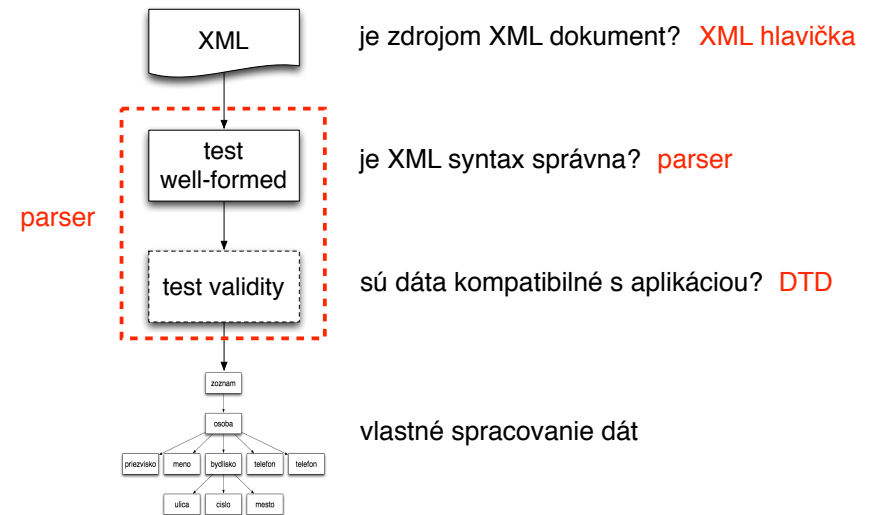


XML a DTD

Michal Kvasnica

Spracovanie XML súborov



DTD: sumarizácia

- DTD definuje, čo musí a čo nesmie obsahovať XML súbor
- DTD validuje správnosť daného XML (zmluva)
- Dva pohľady na DTD:
 - tvorca aplikácie: vytvára a publikuje DTD na to, aby iní mohli poskytnúť našej aplikácii kompatibilné dáta
 - tvorca dát: číta DTD a na základe neho vytvorí vhodný XML dokument

Ako prebieha DTD validácia?

```
<?xml version="1.0" encoding="UTF-8" ?>  
1: <!ELEMENT A (B*,C)>  
2: <!ELEMENT B (#PCDATA)>  
3: <!ELEMENT C (D*,E?)>  
4: <!ELEMENT D (#PCDATA)>  
5: <!ELEMENT E (F+ | #PCDATA)>  
6: <!ELEMENT F (#PCDATA)>
```

```
<A>  
</A>
```

Ako prebieha DTD validácia?

```
<?xml version="1.0" encoding="UTF-8" ?>
1: <!ELEMENT A (B*,C)>
2: <!ELEMENT B (#PCDATA)>
3: <!ELEMENT C (D*,E?)>
4: <!ELEMENT D (#PCDATA)>
5: <!ELEMENT E (F+ | #PCDATA)>
6: <!ELEMENT F (#PCDATA)>
```

```
<A>
  <B>123.4</B>
</A>
```

Ako prebieha DTD validácia?

```
<?xml version="1.0" encoding="UTF-8" ?>
1: <!ELEMENT A (B*,C)>
2: <!ELEMENT B (#PCDATA)>
3: <!ELEMENT C (D*,E?)>
4: <!ELEMENT D (#PCDATA)>
5: <!ELEMENT E (F+ | #PCDATA)>
6: <!ELEMENT F (#PCDATA)>
```

```
<A>
  <B>pisomka</B>
  <C>
    <E>retazec</E>
  </C>
</A>
```

Ako prebieha DTD validácia?

```
<?xml version="1.0" encoding="UTF-8" ?>
1: <!ELEMENT A (B*,C)>
2: <!ELEMENT B (#PCDATA)>
3: <!ELEMENT C (D*,E?)>
4: <!ELEMENT D (#PCDATA)>
5: <!ELEMENT E (F+ | #PCDATA)>
6: <!ELEMENT F (#PCDATA)>
```

```
<A>
  <B></B>
  <C>
    <E>retazec</E>
    <E>retazec</E>
    <E>retazec</E>
  </C>
</A>
```

Ako prebieha DTD validácia?

```
<?xml version="1.0" encoding="UTF-8" ?>
1: <!ELEMENT A (B*,C)>
2: <!ELEMENT B (#PCDATA)>
3: <!ELEMENT C (D*,E?)>
4: <!ELEMENT D (#PCDATA)>
5: <!ELEMENT E (F+ | #PCDATA)>
6: <!ELEMENT F (#PCDATA)>
```

```
<A>
  <B>pisomka</B>
  <C>
    <D>retazec</D>
    <D>retazec</D>
  </C>
</A>
```

Ako prebieha DTD validácia?

```
<?xml version="1.0" encoding="UTF-8" ?>
1: <!ELEMENT A (B*,C)>
2: <!ELEMENT B (#PCDATA)>
3: <!ELEMENT C (D*,E?)>
4: <!ELEMENT D (#PCDATA)>
5: <!ELEMENT E (F+ | #PCDATA)>
6: <!ELEMENT F (#PCDATA)>
```

```
<A>
  <B>text</B>
  <C>
    <D>retazec</D>
    <D>retazec</D>
    <E>text</E>
  </C>
</A>
```

