**XML**

Az XML dokumentumr´ol bizony´ara mindenki tud m´ar legal´abb annyit, hogy sz¨oveges form´atumban adjuk meg a tartalmat ´es a sz¨ovegben jel¨ol˝o elemek vannak, melyek a hat´arolt r´esz valamilyen tulajdons´ag´ara utalnak. A jel¨ol˝o elemekn´el tal´alhat´o nyit´o ´es z´ar´o tag, amelyek a kapcsol´od´o sz¨ovegr´esz elej´et ´es v´eg´et adj´ak meg. A jel¨ol˝o elemtagokat a ’*<*’ ´es ’*>*’ rel´aci´o jelek hat´arolj´ak. A jel¨ol¨o elemek egym´asba ´agyazhat´ok. P´eld´aul az al´abbi kis minta h´arom elemet mutat be, melyek k¨ozu¨l a ’di´ak’ elem mag´aba foglalja a ’n´ev’ ´es ’oszt´aly’ elemeket:

<di´ak>

<n´ev> P´eter </n´ev>

<oszt´aly> 4B </oszt´aly>

</di´ak>

Az XML legfontosabb c´elja, hogy egy plattform-fu¨ggetlen, alkalmaz´as- fu¨ggetlen adatle´ır´o form´atmot adjon, mely alkalmas a web-re ´epu¨l˝o nagy t¨omegu˝ adat´atvitel megval´os´ıt´as´ara ´es a kapott adatok rugalmas, hat´ekony feldogoz´as´ara. Emiatt az XML t¨obb mint adatt´arol´asi szabv´any, hozz´a tar- toznak a szok´asos kezel´est megval´os´ıt´o feldolgoz´asi szabv´anyok is.

**Helyesen form´alt XML dokumentum krit´eriumai**

Az XML dokumentum sz¨oveges ´allom´anyban t´arolt, melynek szok´asos kiterjeszt´ese az ’xml’.

*•*

Az XML-ben a jel¨ol˝o elemek haszn´alatosak a le´ır´o inform´aci´ok, metaada- tok megad´as´ara.

*•*

Az elemek tagjait a ’*<*’ ´es ’*>*’ karakterek hat´arolj´ak. Tartalom szerint az al´abbi elemt´ıpusok ´ertelmezettek:

*•*

* tartalom elem: a felhaszn´al´o a´ltal megadott tartalom tulajdons´agait

´ırja le, a feldolgoz´o programnak sz´ol´o inform´aci´okat tartalmaz

* deklar´aci´os elem: a feldolgoz´onak sz´ol´o instrukci´okat tartalmaz
* megjegyz´es elem: az olvas´onak, a programoz´onak sz´ol´o inform´aci´okat tartamaz

A tartalom elemek lehetnek egytagu´ak ´es k´ettagu´ak. A m´asik k´et el- emt´ıpus egytagu´.

* + - A k´ettagu´ tartalom elem szerkezete:

<elemn´ev jellemz¨ok\_list´aja> tartalom </elemn´ev>

Az elemmel megadott tulajdons´ag a k¨ozrefogott tartalomra vonatkozik. Minden nyit´o elemtaghoz l´eteznie kell egy z´ar´o elemtagnak.

* + - Az egytagu´ tartalom elem szerkezete:

<elemn´ev jellemz¨ok\_list´aja/>

A tartalom elemek neve tetsz˝oleges sz´o lehet. Teh´at a n´ev csak egy sz´ob´ol a´llhat, de a sz´oalakok halmaza nem el˝ore defini´alt. Ez a tulaj- dons´ag l´enyegesen elt´er a m´ar ismert HTML form´atumt´ol, hiszen ott el˝oredefini´alt elemneveket lehet csak haszn´alni. Az XML c´elja, hogy az elem ne csak r¨ogz´ıtett form´atum tulajdons´agok le´ır´as´ara szolg´aljon, hanem tetsz˝oleges tartalom megjel¨ol´est is megval´os´ıthassuk a´ltala. Ekkor az elemn´ev a tartalom szemantikai, jelent´es oldal´at adja meg. Ha p´eld´aul a tartalom a 23-as sz´am, akkor az ’´eletkor’ elemn´ev jelzi, hogy a sz´am egy ´eletkor ´ert´eket takar:

