

## 2. laboratorijska vježba: XML, DTD i XSL

### Cilj vježbe

Upoznavanje s jezikom XML i strukturiranjem informacija. Validacija dobre oblikovanosti i ispravnosti dokumenta pomoću DTD-a. Upoznavanje s transformacijama XSL. Uočavanje razlika jezika XHTML i HTML.

### Priprema za vježbu

Proučiti – XML:

- osnove jezika XML
- XML zaglavlje (`<?xml..?>`)
- korištenje elemenata i atributa
- definicije dobre oblikovanosti i valjanosti
- načine uključivanja DTD-a u XML dokumente (`<!DOCTYPE ...>`).

Proučiti – DTD:

- definiranje elemenata u DTD-u – tako da sadrže neparsirani tekst, druge elemente, itd.
- kod uključivanja drugih elemenata obratiti pažnju na BNF notaciju prilikom specifikacije sadržaja (`*`, `+`, `?`, `|`).
- način definiranja:
  - atributa proizvoljne vrijednosti
  - atributa koji mogu poprimiti samo vrijednosti iz određenog skupa dozvoljenih vrijednosti
  - atributa koji su obavezni, te koji se podrazumijevaju ili su fiksni (`#REQUIRED`, `#IMPLIED`, `#FIXED`)

Proučiti – XSL:

- osnove jezika XSL
- uključivanje XSL stilova u XML dokumente
- definiranje predložaka u XSL-u
- primjena predložaka
- postavljanje prostora imena kako bi se izbjegao konflikt imenovanja
- iteriranje po skupu elemenata korištenjem XSL oznaka *for-each* i sličnih
- provjeru uvjeta XSL oznakom *if*
- dohvaćanje vrijednosti elemenata iz XML-a XSL-om
- postavljanje atributa u izlaznom dokumentu XSL-om

### Zadatak za vježbu

Potrebno je **izraditi DTD** za provjeru odgovarajućeg strukturiranog zapisa podataka u XML dokumentu te **izraditi testni XML** dokument koji služi za ispitivanje primjerenosti izrađenog DTD dokumenta pretpostavljenoj strukturi podataka. Struktura podataka zadana je **inačicom zadatka** koji vam je pridružen. Dokument mora biti **dobro oblikovan i valjan**. DTD dokument mora sadržavati sve navedene podatke u pridruženoj inačici. Za izrađeni XML

dokument izraditi **XSL transformaciju** koja će prikazati podatke iz XML dokumenta kao **tablicu uklopljenu u izgled stranice** pripremljen u prethodnoj vježbi.

Stupci u dokumentu o inačicama i njihov utjecaj na strukturu XML dokumenta:

- **HIJERARHIJSKA RAZINA** – označava razinu hijerarhije podatka u odnosu na korijenski element (atributi se smatraju pod-razinom elementa). Dobro oblikovani XML dokument mora imati samo jedan korijenski element.
- **BROJNOST** – označava može li se element pojaviti samo jednom (1) ili više puta (N) unutar istog nadređenog elementa.
- **ELEMENT ili ATRIBUT** – određuje način zapisa podatka kao elementa ili atributa.
- **OBAVEZNOST POSTOJANJA** – označava obavezno ili opcionalno postojanje elementa ili atributa
- **VRIJEDNOSTI** – označava sadržava li element ili atribut informaciju:
  - **DA** – sadržava standardno upisanu tekstualnu vrijednost (#PCDATA ili CDATA)
  - **SKUP** – sadržava podatak isključivo iz skupa dozvoljenih vrijednosti (enumerirani tip atributa)
  - **NEMA** – element ne sadržava znakovnu vrijednost (#PCDATA), no može sadržavati pod-elemente i attribute
- **PRIMJER SLOBODNOG UPISA VRIJEDNOSTI ili SKUP DOZVOLJENIH VRIJEDNOSTI** – ako je definiran, skup dozvoljenih vrijednosti sadrži popis dozvoljenih vrijednosti atributa.