Ako prebieha DTD validácia?

```
<?xml version="1.0" encoding="UTF-8" ?>
1: <!ELEMENT A (B*,C)>
2: <!ELEMENT B (#PCDATA)>
3: <!ELEMENT C (D*,E?)>
4: <!ELEMENT D (#PCDATA)>
5: <!ELEMENT E (F+ | #PCDATA)>
6: <!ELEMENT F (#PCDATA)>
```

```
<A>
  <B>text</B>
  <C>
    <E>
      <F>text</F>
      <F>text</F>
      <F>text</F>
    </E>
  </C>
</A>
```

Príklad DTD: prechodová charakteristika

- Požiadavky na DTD:
 - hlavný koreňový element: <prech>
 - obsahuje jeden element <system> a jeden element <matica>
 - systém je buď "prenos" alebo "stavovy_opis"
 - pre prenos sa definuje "citateľ" a "menovateľ" ako "matica"
 - pre stavový opis sa definujú matice "A", "B", "C", "D"

- Požiadavky na DTD:

- hlavný koreňový element: <prech>
- obsahuje jeden element <system> a jeden element <matica>
- systém je buď "prenos" alebo "stavovy_opis"
- pre prenos sa definuje "citateľ" a "menovateľ" ako "matica"
- pre stavový opis sa definujú matice "A", "B", "C", "D"

```
<?xml version="1.0" encoding="UTF-8" ?>
<!ELEMENT prech (system,matica)>
<!ELEMENT system (prenos | stavovy_opis)>
<!ELEMENT prenos (citateľ,menovateľ)>
<!ELEMENT citateľ (matica)>
<!ELEMENT menovateľ (matica)>
<!ELEMENT stavovy_opis (A,B,C,D)>
<!ELEMENT A (matica)>
<!ELEMENT B (matica)>
<!ELEMENT C (matica)>
<!ELEMENT D (matica)>
<!ELEMENT matica (riadok*)>
<!ELEMENT riadok (stlpec*)>
<!ELEMENT stlpec (#PCDATA)>
```

Príklad DTD: databáza kníh a vydavateľov

- Požiadavky na DTD:
 - hlavný koreňový element: <db>
 - môže obsahovať 0, 1 alebo viac kníh a 0, 1 alebo viac vydavateľov
 - kniha má elementy: titulok, rok publikovania, vydavateľstvo (všetky povinné)
 - vydavateľ je buď "meno" a "priezvisko", alebo záznam v tvare "firma"
 - "firma" povinne obsahuje názov, sídlo a alternatívne aj telefónny kontakt (všetko ako reťazce)

- Požiadavky na DTD:
 - hlavný koreňový element: <db>
 - môže obsahovať 0, 1 alebo viac kníh a 0, 1 alebo viac vydavateľov
 - kniha má elementy: titulok, rok publikovania, vydavateľstvo (všetky povinné)
 - vydavateľ je buď "meno" a "priezvisko", alebo záznam v tvare "firma"
 - "firma" povinne obsahuje názov, sídlo a alternatívne aj telefónny kontakt (reťazce)

```
<?xml version="1.0" encoding="UTF-8" ?>
<!ELEMENT db (kniha*,vydavatel*)>
<!ELEMENT kniha (titulka,rok,vydavatelstvo)>
<!ELEMENT titulka (#PCDATA)>
<!ELEMENT rok (#PCDATA)>
<!ELEMENT vydavatelstvo (#PCDATA)>
<!ELEMENT vydavatel ((meno,priezvisko)|firma)>
<!ELEMENT meno (#PCDATA)>
<!ELEMENT priezvisko (#PCDATA)>
<!ELEMENT firma (nazov,sidlo,telefon?)>
<!ELEMENT nazov (#PCDATA)>
<!ELEMENT sidlo (#PCDATA)>
<!ELEMENT telefon (#PCDATA)>
```

```
<?xml version="1.0" encoding="UTF-8" ?>
<!ELEMENT db (kniha*,vydavatel*)>
<!ELEMENT kniha (titulka,rok,vydavatelstvo)>
<!ELEMENT titulka (#PCDATA)>
<!ELEMENT rok (#PCDATA)>
<!ELEMENT vydavatelstvo (#PCDATA)>
<!ELEMENT vydavatel ((meno,priezvisko)|firma)>
<!ELEMENT meno (#PCDATA)>
<!ELEMENT priezvisko (#PCDATA)>
<!ELEMENT firma (nazov,sidlo,telefon?)>
<!ELEMENT nazov (#PCDATA)>
<!ELEMENT sidlo (#PCDATA)>
<!ELEMENT telefon (#PCDATA)>

<?xml version="1.0" encoding="UTF-8" ?>
<db>
  <vydavatel>
    <meno>Jan</meno>
    <priezvisko>Knihovy</priezvisko>
  </vydavatel>
  <vydavatel>
    <firma>
      <nazov>Knihy noveho milenia</nazov>
      <sidlo>Knizna 24, Bratislava</sidlo>
    </firma>
  </vydavatel>
</db>
```

Asociácia DTD k XML dokumentu

privátne DTD

koreňový element

cesta k DTD súboru

```
<?xml version="1.0" ?>
<!DOCTYPE db SYSTEM "databaza.dtd">
<db>
  <vydavatel>
    <meno>Jan</meno>
    <priezvisko>Knihovy</priezvisko>
  </vydavatel>
  <vydavatel>
    <firma>
      <nazov>Knihy noveho milenia</nazov>
      <sidlo>Knizna 24, Bratislava</sidlo>
    </firma>
  </vydavatel>
</db>
```