*•*

<´eletkor> 23 </´eletkor>

* + - A deklar´aci´os elem szerkezete:

<?elemn´ev jellemz¨ok\_list´aja ?>

Az XML dokumentum elej´an ´all´o ’xml’ elemnevu˝ deklar´aci´o mellett a dokumentum egyedi deklar´aci´os elemeket is tartalmazhat

* + - A megjegyz´es elem szerkezete:

<! megjegyz´es >

A szabv´any alak, amely biztos´ıtja, hogy az elem nem egy speci´alis feldolgoz´o programnak sz´ol, a k¨ovetkez˝o alaku´:

<!-- megjegyz´es -->

A jellemz˝ok az elem valamely tulajdons´ag´at, viselked´esi param´eter´et adj´ak meg. Egy jellemz˝o megad´asa egy p´arossal t¨ort´enik, a p´arosban a jellemz˝o nev´et ´es az ´ert´ek´et t´ars´ıtjuk ¨ossze. A megad´as form´atuma:

*•*

jellemz¨o\_neve = "´ert´ek"

A jellemz˝oh¨oz kapcsolt ´ert´eket teh´at mind´ıg id´ez˝ojelek k¨oz¨ott kell sz- erepeltetni. A lista tetsz˝oleges darabsz´amu´ p´arosokat tartalmaz, azaz lehet u¨res is.

Az XML dokumentum els˝o sor´anak egy XML deklar´aci´os elemet kell tartalmaznia:

*•*

<?xml version="vv" ?>

ahol xx egy verzi´osz´amot ad meg. Ezzel az elemmel jelezzu¨k az XML feldolgoz´onak, hogy a forr´asf´ajl szintaktik´aja mely XML szabv´any verzi´onak felel meg. A jelenleg haszn´alatos szabv´any az 1.0:

<?xml version="1.0" ?>

* + - Az XML dokumentumban a tartalom elemek egym´asba a´gyazhat´ok:

egy k´ettagu´ tartalom elem mag´aba foglalhat egy u´jabb tartalom elemet:

<fo\_elem>

<al\_elem\_1> ... </al\_elem\_1>

<al\_elem\_2/>

</fo\_elem>

A tartalmazott elemet gyerek elemnek, a tartalmaz´ot szu¨l˝o elemnek nevezzu¨k. A szu¨l˝o ´es annak minden szu¨l˝oje az elem o˝seinek halmaz´at alkotja.

Minden elemnek csak egy szu¨l˝oje lehet, nem lehet egy elem k´et ku¨l¨onb¨oz˝o szl˝oben is gyermek. Ezen megk¨ot´esb˝ol ered˝oen a tartalom elemek egy hierarchi´at alkotnak.

*•*

A dokumentum gy¨ok´er eleme az a tartalom elem, amelyhez nem l´etezik szu¨l˝o elem. Az XML dokumentum csak egy gy¨ok´er elemet tartal- mazhat. Az al´abbi p´elda egy hosszabb XML mint´at ad meg:

*•*

<?xml version="vv" ?>

<!-- ez egy minta KL:2007.07.01 -->

<k¨onyvek>

<k¨onyv ikod="1">

<cim> XML alapok </cim>

<ar pnem="Ft"> 5600 </ar>

</k¨onyv>

<k¨onyv ikod="2">

<cim> Adatkezel´esi alapok </cim>

<ev>2003</ev>

<olvasoterem/>

</k¨onyv>

</k¨onyvek>

A XML dokumentumban bizonyos karakterek szerar´ator szerepet t¨oltenek be, Ilyen karakter p´eld´aul a ’*<*’ jel. Ez a karakter ebben az alak- ban teh´at nem fordulhat el˝o az elemnevekben, a norm´al tartalom le´ır´o sz¨ovegben. A foglalt karakterek halmaza az al´abbi karakterekre terjed ki:

*•*

**–** *<*

**–** *>*

**–** &

**–** *j*

**–** ”

Az XML dokumentumba beilleszthet˝o olyan tetsz˝oleges sz¨ovegr´esz, melyet a XML feldolgoz´o teljes eg´esz´eben tartalom sz¨ovegnek tekint, nem keres benne szepar´ator elemeket. A sz¨ovegr´esz emiatt tartalmazhat foglalt karatereket is. A sz¨ovegr´esz megad´asa a CDATA elemmel t¨ort´enik:

*•*

<![CDATA[ sz¨oveg ]]>

A tartalmazott sz¨ovegre vonatoz´o egyedu¨li megk¨ot´es, hogy ne forduljon el˝o benne a ’]] *>*’ jelsorozat.

A foglalt karakterek tartalom sz¨ovegbe val´o beillszet´es´ere egyed-szimb´olumokat is lehet haszn´alni. A szimb´olumot egy &*kod*; form´atumban lehet megadni,

*•*

ahol a k´od ´ert´eke lehet sz´amk´od vagy sz¨oveg. A foglalt karakterekhez tartoz´o szimb´olumok:

**–** *<* : &*lt*;

**–** *>* : &*gt*;

**–** & : &*amp*;

* *j* : &*apos*;
* ” : &*quot*;

# XML n´evterek

Az XML dokuemntumban felhaszn´alt elemnevek az alapadatok jelent´es´ere utalnak. Mivel rendszerint a term´eszetes nyelv szavait haszn´alj´ak fel az el- emnevek megad´as´an´al, r¨ogt¨on felmeru¨l a t¨obb´ertelmu˝s´eg probl´em´aja. Hiszen egy fogalom, mint p´eld´aul a ’jegy’, nagyon sok ´ertelemben haszn´alhat´o:

´erdemjegy, utaz´asi jegy, jellemz˝o. Emiatt egy adott elemn´ev ku¨l¨onb¨oz˝o alkalmaz´asi k¨ogyezetben, kontexusban, m´as ´es m´as szemantikai jelent´essel b´ırhat.

Az XML n´evt´er egy t´emateru¨letet azonos´ıt. Az elemeket t´ars´ıtani lehet a n´evterekkel, minden elemhez egy n´evt´er rendelhet˝o. A n´evt´er neve maga is tetsz˝oleges n´ev lehet. L´athat´o azonban, hogy ez esetben a n´evu¨tk¨oz´esi probl´em´at nem oldottuk meg, csak ´attoltuk egy m´asik szintre. Hogy ne kell- jen magasabb szintu˝ n´evterek l´ancolat´at l´etrehozni, a XML gyakorlatban a n´evterek elnevez´es´en´el az egyedis´egere t¨orekednek. Hogy min´el kisebb es´elye legyen az ütközésnek, URL formátumot szokás használni. A URL ´allom´any specifikus r´esze rendszerint a feldolgoz´o program t´argyk¨or´ere utal. A n´evt´erhez hozz´arendelt URL ´ert´eknek nem kell l´etez˝o objektumra mutatnia. N´eh´any k¨ozismert n´evt´er azonos´ıt´o:

* + - <http://www.w3.org/2000/10/XMLSchema>: W3C XMLSchema n´evtere
    - urn:schemas-microsoft-com:xml-data : az MS XML s´ema n´evtere
    - <http://www.w3.org/1999/XSL/Transform>: az XSL feldolgoz´o n´evtere
    - <http://www.w3.org/1999/xlink>: az XLink feldolgoz´o n´evtere