**5 elemenata ili atributa**, koji nisu definirani u dokumentu o inačicama, potrebno je **definirati proizvoljno**, no da s preostalim elementima i atributima čine logičku cjelinu. Provjeravat će se smislenost i logičnost tih elemenata i atributa u kontekstu smisla cijele inačice, a uspoređivat će se i potencijalno prepisivanje takvih proizvoljnih tipova unutar rješenja iste inačice, ali i u rješenjima različitih inačica.

Testni XML dokument **mora sadržavati barem 5 podataka prve razine** (npr. 5 osoba u imeniku i sl.), biti kodiran **kodnom stranicom UTF-8** i **mora** u odnosu na pridruženi DTD dokument **biti valjan**. Provjeru valjanosti možete vršiti *pouzdanim* alatom po vašem izboru.

Potrebno je izraditi **XSL transformaciju** za prikaz podataka **u HTML tablici u XHTML dokumentu**. Rezultat XSL transformacije će se **provjeriti** pomoću **Web preglednika** (Internet Explorer, Firefox, ...). Stranica koju izrađujete se naravno razlikuje ovisno o inačici zadatka koji ste dobili. Dodatno, Web stranicima iz 1. lab. vježbe treba unutar izbornika navigacije postaviti link na novoizrađenu stranicu, odnosno točnije, na datoteku *.xml* koja sadrži poziv transformacije *.xsl*. Novoizrađena Web stranica treba slijediti dizajn Web stranica iz 1. lab. vježbe. Unutar XSL datoteke, osim same transformacije potrebno je uključiti ostale dijelove XHTML stranice, poput kodne stranice, CSS-a, navigacije, itd.

Prilikom izrade tablice, uzmite u obzir **preglednost** prikaza podataka. U slučaju da je prikazivanje svih podataka nepregledno (ne stanu svi stupci na zaslon), izaberite **podskup** podataka za prikazivanje i prikažite **samo** njih.

Tablica mora imati **zaglavni redak** s nazivom pojedinih stupaca.

## Predaja vježbe

Predaja vježbe podrazumijeva:

1. predočavanje datoteke s DTD-om, objašnjenje pojedinih rješenja u DTD-u
2. predočavanje datoteke s testnim XML-om,
3. dokazivanje valjanosti testnog XML dokumenta (alatom po vašem izboru)
4. promjene u DTD-u i testnom XML dokumentu te ponovnu provjeru valjanosti testnog dokumenta
5. predočavanje XSL transformacije testnog XML dokumenta u Web pregledniku
6. eventualne promjene u testnom XML-u ili definiciji XSL transformacije te ponovno provođenje transformacije

Vježba se predaje na lokalnom računalu. Datoteke koje treba predati i prikazati su sljedećih naziva:

- podaci.xml
- podaci.dtd
- prikaz.xsl

U ovoj će vježbi biti potrebne i druge datoteke (iz vježbe 1.). Na poslužitelj Moodle treba postaviti ZIP arhivu sa svim datotekama potrebnima za izvođenje.

### **Provjera valjanosti XML dokumenta**

Izbor alata kojim se vrši provjera valjanosti je slobodan. Uvjet je da je korištenje tog alata legalno. Predlažemo tri alata/načina provođenja provjere:

1. **Na računalu pinus**, u datoteci `~marvin/or/validate.sh` nalazi se gotova skripta `“validate.sh”`. Kopirajte datoteku u vaš direktorij naredbom:

```
$> cp ~marvin/or/validate.sh .
```

Napomena: Obratite pažnju na razmak i točku nakon imena! Oznaka `"$>"` ili slična, označava spremnost za rad, te je ne upisujete!

Skriptu pokrećite iz trenutnog direktorija:

```
$> ./validate <ime_xml_datoteke>
```

XML datoteka mora se nalaziti u trenutnom direktoriju. Ako niste naveli putanju do DTD datoteke u `<!DOCTYPE ...>`, tada i ona mora biti u trenutnom direktoriju.

