

XML Schema (Vežbe 3)

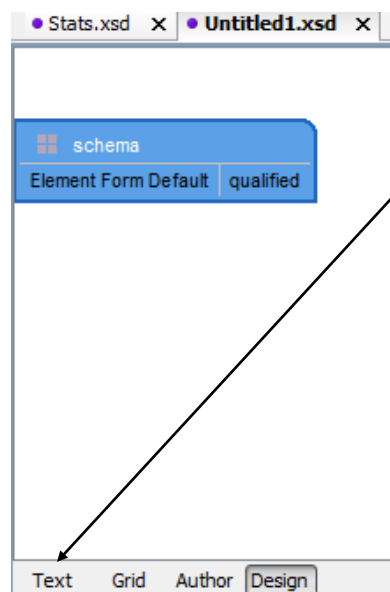
Materijali

Šta nedostaje DTD-u?

- **Ne-XML sintaksa**
- **Slaba podrška za namespace tj. prostore imena**
 - Svako može da definiše sopstvenu XML gramatiku, pa ako želimo da koristimo različite gramatike može doći do poklapanja imena. Da bismo to izbegli koristimo namespace. Namespace-ovi omogućavaju jedinstvenost naziva XML elemenata dodavanjem prefiksa.
- **Nema tipizacije podataka**
 - Naročito za sadržaj elemenata. Npr. nije moguće definisati datumska ili numerička polja da budu u odgovarajućem formatu (npr. dd.MM.yyyy).
- **Ograničena proširivost**
 - Reuse tipova podataka nije moguć.
- **Ograničene mogućnosti opisa strukture podataka**
 - Ne može se nametnuti broj podelemenata bez nametanja redosleda. Takođe, ne može se nametnuti redosled i broj podelemenata kada se koristi mešani sadržaj.

Formiranje XML Schema-e

- **Kreirati novu šemu u Oxygen-u** (File > New > XML Schema) (slika 1):



- XML šema predstavlja još jedan XML dokument čiji je korenski element "schema".
- Kreiranjem XML šeme u Oxygen-u se otvara grafički prikaz, za prebacivanje u tekstualni režim koriste se tabulatori koji se nalaze na dnu grafičkog prikaza.

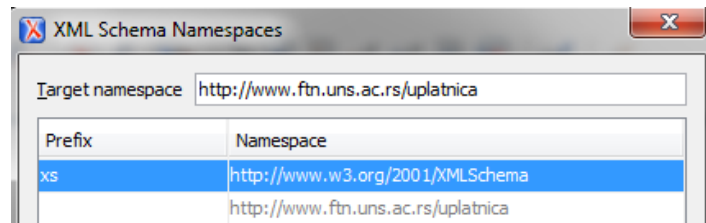
```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema" elementFormDefault="qualified"> {
3
```

- Šema rečnik se po defaultu nalazi u **http://www.w3.org/2001/XMLSchema** namespace-u, a u našoj šemi je on kvalifikovanim prefiksom **xs**. Pošto se svi elementi za konstrukciju šeme nalaze u datom namespace-u, to ujedno znači da će prilikom referenciranja njihova imena biti kvalifikovana odgovarajućim prefiksom, npr. xs:element, xs:sequence, ...