Privátne DTD určené pomocou URL

```
<?xml version="1.0" ?>
<!DOCTYPE db SYSTEM "http://kirp.sk/databaza.dtd">
<db>
  <vydavatel>
    <meno>Jan</meno>
    <priezvisko>Knihovy</priezvisko>
  </vydavatel>
  <vydavatel>
    <firma>
      <nazov>Knihy noveho milenia</nazov>
      <sidlo>Knizna 24, Bratislava</sidlo>
    </firma>
  </vydavatel>
</db>
```

Verejné DTD na lokálnom disku

verejný identifikátor

```
<?xml version="1.0" ?>
<!DOCTYPE db PUBLIC "knizna databaza" "databaza.dtd">
<db>
  <vydavatel>
    <meno>Jan</meno>
    <priezvisko>Knihovy</priezvisko>
  </vydavatel>
  <vydavatel>
    <firma>
      <nazov>Knihy noveho milenia</nazov>
      <sidlo>Knizna 24, Bratislava</sidlo>
    </firma>
  </vydavatel>
</db>
```

Verejné DTD pomocou URL

verejný identifikátor

```
<?xml version="1.0" ?>
<!DOCTYPE db PUBLIC "knizna databaza" "http://kirp.sk/databaza.dtd">
<db>
  <vydavatel>
    <meno>Jan</meno>
    <priezvisko>Knihovy</priezvisko>
  </vydavatel>
  <vydavatel>
    <firma>
      <nazov>Knihy noveho milenia</nazov>
      <sidlo>Knizna 24, Bratislava</sidlo>
    </firma>
  </vydavatel>
</db>
```

Vložené DTD

```
<?xml version="1.0" ?>
<!DOCTYPE db [
  <!ELEMENT db (kniha*,vydavatel*)>
  <!ELEMENT kniha (titulka,rok,vydavatelstvo)>
  <!ELEMENT titulka (#PCDATA)>
  ...
]>
<db>
  <vydavatel>
    <meno>Jan</meno>
    <priezvisko>Knihovy</priezvisko>
  </vydavatel>
  <vydavatel>
    <firma>
      <nazov>Knihy noveho milenia</nazov>
      <sidlo>Knizna 24, Bratislava</sidlo>
    </firma>
  </vydavatel>
</db>
```

XML obsahuje nielen elementy...

```
<?xml version="1.0" ?>
<!DOCTYPE db SYSTEM "databaza.dtd">
<db>
  <vydavatel id="1">
    <meno>Jan</meno>
    <priezvisko>Knihovy</priezvisko>
  </vydavatel>
  <vydavatel id="2">
    <firma typ="sro" ICO="102030">
      <nazov>x &lt; y</nazov>
      <sidlo>Knizna 24, Bratislava</sidlo>
    </firma>
  </vydavatel>
</db>
```

... atribúty ...

```
<?xml version="1.0" ?>
<!DOCTYPE db SYSTEM "databaza.dtd">
<db>
  <vydavatel id="1">
    <meno>Jan</meno>
    <priezvisko>Knihovy</priezvisko>
  </vydavatel>
  <vydavatel id="2">
    <firma typ="sro" ICO="102030">
      <nazov>x &lt; y</nazov>
      <sidlo>Knizna 24, Bratislava</sidlo>
    </firma>
  </vydavatel>
</db>
```

... a entity

```
<?xml version="1.0" ?>
<!DOCTYPE db SYSTEM "databaza.dtd">
<db>
  <vydavatel id="1">
    <meno>Jan</meno>
    <priezvisko>Knihovy</priezvisko>
  </vydavatel>
  <vydavatel id="2">
    <firma typ="sro" ICO="102030">
      <nazov>x &lt; y</nazov>
      <sidlo>Knizna 24, Bratislava</sidlo>
    </firma>
  </vydavatel>
</db>
```

DTD: definícia atribútov

```
<!ATTLIST element atribut typ hodnota_atributu >
```

Nepovinný atribút s defaultnou hodnotou

```
<!ELEMENT firma (#PCDATA)>  
<!ATTLIST firma typ CDATA "sro">
```

Valid	Invalid
<pre><firma typ="sro"> x &lt; y </firma> <firma typ="as"> x &lt; y </firma> <firma> x &lt; y </firma></pre>	<pre><firma id="1"> x &lt; y </firma> <firma typ="as" id="1"> x &lt; y </firma></pre>

Defaultná hodnota môže byť aj prázdna

```
<!ELEMENT firma (#PCDATA)>  
<!ATTLIST firma typ CDATA "">
```

Valid	Invalid
<pre><firma typ="sro"> x &lt; y </firma> <firma typ="as"> x &lt; y </firma> <firma> x &lt; y </firma></pre>	<pre><firma id="1"> x &lt; y </firma> <firma typ="as" id="1"> x &lt; y </firma></pre>

CDATA vs PCDATA

```
<!ELEMENT firma (#PCDATA)>  
<!ATTLIST firma typ CDATA "">
```

Valid	Invalid
<pre><firma typ="sro"> x &lt; y </firma> <firma typ="as"> x &lt; y </firma> <firma> x &lt; y </firma></pre>	<pre><firma id="1"> x &lt; y </firma> <firma typ="as" id="1"> x &lt; y </firma></pre>

Nepovinný atribút bez defaultnej hodnoty

```
<!ELEMENT firma (#PCDATA)>  
<!ATTLIST firma typ CDATA #IMPLIED>
```

Valid	Invalid
<pre><firma typ="sro"> x &lt; y </firma> <firma> x &lt; y </firma></pre>	<pre><firma id="1"> x &lt; y </firma> <firma typ="sro" id="1"> x &lt; y </firma></pre>