Mint a p´eld´ab´ol is l´athat´o, hogy az URL form´atumu´ n´evt´er azonos´ıt´ok

igen tereb´elyesek az egyedis´eg biztos´ıt´asa v´egett. Ha ezen azonos´ıt´okat adn´ank meg az egyes tartalom elemekn´el, akkor bizony igencsak elnyomn´a a t¨obbi adatot a n´evt´er megad´asa. Hogy ne uralkodjon el ennyire a n´evt´er kijel¨ol´es, az XML szabv´any lehet˝ov´e teszi a n´evterekhez egy lok´alis aliasn´ev, r¨ovid´ıt´es hozz´arendel´es´et. Ezt¨ovet˝oen elegend˝o csak a r¨ovidebb n´evt´eralias haszn´alata. A szabv´any emelett m´eg azt is lehet˝ov´e teszi, hogy alap´ertelmez´esi n´everet hozzunk l´etre, amikor nem kell m´eg a n´evt´eraliast sem megadni. A n´evterek

megad´asa elemjellemz˝on keresztu¨l t¨ort´enik. A vonatkoz´o elemjellemz˝o azonos´ıt´o neve: ’xmlns’. A n´evt´er megad´as form´atuma:

<elemn´ev ... xmlns:alias = URL ...>

ahol a megadott alias haszn´alhat´o a tov´abbiakban a kapcsol´od´o URL helyet. Ha az ’:alias’ r´esz elmarad, akkor a megadott URL lesz az elemek alap´ertelmez´esi n´evtere. Az alap´ertelemz´esi n´evt´er megad´as´anak szintak- tik´aja

<elemn´ev ... xmlns = URL ...>

Egy elemn´el t¨obb n´evt´er kijel¨ol´es is ´elhet. A n´evt´er hozz´arendel´ese egy tartalomelemhez a

<alias:elemn´ev ...>

form´atumban t¨ort´enik. Ha k´ettagu´ az elem, akkor a z´ar´o tagn´al is szere- peltetni kell a nyit´o tagban megadott alias nevet. Az alap´ertelemz´esi n´evt´er l´etez´ese eset´en a norm´al

<elemn´ev ...>

elemmegad´as az elemet az alap´ertelmez´esi n´evt´erhez rendeli. Az alap´ertelmez´esi n´evt´er haszn´alat´anak egyik saj´atoss´aga, hogy az alap´ertelemz´esi n´evt´er csak

az elemekre vonatkozik, a kapcsol´od´o elemjellem˝okre m´ar nem. Az aliasn´ev elemjellemz˝oh¨oz t¨ort´en˝o rendel´esekor a elemjellemz˝o neve el˝ott kell megadni az alias azonos´ıt´ot:

<elemn´ev alias:jellemzo\_nev=’ertek’ ...>

Az aliasn´ev ´erv´enyess´egi k¨ore azon elemre ´es annak befoglalt elemeire ter- jed ki, amelyben defini´alt´ak.

Egy bels˝o elemben a k´ıvu¨l ´ertelmezett aliasn´ev felu¨l´ırhat´o egy u´j ´ertelmez´essel.

<ktar:konyvek [xmlns:kt=’http://iit.uni-miskolc.hu/kovacs/konyvtar’](http://iit.uni-miskolc.hu/kovacs/konyvtar) [xmlns:vir=’http://uni-miskolc.hu/nyilvanartas’](http://uni-miskolc.hu/nyilvanartas) [xmlns=’http://iit.uni-miskolc.h](http://iit.uni-miskolc.hu/root)u/root’>

<ktar:konyv id=’1’>

<ktar:ISBN> I11233 </ktar:ISBN>

<vir:olvaso> UZI234 <uid> olg45 </uid> </vir:olvaso>

<vir:nkod [xmlns:vir=’http://uni-miskolc.hu/neptun’>](http://uni-miskolc.hu/neptun) OPUZT7 </vir:nkod>

</ktar:konyv>

<ktar:konyv id=’3’>

<ktar:ISBN> TF7671 </ktar:ISBN>

<vir:olvaso> JII781 <uid> z78hz </uid> </vir:olvaso>

<vir:nkod [xmlns:vir=’http://uni-miskolc.hu/neptun’>](http://uni-miskolc.hu/neptun) LOP86J </vir:nkod>

</ktar:konyv>

</ktar:konyvek>