Slika 1: Kreiranje nove XML šeme

- **Dodeliti šemi target namespace**

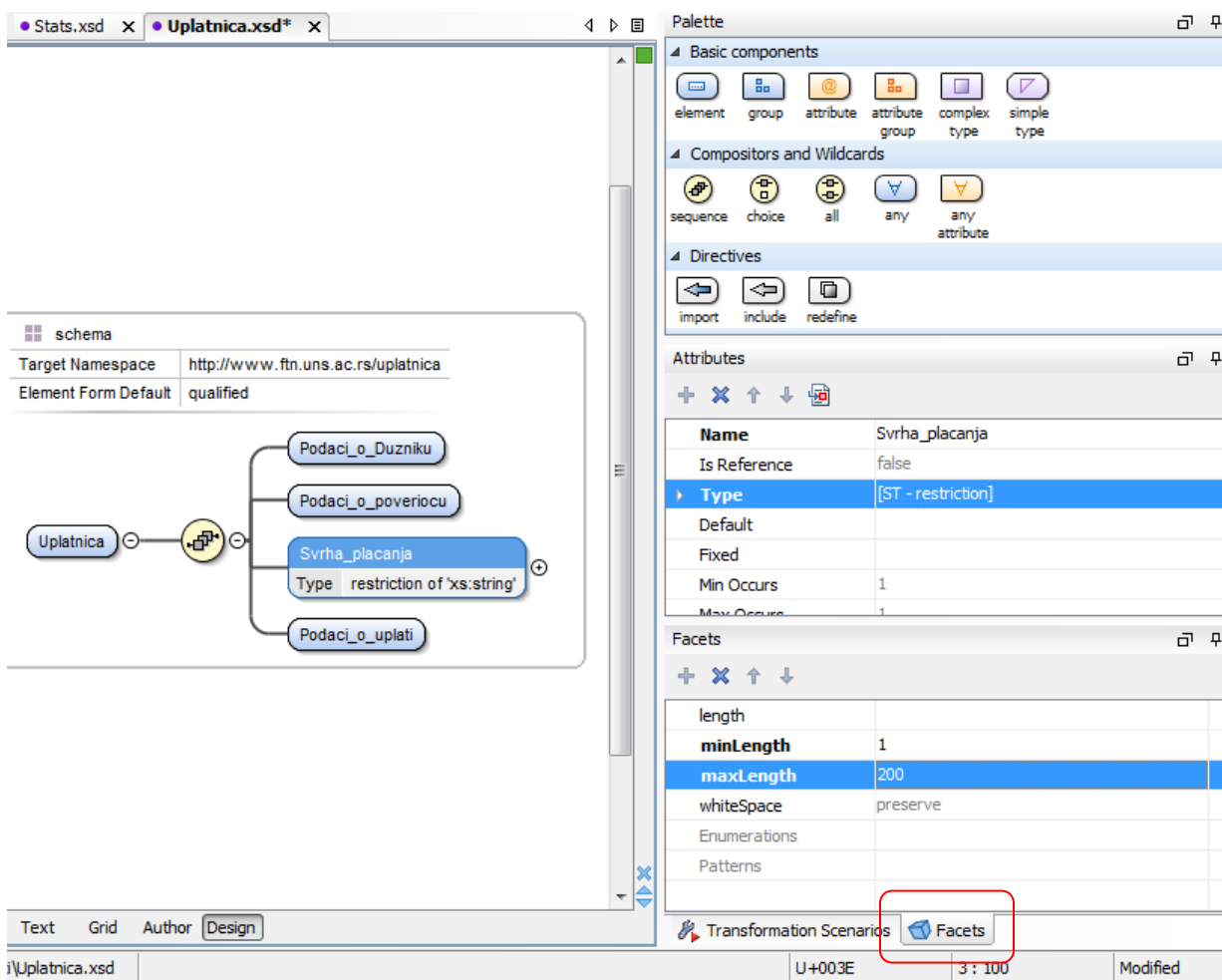
- Dvoklik na "schema" element grafičkog prikaza otvara sledeći prozor (slika 2):



Slika 2: Postavljanje target namespace-a

- Ovim smo rekli da se naš rečnik nalazi u <http://www.ftn.uns.ac.rs/uplatnica> targetNamespace-u što ujedno predstavlja prostor imena, jedinstveni identifikator XML gramatike tj. identifikator svih XML elemenata koji će biti kreirani u datoj šemi.
 - Nismo definisali prefiks za target namespace jer se u konkretnom slučaju default namespace poklapa sa target namespace-om.
 - Sadržaj atributa `xmlns="http://www.ftn.uns.ac.rs/uplatnica"` predstavlja mapiranje na default namespace dokumenta. Elementi koji se navode bez prefiksa se traže u ovom namespace-u.
- **Dodeliti prefiks target namespace-u**
 - Dvoklikom na tekstualno polje "Prefix" dijaloga sa prethodne slike dodeliti prefiks (npr. "upl") target namespace-u.
- **Napraviti novi globalni element "Uplatnica"**
 - Desnim klikom na grafički prikaz odabrati **New Global > Element** i dodeliti mu ime npr. **Uplatnica**.
 - Uplatnicu ćemo definisati kao sekvencu elemenata **Podaci_o_duzniku**, **Podaci_o_poveriocu**, **Svrha_placanja** i **Podaci_o_uplati**:
 - Desni klik na element **Uplatnica**, odabrati **Append Child > Sequence**.
 - Desni klik na sekvencu, odabrati **Append Child > Element** i dodeliti mu ime **Podaci_o_duzniku**.
 - Element **Podaci_o_duzniku** je **lokalni** element u okviru elementa **Uplatnica** što znači da ne može biti referenciran ni unutar, niti izvan svog parent-a, za razliku od **globalnih** elemenata.
 - Sličnu proceduru ponoviti za **Podaci_o_poveriocu**, **Svrha_placanja** i **Podaci_o_uplati**.
 - **Napomena:** Pored sekvence (kod koje je važan redosled elemenata) možemo odabrati i izbor (Choice) kao i neuređen skup (All) kod koga redosled navođenja elemenata nije bitan za validaciju.
 - Za svaki od elemenata se može definisati i broj pojavljivanja datog elementa u sekvenci postavljanjem kardinaliteta (**Min Occurs** i **Max Occurs** u okviru **Attributes** panela, desno od grafičkog prikaza), podrazumevane vrednosti su Min Occurs: 1, Max Occurs: 1.
- Elementu **"Svrha_placanja"** dodeliti tip string

- U okviru **Attributes** panela, postaviti **xs:string** za tip elementa **Svrha_placanja** (slika 3) dvoklikom na stavku **Type**. Prefiks **xs** se koristi jer je tip **string** definisan u XML Schema namespace-u, koji je kvalifikovan prefiksom **xs** (tačka 1).