Povinný atribút

```
<!ELEMENT firma (#PCDATA)>
<!ATTLIST firma typ CDATA #REQUIRED>
```

Valid	Invalid
<pre><firma typ="sro"> </firma> <firma typ="as"> x &lt; y </firma></pre>	<pre><firma> x &lt; y </firma> <firma id="1"> x &lt; y </firma> <firma typ="sro" id="1"> x &lt; y </firma></pre>

Povinný atribút s fixnou hodnotou

```
<!ELEMENT firma (#PCDATA)>
<!ATTLIST firma typ CDATA #FIXED "sro">
```

Valid	Invalid
<pre><firma typ="sro"> </firma> <firma typ="sro"> x &lt; y </firma></pre>	<pre><firma typ="as"> x &lt; y </firma></pre>

Atribút s hodnotou typu "výber"

```
<!ELEMENT firma (#PCDATA)>
<!ATTLIST firma typ (sro|as|vos) "sro">
```

Valid	Invalid
<pre><firma> </firma> <firma typ="sro"> </firma> <firma typ="as"> </firma> <firma typ="vos"> </firma></pre>	<pre><firma typ="ltd"> </firma> <firma typ="inc"> </firma> <firma ICO="4"> </firma> <firma typ="as" ICO="4"> </firma></pre>

Element s viacerými atribútmi

```
<!ELEMENT firma (#PCDATA)>
<!ATTLIST firma typ CDATA "sro"
              ICO CDATA #REQUIRED
>
```

Valid	Invalid
<pre><firma ICO="2"> </firma> <firma typ="sro" ICO="2"> </firma> <firma typ="as" ICO="3"> </firma></pre>	<pre><firma typ="as"> </firma> <firma> </firma> <firma ICO="4" id="1"> </firma></pre>

Element s viacerými atribútmi

```
<!ELEMENT firma (#PCDATA)>
<!ATTLIST firma typ CDATA "sro"
              ICO CDATA #REQUIRED
>
```

Valid	Invalid
<pre><firma ICO="2"> </firma> <firma typ="sro" ICO="2"> </firma> <firma typ="as" ICO="3"> </firma> <firma ICO="3" typ="as"> </firma></pre>	<pre><firma typ="as"> </firma> <firma> </firma> <firma ICO="4" id="1"> </firma></pre>

Príklad: matica

Požiadavky na DTD:

- hlavný koreňový element: <matica>
- matica obsahuje žiaden, jeden alebo viac riadkov
- povinné atribúty: "nr", "nc"
- každý riadok obsahuje žiaden, jeden alebo viac stĺpcov
- každý stĺpec obsahuje reťazec
- voliteľný atribút "obsah": výber z "bin", "dec", "hex"

```
<?xml version="1.0" encoding="UTF-8" ?>
<!ELEMENT matica (riadok*)>
<!ATTLIST matica nr CDATA #REQUIRED
                 nc CDATA #REQUIRED
>
<!ELEMENT riadok (stlpec*)>
<!ELEMENT stlpec (#PCDATA)>
<!ATTLIST stlpec obsah (bin|dec|hex) #IMPLIED >
```

Príklad: matica

Požiadavky na DTD:

- hlavný koreňový element: <matica>
- matica obsahuje žiaden, jeden alebo viac riadkov
- povinné atribúty: "nr", "nc"
- každý riadok obsahuje žiaden, jeden alebo viac stĺpcov
- každý stĺpec obsahuje reťazec
- voliteľný atribút "obsah": výber z "bin", "dec", "hex", default: "bin"

```
<?xml version="1.0" encoding="UTF-8" ?>
<!ELEMENT matica (riadok*)>
<!ATTLIST matica nr CDATA #REQUIRED
                 nc CDATA #REQUIRED
>
<!ELEMENT riadok (stlpec*)>
<!ELEMENT stlpec (#PCDATA)>
<!ATTLIST stlpec obsah (bin|dec|hex) "bin" >
```

Atribút s hodnotou "unikátne ID"

```
<!ELEMENT zoznam (firma*)>
<!ELEMENT firma (#PCDATA)>
<!ATTLIST firma ICO ID #REQUIRED>
```

Valid	Invalid
<p>ID sa vždy musí začínať písmenom!</p> <pre><zoznam> <firma ICO="SK01"> Apple </firma> <firma ICO="SK23"> Microsoft </firma> </zoznam></pre>	<pre><firmy> <firma ICO="SK01"> Apple </firma> <firma ICO="SK01"> Microsoft </firma> </firmy></pre>

Atribút s hodnotou "referencia ID"

```
<!ELEMENT zoznam (firma*,clovek*)>
<!ELEMENT firma (#PCDATA)>
<!ATTLIST firma ICO ID #REQUIRED>
<!ELEMENT clovek (#PCDATA)>
<!ATTLIST clovek kde IDREF #REQUIRED>
```

Valid	Invalid
<pre><zoznam> <firma ICO="SK01"> Apple </firma> <clovek kde="SK01"> Steve Jobs </clovek> <clovek kde="SK01"> Bill Gates </clovek> </zoznam></pre>	<pre><zoznam> <firma ICO="SK01"> Apple </firma> <clovek kde="SK23"> Steve Jobs </clovek> <clovek kde="SK01"> Bill Gates </clovek> </zoznam></pre>