2. Na računalu s Windows operacijskim sustavom možete koristiti jednostavan i besplatan alat **AltovaXML**. Instalacijsku datoteku i upute preuzmite s <http://www.altova.com/altovaxml.html>
3. Na 49. slajdu predavanja o DTD-u nalazi se izvorni kod i primjer korištenja vrlo jednostavne **PHP skripte** za provjeru valjanosti XML dokumenta. Za korištenje ove skripte, trebate imati instaliran PHP (<http://www.php.net>, verzija 5.2.5) za operacijski sustav koji koristite na računalu.

## Transformacija XML dokumenta

Transformacija dokumenta treba se provoditi korištenjem jednog od Web preglednika. Kako bi preglednik znao provesti XSL transformaciju XML dokumenta, u XML dokumentu je to potrebno označiti:

```
<?xml-stylesheet type="text/xsl" href="datoteka_s_opisom_transformacije.xsl"?>
```

XSL datoteka je također valjani XML dokument čiji je korijenski element *stylesheet* iz prostora imena *xsl*:

```
<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform" version="1.0"> ...  
</xsl:stylesheet>
```

Postavljanje tipa izlaznog dokumenta (DOCTYPE) izvodi se sa:

```
<xsl:output method="xml" indent="yes" doctype-system="http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd" doctype-public="-//W3C//DTD XHTML 1.0 Strict//EN"/>
```

Primjena predloška na pojedini element pokreće se sa:

```
<xsl:template match="/"> ... </xsl:template>
```

...gdje je izraz u atributu *match* odabir elementa na koji se predložak primjenjuje. "/" označava sve elemente.

Imenički prostor oznaka HTML-a može se postaviti na elementu *<html>*:

```
<html xmlns="http://www.w3.org/1999/xhtml">
```

Iteriranje po podelementima nekog elementa izvodi se oznakom *for-each* koja je slična *for* petlji u programskim jezicima:

```
<xsl:for-each select="element"> ... </xsl:for-each>
```

Dohvaćanje sadržaja elementa XML-a izvodi se oznakom *value-of*:

```
<xsl:value-of select="naziv" />
```

Dodavanje atributa elementu izlaznog dokumenta potrebno je posebno navesti. Primjer dodavanja atributa *href* elementu *a* i postavljanja sadržaja za slanje maila:

```
<a>  
  <xsl:attribute name="href">  
    mailto:<xsl:value-of select="ime" />.<xsl:value-of select="prezime" />@fer.hr  
  </xsl:attribute>  
  Elektronička pošta za: <xsl:value-of select="ime" /> <xsl:value-of select="prezime" />  
</a>
```

## Ispitno gradivo vježbe

Ispitno gradivo uključuje sve navedeno u pripremi za vježbu, te detaljno razumijevanje napisanog rješenja i snalaženja u prepravcima istog.

### Primjeri pitanja:

1. Što je XML, a što DTD?
2. Što znače pojedine stavke u zaglavlju XML dokumenta?
3. Kako je sve moguće uključiti DTD u XML dokument?

4. Što znači dobro oblikovan, a što ispravan valjan?
5. Kako se u DTD-u deklariraju elementi, a kako atributi elemenata?
6. Kako se pišu komentari u DTD-u i XML-u?
7. Kako odlučiti da li bi nešto trebalo biti atribut ili element?
8. Kako se XML dokumentu pridružuje XSL predložak stila?
9. Kako se XSL-om može stvoriti popis svih zapisa (podataka) odjednom?
10. Kako se XSL-om može stvoriti popis svih elemenata istog roditelja?
11. Kako se XSL-om može stvoriti popis svih elemenata čiji atribut ima jednaku vrijednost?
12. Kako bi drugom bojom označili elemente posebnih značajki (npr. proizvode s niskom cijenom)?

## Linkovi i literatura

- W3Schools XML Tutorial - <http://www.w3schools.com/xml/>
- W3Schools XSLT Tutorial - <http://www.w3schools.com/xsl/>
- Norman Walsh – XSL Tutorial - <http://nwalsh.com/docs/tutorials/xsl/>
- Altova XML - <http://www.altova.com/altovaxml.html>