Slika 3: Postavljanje prostog tipa

- Odabirom kartice **Facets** (uokvireno crvenim na slici 3), nad već definisanim tipom moguće je postavljati ograničenja (restrikcije). Ovde je minimalna dužina stringa (koji će biti sadržan elementom **Svrha_placanja**) postavljena na 1, a maksimalna na 200.
- **Definisati novi globalni element "Adresa"**
 - Specificirati adresu kao sekvencu elemenata: **Mesto**, **Postanski_broj**, **Ulica** i **Broj**.
 - Elementima **Mesto** i **Ulica** dodeliti tip **string**.
 - Elementu **Postanski_broj** dodeliti tip **int** (takođe definisan u okviru XML Scheme). Staviti ograničenje da najmanji poštanski broj koji je moguće uneti bude 11000 (ograničenje **minInclusive**), a najveći 40000 (ograničenje **maxInclusive**)
 - Element **Broj** postaviti da mora sadržati pozitivan broj (tip **int** uz ograničenje ili **positiveInteger**)
- **Definisati novi kompleksni tip "Tlice"**
 - Kompleksni tipovi, za razliku od prostih, mogu imati podelemente i atribute.

- Za definisanje kompleksnog tipa, desni klik na grafički prikaz **New Global > Complex Type**.
- Novom kompleksnom tipu dodeliti naziv **TLice** označiti ga apstraktnim (postaviti **Abstract** na **true** u **Attributes** panelu).
- **Tip "TLice" definisati kao sekvencu elemenata**
 - U sekvencu dodati referencu na globalni element **Adresa**. Postupak dodavanja je kao kod lokalnog elementa, voditi računa o prefiksima (odabrati **upl:Adresa**).
 - Primetiti da je definisani element **Adresa** kvalifikovan prefiksom **upl** – prefiks koji označava target namespace, tj. prostor imena naše šeme (svi elementi koje definišemo u okviru ove šeme pripadaće definisanom target namespace-u).
 - Pored reference na globalni element **Adresa**, dodati i novi lokalni element **Broj_racuna**.
- **Polimorfizam**
 - Elementima **Podaci_o_duzniku** i **Podaci_o_poveriocu** dodeliti tip **TLice**, na isti način kao što je izvršena dodela prostih tipova (npr. tip string). Ovaj tip će se nalaziti pod prefiksom **upl**, **upl:TLice**).
 - Tip **TLice** predstavlja apstrakciju jer omogućava modelovanje kako pravnih tako i fizičkih lica (videti naredne tačke).
- **Proširivanje postojećeg tipa**
 - Kreirati kompleksni tip **TPravno_lice**. **TPravno_lice** treba da nasleđuje **TLice** (atribut **Base Type** postaviti na **TLice** u **Attributes** panelu).
 - Pored adrese (koju **TPravno_lice** nasleđuje iz **TLice**), proširiti **TPravno_lice** dodavanjem sekvence elemenata, sa jednim elementom **Naziv** tipa **string**.
- **Definisati kompleksni tip "TFizicko_lice"**
 - Analogno kao i **TPravno_lice** definisati **TFizicko_lice**. **TFizicko_lice** će umesto **Naziva** imati **Ime** i **Prezime** kao lokalne elemente (tipa **xs:string**).
- **Definisanje prostog tipa**
 - Definisati prost tip za broj računa. Prosti tipovi ne mogu imati podelemente ili attribute, a možemo ih kreirati:
 - **restrikcijom** ugrađenih tipova (navođenje ograničenja na vrednost – min dužina, max dužina, enumeracija, ...)
 - **lista** (definisanje konačne liste mogućih vrednosti) – predstavlja listu elemenata nekog tipa koji će biti razdvojeni razmacima npr. <bingo>4 9 12</bingo>
 - **unija** (kombinovanje vrednosti različitih osnovnih tipova) – vrednost novog tipa može biti vrednost bilo kog tipa koji je član unije
 - Prosti tip kreirati desnim klikom na grafički prikaz **New Global > Simple type**.
 - Tip nazvati **TBroj_racuna**. Broj računa se zadaje u sledećem formatu: 3 cifre –13 cifara – 2 cifre (npr. 840-1234567890123-28). Ovo možemo specificirati restrikcijom tipa **string** postavljanjem **RegExp Pattern**-a:
 - Dodeliti prostom tipu **TBroj_racuna** tip **string**
 - Nakon dvoklika na **Patterns** u **Facetes** tabu (gde se postavljaju ograničenja) uneti sledeći šablon: \d{3}-\d{1,13}-\d{2}