Atribút s hodnotou "viaceré referencie ID"

```
<!ELEMENT zoznam (clovek*,firma*)>
<!ELEMENT clovek (#PCDATA)>
<!ATTLIST clovek id ID #REQUIRED>
<!ELEMENT firma (#PCDATA)>
<!ATTLIST firma ludia IDREFS #REQUIRED>
```

Valid	Invalid
<pre><zoznam> <clovek id="x1"> Steve Jobs </clovek> <clovek id="x2"> Bill Gates </clovek> <firma ludia="x1 x2"> Apple </firma> </zoznam></pre>	<pre><zoznam> <clovek id="x1"> Steve Jobs </clovek> <clovek id="x2"> Bill Gates </clovek> <firma ludia="x3 x1"> Apple </firma> </zoznam></pre>

Príklad: matrika

- Požiadavky na DTD:
 - hlavný koreňový element <ludia> obsahuje 1 alebo viac <clovek>
 - <clovek> má nasledovné atribúty:
 - povinné unikátne ID "meno"
 - povinné pohlavie "muz" alebo "zena"
 - voliteľný odkaz na "otec" a "matka"
 - voliteľný odkaz na viacero "deti"

```
<?xml version="1.0" encoding="UTF-8" ?>
<!ELEMENT ludia (clovek+)>
<!ELEMENT clovek (#PCDATA)>
<!ATTLIST clovek meno ID #REQUIRED
           pohlavie (muz|zena) #REQUIRED
           otec IDREF #IMPLIED
           matka IDREF #IMPLIED
           deti IDREFS #IMPLIED
>
```

Príklad: matrika

```
<?xml version="1.0" encoding="UTF-8" ?>
<!ELEMENT ludia (clovek+)>
<!ELEMENT clovek (#PCDATA)>
<!ATTLIST clovek meno ID #REQUIRED
           pohlavie (muz|zena) #REQUIRED
           otec IDREF #IMPLIED
           matka IDREF #IMPLIED
           deti IDREFS #IMPLIED
>

<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE ludia SYSTEM "ludia.dtd">
<ludia>
  <clovek meno="Eva" pohlavie="zena"></clovek>
  <clovek meno="Adam" pohlavie="muz"
    deti="Jan Lenka"></clovek>
  <clovek meno="Jan" pohlavie="muz"
    matka="Eva" otec="Adam"></clovek>
  <clovek meno="Lenka" pohlavie="zena"
    matka="Eva" otec="Adam"></clovek>
</ludia>
```

DTD kontroluje iba unikátnosť, nie význam!

```
<?xml version="1.0" encoding="UTF-8" ?>
<!ELEMENT ludia (clovek+)>
<!ELEMENT clovek (#PCDATA)>
<!ATTLIST clovek meno ID #REQUIRED
           pohlavie (muz|zena) #REQUIRED
           otec IDREF #IMPLIED
           matka IDREF #IMPLIED
           deti IDREFS #IMPLIED
>

<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE ludia SYSTEM "ludia.dtd">
<ludia>
  <clovek meno="Eva" pohlavie="zena"></clovek>
  <clovek meno="Adam" pohlavie="muz"
    deti="Jan Lenka"></clovek>
  <clovek meno="Jan" pohlavie="muz"
    matka="Eva" otec="Lenka"></clovek>
  <clovek meno="Lenka" pohlavie="zena"
    matka="Eva" otec="Adam"></clovek>
</ludia>
```

ID a IDREF sa nedá ohraničiť na určitú vetvu stromu!

```
<zoznam>
  <firma ICO="SK01">
    <oddelenie id="vyskum"/>
    <oddelenie id="vyvoj"/>
    <clovek kde="vyskum vyvoj"/>
  </firma>
  <firma ICO="SK02">
    <oddelenie id="vyroba"/>
    <oddelenie id="vyvoj"/>
    <clovek kde="vyskum"/>
  </firma>
</zoznam>
```

XML obsahuje aj entity

```
<?xml version="1.0" ?>
<!DOCTYPE db SYSTEM "databaza.dtd">
<db>
  <vydavatel id="1">
    <meno>Jan</meno>
    <priezvisko>Knihovy</priezvisko>
  </vydavatel>
  <vydavatel id="2">
    <firma typ="sro" ICO="102030">
      <nazov>x &lt; y</nazov>
      <sidlo>Knizna 24, Bratislava</sidlo>
    </firma>
  </vydavatel>
</db>
```

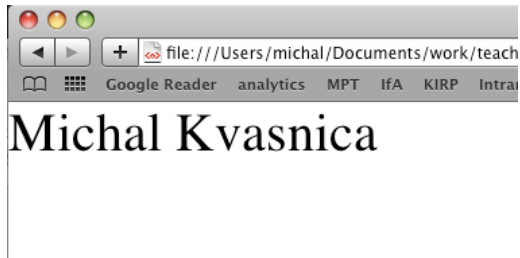
XML entity

- Entita = skratka pre jeden alebo viac znakov
- Preddefinované jednoznakové entity:

<	<
>	>
&	&
'	'
"	"

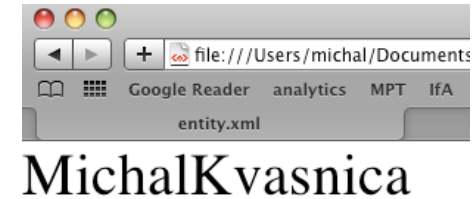
Definícia vlastných entít

```
<?xml version="1.0" encoding="UTF-8" ?>
<!DOCTYPE clovek [
<!ELEMENT clovek (#PCDATA)>
<!ENTITY tvorca_dtd "Michal Kvasnica">
]>
<clovek>
  &tvorca_dtd;
</clovek>
```



