosító a hálózaton), *nem kötelező*
11. **Source** (forrás), *kötelező*
12. **Language** (nyelv), *kötelező*
13. **Relation** (kapcsolatok), *nem kötelező*
14. **Coverage** (hatókör), *nem kötelező*
A földrajzi hatókör leírásához:
http://www.getty.edu/research/conducting_research/vocabularies/tgn/index.html
15. **Rights** (jogok), *kötelező*

Példa:

```
<?xml version="1.0" encoding="UTF-8"?>
<simpledc
xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="http://dublincore.org/schemas/xmls/qdc/2003/04/02/simpledc.xsd">
  <dc:title xml:lang="hu">A Lundi alapelvek</dc:title>
  <dc:creator xml:lang="hu">Markója Szilárd</dc:creator>
  <dc:subject xml:lang="hu">digitalizálás, digitális tartalomfejlesztés
összehangolása nemzetközi szinten</dc:subject>
  <dc:description xml:lang="hu">Az e-kultúra előretörésével világossá
vált, hogy az európai kulturális műhelyek, a könyvtárak, egyetemek,
tartalomszolgáltatók nem fognak tudni megbirkózni a még zömmel papír alapú
tudásbázisok átalakításával.</dc:description>
  <dc:publisher xml:lang="hu">Hallgatói Információs
Központ</dc:publisher>
  <dc:contributor></dc:contributor>
  <dc:type>Text</dc:type>
  <dc:format>text/xml</dc:format>
  <dc:identifier>[URI]</dc:identifier>
  <dc:source>eredeti [ISBN] [ISSN]</dc:source>
  <dc:date>2005-03-29</dc:date>
  <dc:coverage>2001 - 2004, Magyarország</dc:coverage>
  <dc:language>hu</dc:language>
  <dc:relation></dc:relation>
</simpledc>
```

6. Mintadokumentumok

A kiadók részére [mintadokumentumokat, példafájlokat](#) teszünk közzé az interneten. A dokumentumok formai helyességének ellenőrzésére szolgáló XSD (XML Schema Definition) fájlok az egyes szabványokat (Docbook XML, Dublin Core, MARCXML, MathML) karbantartó, felügyelő szervezetek honlapjairól tölthetők le. Az XSD fájlok letöltése az elektronikus könyvek szerkesztése közben praktikus lehet, azonban a leadott anyag ellenőrzésekor a következő XSD-k szerinti validálás a normatív:

Könyv: <http://docbook.org/xsd/4.4/docbook.xsd>

Dublin Core metaadatok: <http://dublincore.org/schemas/xmls/qdc/2003/04/02/simpledc.xsd>

MARC metaadatok: <http://www.loc.gov/standards/marcxml/schema/MARC21slim.xsd>

MathML formulák: <http://www.w3.org/Math/XMLSchema/mathml2/mathml2.xsd>