- Pored **Pattern**-a, može se definisati i enumeracija (element uzima jednu od definisanih vrednosti), dodavanjem željenih vrednosti u **Enumeration** polje u okviru **Facets** kartice.
- **Dodeliti tip elementu "Broj_racuna"**
 - Tip elementa **Broj_racuna** postaviti da bude **TBroj_racuna** i dodeliti mu default vrednost 000-000000000000-00 (postaviti **Default** atribut u okviru **Attributes** panela).
 - Provera broja računa se vrši računanjem **checksum** cifre po modelu 97 na sledeći način:
 - 1) broj se pomnoži sa 100
 - 2) zatim se rezultat podeli sa 97
 - 3) a ostatak pri deljenju oduzme od 98
 - 4) rezultat je checksum od 2 cifre
 - Da li je ovu proveru moguće izvršiti u XML šemi?
- **Specificirati element "Podaci_o_uplati"**
 - Definisati kao sekvencu elemenata: **Sifra_placanja** (xs:int), **Iznos** (xs:int), **Podaci_o_zaduzenju** (xs:string) i **Podaci_o_odobrenju** (xs:string).
 - Svim navedenim lokalnim elementima (zasad) dodeliti proste tipove.
 - Svi navedeni elementi, kao i ostatak XML gramatike, će biti prošireni u okviru 1. zadatka.
- **Napraviti XML na osnovu date XML šeme**
 - Odabrati **File > New > XML Document** i kliknuti na **Customize**.
 - Postupak je sličan kao kod DTD-a, odabrati šemu i korenski element za XML.
 - Obratiti pažnju da se kao korenski element mogu uzeti samo definisani **globalni elementi**.
 - Za korenski element uzeti element **Uplatnica**, nakon generisanja XML popuniti na sledeći način:

```
<Uplatnica xmlns="http://www.ftn.uns.ac.rs/uplatnica"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.ftn.uns.ac.rs/uplatnica
Uplatnica.xsd">
```

```
<Podaci_o_Duzniku xsi:type="TPravno_lice">
  <Adresa>
    <Mesto>Novi Sad</Mesto>
    <Postanski_broj>21000</Postanski_broj>
    <Ulica>Trg Dositeja Obradovica</Ulica>
    <Broj>6</Broj>
  </Adresa>
  <Naziv>FTN</Naziv>
</Podaci_o_Duzniku>
```

```
<Podaci_o_poveriocu xsi:type="TFizicko_lice">
  <Adresa>
    <Mesto>Novi Sad</Mesto>
    <Postanski_broj>21000</Postanski_broj>
    <Ulica>Tu negde</Ulica>
    <Broj>1</Broj>
  </Adresa>
  <Ime>Petar</Ime>
  <Prezime>Petrovic</Prezime>
</Podaci_o_poveriocu>
```

Povezivanje XML dokumenta i šeme se vrši preko **schemaLocation** atributa u okviru korenskog elementa. U njemu se navodi target namespace šeme na osnovu koje se instancira XML, kao njena lokacija. Ovim se definiše koji od konkretnih tipova (koji nasleđuju **TLice**) se instancira.

```
<Svrha_placanja>Uplata po fakturi br. 16</Svrha_placanja>
<Podaci_o_uplati>
  <Sifra_placanja>153</Sifra_placanja>
  <Iznos>8000,00</Iznos>
  <Podaci_o_zaduzenju>
    840-1234567890123-28
  </Podaci_o_zaduzenju>
  <Podaci_o_odobrenju>
    370-1234567890123-72
  </Podaci_o_odobrenju>
</Podaci_o_uplati>
</Uplatnica>
```

Obratiti pažnju

- **schemaLocation** atribut se nalazi u **xsi** namespace-u (gde su definisani svi gradivni elementi same XML šeme) koji je definisan sa `xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"`
- default namespace u definisanom XML-u je `xmlns="http://www.ftn.uns.ac.rs/uplatnica"`
Obratiti pažnju da je to isti namespace kao target namespace šeme prema kojoj se pravi XML. Elementi default namespace-a se ne navode sa prefiksom, a pošto default namespace ukazuje na prostor imena uplatnice (Uplatnica.xsd), gradivni elementi ove šeme se u datom XML dokumentu ne moraju kvalifikovati prefiksom (npr. element **Uplatnica** se navodi bez prefiksa).