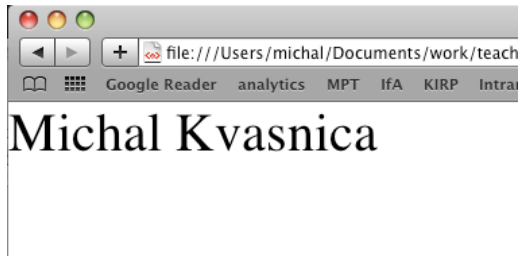
Entity môžu ísť aj za sebou

```
<?xml version="1.0" encoding="UTF-8" ?>
<!DOCTYPE clovek [
<!ELEMENT clovek (#PCDATA)>
<!ENTITY tvorca_dtd_meno "Michal">
<!ENTITY tvorca_dtd_priezvisko "Kvasnica">
]>
<clovek>
  &tvorca_dtd_meno;&tvorca_dtd_priezvisko;
</clovek>
```



Pozor na medzery

```
<?xml version="1.0" encoding="UTF-8" ?>
<!DOCTYPE clovek [
<!ELEMENT clovek (#PCDATA)>
<!ENTITY tvorca_dtd_meno "Michal">
<!ENTITY tvorca_dtd_priezvisko "Kvasnica">
]>
<clovek>
  &tvorca_dtd_meno; &tvorca_dtd_priezvisko;
</clovek>
```



Entita "nbsp" nie je štandardne definovaná!

```
<?xml version="1.0" encoding="UTF-8" ?>
<!DOCTYPE clovek [
<!ELEMENT clovek (#PCDATA)>
<!ENTITY tvorca_dtd_meno "Michal">
<!ENTITY tvorca_dtd_priezvisko "Kvasnica">
<!ENTITY nbsp " ">
]>
<clovek>
  &tvorca_dtd_meno; &nbsp; &tvorca_dtd_priezvisko;
</clovek>
```

Príklad: evidencia stavebného materiálu

- Hlavný element <stavba> obsahuje 1+ <material>
- <material> obsahuje <opis>, <dodavatel>, <instalater>, <kontrolor>
- <opis> je string
- dodavatel, instalater a kontrolor su:
 - bud <meno> a <priezvisko>
 - alebo <firma>: atribút ICO je povinné ID, atribút typ=(sro|as) #IMPLIED

```
<!ELEMENT stavba (material+)>
<!ELEMENT material (opis,dodavatel,instalater,kontrolor)>
<!ELEMENT opis (#PCDATA)>
<!ELEMENT dodavatel ((meno,priezvisko)|firma)>
<!ELEMENT instalater ((meno,priezvisko)|firma)>
<!ELEMENT kontrolor ((meno,priezvisko)|firma)>
<!ELEMENT meno (#PCDATA)>
<!ELEMENT priezvisko (#PCDATA)>
<!ELEMENT firma (#PCDATA)>
<!ATTLIST firma ICO ID #REQUIRED
                typ (sro|as) #IMPLIED
>
```

```
<!ELEMENT stavba (material+)>
<!ELEMENT material (opis,dodavatel,instalater,kontrolor)>
<!ELEMENT opis (#PCDATA)>
<!ELEMENT dodavatel ((meno,priezvisko)|firma)>
<!ELEMENT instalater ((meno,priezvisko)|firma)>
<!ELEMENT kontrolor ((meno,priezvisko)|firma)>
<!ELEMENT meno (#PCDATA)>
<!ELEMENT priezvisko (#PCDATA)>
<!ELEMENT firma (#PCDATA)>
<!ATTLIST firma ICO ID #REQUIRED
                typ (sro|as) #IMPLIED>
```

```
    <stavba>
      <material>
        <opis>klince</opis>
        <dodavatel>
          <meno>Janko</meno>
          <priezvisko>Klincovy</priezvisko>
        </dodavatel>
        <instalater>
          <meno>Janko</meno>
          <priezvisko>Klincovy</priezvisko>
        </instalater>
        <kontrolor>
          <meno>Janko</meno>
          <priezvisko>Klincovy</priezvisko>
        </kontrolor>
      </material>
    </stavba>
```

Je možné zjednodušiť opakovaný zápis?

```
<!ELEMENT stavba (material+)>
<!ELEMENT material (opis,dodavatel,instalater,kontrolor)>
<!ELEMENT opis (#PCDATA)>
<!ELEMENT dodavatel ((meno,priezvisko)|firma)>
<!ELEMENT instalater ((meno,priezvisko)|firma)>
<!ELEMENT kontrolor ((meno,priezvisko)|firma)>
<!ELEMENT meno (#PCDATA)>
<!ELEMENT priezvisko (#PCDATA)>
<!ELEMENT firma (#PCDATA)>
<!ATTLIST firma ICO ID #REQUIRED
                typ (sro|as) #IMPLIED
>
```

Je možné zjednodušiť opakovaný zápis?

```
<!ENTITY % firmy "(meno,priezvisko)|firma">
<!ELEMENT stavba (material+)>
<!ELEMENT material (opis,dodavatel,instalater,kontrolor)>
<!ELEMENT opis (#PCDATA)>
<!ELEMENT dodavatel ((meno,priezvisko)|firma)>
<!ELEMENT instalater ((meno,priezvisko)|firma)>
<!ELEMENT kontrolor ((meno,priezvisko)|firma)>
<!ELEMENT meno (#PCDATA)>
<!ELEMENT priezvisko (#PCDATA)>
<!ELEMENT firma (#PCDATA)>
<!ATTLIST firma ICO ID #REQUIRED
                typ (sro|as) #IMPLIED
>
```

Je možné zjednodušiť opakovaný zápis?

špeciálne označenie skratkovej entity

```
<!ENTITY % firmy "(meno,priezvisko)|firma">
<!ELEMENT stavba (material+)>
<!ELEMENT material (opis,dodavatel,instalater,kontrolor)>
<!ELEMENT opis (#PCDATA)>
<!ELEMENT dodavatel ((meno,priezvisko)|firma)>
<!ELEMENT instalater ((meno,priezvisko)|firma)>
<!ELEMENT kontrolor ((meno,priezvisko)|firma)>
<!ELEMENT meno (#PCDATA)>
<!ELEMENT priezvisko (#PCDATA)>
<!ELEMENT firma (#PCDATA)>
<!ATTLIST firma ICO ID #REQUIRED
                typ (sro|as) #IMPLIED
>
```

Je možné zjednodušiť opakovaný zápis?

potom nasleduje medzera

```
<!ENTITY % firmy "(meno,priezvisko)|firma">
<!ELEMENT stavba (material+)>
<!ELEMENT material (opis,dodavatel,instalater,kontrolor)>
<!ELEMENT opis (#PCDATA)>
<!ELEMENT dodavatel ((meno,priezvisko)|firma)>
<!ELEMENT instalater ((meno,priezvisko)|firma)>
<!ELEMENT kontrolor ((meno,priezvisko)|firma)>
<!ELEMENT meno (#PCDATA)>
<!ELEMENT priezvisko (#PCDATA)>
<!ELEMENT firma (#PCDATA)>
<!ATTLIST firma ICO ID #REQUIRED
                typ (sro|as) #IMPLIED
>
```

Je možné zjednodušiť opakovaný zápis?

po nej názov entity

```
<!ENTITY % firmy "(meno,priezvisko)|firma">
<!ELEMENT stavba (material+)>
<!ELEMENT material (opis,dodavatel,instalater,kontrolor)>
<!ELEMENT opis (#PCDATA)>
<!ELEMENT dodavatel ((meno,priezvisko)|firma)>
<!ELEMENT instalater ((meno,priezvisko)|firma)>
<!ELEMENT kontrolor ((meno,priezvisko)|firma)>
<!ELEMENT meno (#PCDATA)>
<!ELEMENT priezvisko (#PCDATA)>
<!ELEMENT firma (#PCDATA)>
<!ATTLIST firma ICO ID #REQUIRED
                typ (sro|as) #IMPLIED
>
```

Je možné zjednodušiť opakovaný zápis?

a nakoniec obsah entity ako string

```
<!ENTITY % firmy "(meno,priezvisko)|firma">
<!ELEMENT stavba (material+)>
<!ELEMENT material (opis,dodavatel,instalater,kontrolor)>
<!ELEMENT opis (#PCDATA)>
<!ELEMENT dodavatel ((meno,priezvisko)|firma)>
<!ELEMENT instalater ((meno,priezvisko)|firma)>
<!ELEMENT kontrolor ((meno,priezvisko)|firma)>
<!ELEMENT meno (#PCDATA)>
<!ELEMENT priezvisko (#PCDATA)>
<!ELEMENT firma (#PCDATA)>
<!ATTLIST firma ICO ID #REQUIRED
                typ (sro|as) #IMPLIED
>
```

Použitie “opakovecej” entity

```
<!ENTITY % firmy "(meno,priezvisko)|firma">
<!ELEMENT stavba (material+)>
<!ELEMENT material (opis,dodavatel,instalater,kontrolor)>
<!ELEMENT opis (#PCDATA)>
<!ELEMENT dodavatel (%firmy;)>
<!ELEMENT instalater ((meno,priezvisko)|firma)>
<!ELEMENT kontrolor ((meno,priezvisko)|firma)>
<!ELEMENT meno (#PCDATA)>
<!ELEMENT priezvisko (#PCDATA)>
<!ELEMENT firma (#PCDATA)>
<!ATTLIST firma ICO ID #REQUIRED
                typ (sro|as) #IMPLIED
>
```

Použitie “opakovecej” entity

```
<!ENTITY % firmy "(meno,priezvisko)|firma">
<!ELEMENT stavba (material+)>
<!ELEMENT material (opis,dodavatel,instalater,kontrolor)>
<!ELEMENT opis (#PCDATA)>
<!ELEMENT dodavatel (%firmy;)>
<!ELEMENT instalater (%firmy;)>
<!ELEMENT kontrolor ((meno,priezvisko)|firma)>
<!ELEMENT meno (#PCDATA)>
<!ELEMENT priezvisko (#PCDATA)>
<!ELEMENT firma (#PCDATA)>
<!ATTLIST firma ICO ID #REQUIRED
                typ (sro|as) #IMPLIED
>
```

Použitie “opakovecej” entity

```
<!ENTITY % firmy "(meno,priezvisko)|firma">
<!ELEMENT stavba (material+)>
<!ELEMENT material (opis,dodavatel,instalater,kontrolor)>
<!ELEMENT opis (#PCDATA)>
<!ELEMENT dodavatel (%firmy;)>
<!ELEMENT instalater (%firmy;)>
<!ELEMENT kontrolor (%firmy;)>
<!ELEMENT meno (#PCDATA)>
<!ELEMENT priezvisko (#PCDATA)>
<!ELEMENT firma (#PCDATA)>
<!ATTLIST firma ICO ID #REQUIRED
                typ (sro|as) #IMPLIED
>
```

“Opakovacia” entita sa správa ako substitúcia

```
<!ENTITY % firmy "(meno,priezvisko)|firma">
<!ELEMENT stavba (material+)>
<!ELEMENT material (opis,dodavatel,instalater,kontrolor)>
<!ELEMENT opis (#PCDATA)>
<!ELEMENT dodavatel ((%firmy;)+)>
<!ELEMENT instalater ((%firmy;)+)>
<!ELEMENT kontrolor ((%firmy;)+)>
<!ELEMENT meno (#PCDATA)>
<!ELEMENT priezvisko (#PCDATA)>
<!ELEMENT firma (#PCDATA)>
<!ATTLIST firma ICO ID #REQUIRED
                typ (sro|as) #IMPLIED
>
```

Prečo sme pridali zátvorky?

```
<!ENTITY % firmy "(meno,pr)|firma">
<!ELEMENT dodavatel (%firmy;+)>
<!ELEMENT dodavatel ((meno,pr)|firma+)>
```

Valid	Invalid
<pre><dodavatel> <meno></meno> <pr></pr> </dodavatel></pre>	<pre><dodavatel> <meno></meno> <pr></pr> <meno></meno> <pr></pr> </dodavatel></pre>
<pre><dodavatel> <firma></firma> <firma></firma> <firma></firma> </dodavatel></pre>	<pre><dodavatel> <meno></meno> <pr></pr> <firma></firma> </dodavatel></pre>

Prečo sme pridali zátvorky?

```
<!ENTITY % firmy "(meno,pr)|firma">
<!ELEMENT dodavatel ((%firmy;)+)>
<!ELEMENT dodavatel (((meno,pr)|firma)+)>
```

Valid	Invalid
<pre><dodavatel> <meno></meno> <pr></pr> <meno></meno> <pr></pr> <firma></firma> </dodavatel></pre>	<pre><dodavatel> <meno></meno> </dodavatel></pre>
<pre><dodavatel> <firma></firma> <firma></firma> <firma></firma> </dodavatel></pre>	<pre><dodavatel> <pr></pr> <firma></firma> </dodavatel></pre>

Nástroj xmllint

```
$ xmllint --noout --dtdvalid sablona.dtd subor.xml
```

Príklad: adresa servera

```
<?xml version="1.0" encoding="UTF-8" ?>
<!ELEMENT Server (Name, Address)>
<!ELEMENT Name (#PCDATA)>
<!ELEMENT Address (#PCDATA)>
```

```
<?xml version="1.0" ?>
<Server>
  <Name>OurWebServer</Name>
  <Address>123.45.67.8</Address>
</Server>
```


Príklad: adresa budovy

```
<?xml version="1.0" encoding="UTF-8" ?>
<!ELEMENT Address (Street, City, Country, ZIP)>
<!ELEMENT Street (#PCDATA)>
<!ELEMENT City (#PCDATA)>
<!ELEMENT Country (#PCDATA)>
<!ELEMENT ZIP (#PCDATA)>
```

```
<?xml version="1.0" encoding="UTF-8" ?>
<Address>
  <Street>Wilhelminenstr. 7</Street>
  <City>Darmstadt</City>
  <Country>Germany</Country>
  <ZIP>D-64285</ZIP>
</Address>
```

Čo ak ich teraz skombinujeme?

```
<?xml version="1.0" encoding="UTF-8" ?>
<!ELEMENT Department (Address, Server)>
<!ELEMENT Address (Street, City, Country, ZIP)>
<!ELEMENT Street (#PCDATA)>
<!ELEMENT City (#PCDATA)>
<!ELEMENT Country (#PCDATA)>
<!ELEMENT ZIP (#PCDATA)>
<!ELEMENT Server (Name, Address)>
<!ELEMENT Name (#PCDATA)>
<!ELEMENT Address (#PCDATA)>
```

```
<?xml version="1.0" encoding="UTF-8" ?>
<Department>
  <Address>
    <Street>Wilhelminenstr. 7</Street>
    <City>Darmstadt</City>
    <Country>Germany</Country>
    <ZIP>D-64285</ZIP>
  </Address>
  <Server>
    <Name>OurWebServer</Name>
    <Address>123.45.67.8</Address>
  </Server>
</Department>
```

V DTD sa nesmú redefinovať elementy

```
<?xml version="1.0" encoding="UTF-8" ?>
<!ELEMENT Department (Address, Server)>
<!ELEMENT Address (Street, City, Country, ZIP)>
<!ELEMENT Street (#PCDATA)>
<!ELEMENT City (#PCDATA)>
<!ELEMENT Country (#PCDATA)>
<!ELEMENT ZIP (#PCDATA)>
<!ELEMENT Server (Name, Address)>
<!ELEMENT Name (#PCDATA)>
<!ELEMENT Address (#PCDATA)>
```

```
department.dtd:10: validity error :
Redefinition of element Address
<!ELEMENT Address (#PCDATA)>
      ^
department.xml:11: element Address: validity
error : Element Address content does not
follow the DTD, expecting (Street , City ,
Country , ZIP), got (CDATA)
```

```
    <Name>OurWebServer</Name>
    <Address>123.45.67.8</Address>
  </Server>
</Department>
```