
Laboratorium 4 – Projektowanie języków XML

DTD

Co poznamy na zajęciach i do czego się to przyda?

Dlaczego strona html musi zaczynać się od tagu <html>? Bo nakazuje jej to pewny szablon zwany Document Type Definition (DTD), który mówi jakie tagi są dozwolone, a jakie nie. DTD może też definiować (w ograniczonym stopniu) strukturę dokumentu, sposób zagnieżdżania czy rodzaj danych. DTD pilnują poprawności XML, więc są przydatne przy wielu technologiach.

Dzisiaj poznamy jak tworzyć DTD i walidować pliki XML z DTD.

Zadanie 1

Zapoznaj się z materiałami na stronie:

- http://www.w3schools.com/xml/xml_dtd_intro.asp (samouczek DTD: rozdział DTD Intro, aż do DTD Examples w zakładce XML DTD)

Spróbuj znaleźć i inne ciekawe tutoriale.

Pobierz plik uczelnia.dtd (<https://inf.ug.edu.pl/~gmadejsk/xml/docs/uczelnia.dtd>) i zapoznaj się z nim. Wyjaśnij:

- Co oznaczają symbole: #PCDATA ? * + | EMPTY
- Czy w dokumencie DTD jest encja? Gdzie jest zadeklarowana, a gdzie wykorzystana?

Zadanie 2

Stwórz plik uczelnia.xml i zadeklaruj w nim walidację względem ściągniętego szablonu uczelnia.dtd (szablon dtd pozostaje w osobnym pliku czyli to DTD zewnętrzne) wpisując w odpowiednim miejscu:

```
<!DOCTYPE uczelnia SYSTEM "uczelnia.dtd">
```

Plik uczelnia.xml powinien być skonstruowany i wypełniony przez ciebie sensownymi danymi (np. z Uniwersytetu Gdańskiego).

Zwaliduj plik względem szablonu komendą:

```
xmllint --valid uczelnia.xml
```

lub

```
xmllint --noout --valid uczelnia.xml
```

Zadanie 3

Pobierz plik biblioteka.xml (<https://inf.ug.edu.pl/~gmadejsk/xml/docs/biblioteka.xml>) i zapoznaj się z nim. Zmodyfikuj go dodając DTD wewnętrzne opisujące strukturę pliku:

- Tytuł, autor i opis zawierają tekst.
- Książka ma kilka atrybutów:
 - id – jest wymagane, atrybut tekstowy
 - dzial – jest wymagany, atrybut tekstowy
 - oprawa – jest wymagany, atrybut przyjmuje tylko wartości twarda lub miękka
 - opcja – jest opcjonalny, tekstowy
- Dział jest elementem tekstowym i ma tekstowy atrybut wymagany: kod.

Zwaliduj plik xml, nie zapomnij o dodaniu --valid.

Zadanie 4

Zmodyfikuj poprzednie zadanie tak, aby id książki było atrybutem typu ID (w DTD), kod działu też był atrybutem typu ID.

W pliku DTD powinien pojawić się też plik IDREF odwołujący się do jakiegoś ID, znajdź go i nadaj mu typ IDREF. Nie będę podpowiadać jaki to atrybut, odwołam się do inteligencji czytelnika ;)

Zwaliduj plik xml, nie zapomnij o dodaniu --valid.

Sprawdź czy plik się będzie walidował gdy zmienisz wartość jednego atrybutu dzial na nieistniejący np. xyz.

Sprawdzenie

Po wykonaniu zadań poproś prowadzącego o sprawdzenie pliku uczelnia.xml. Omów po krótku znaczenie symboli z uczelnia.dtd. Pokaż i zwaliduj uzupełniony plik biblioteka.xml.

Podsumowanie

Na dzisiejszy zajęciach poznaliśmy jak opisywać pliki XML za pomocą szablonów DTD (zewnętrznie i wewnętrznie). Znamy składnię DTD i umiemy korzystać z narzędzia walidującego.

Na kolejnych zajęciach...

Na kolejnych zajęciach prawdopodobnie zajmiemy się schematami XSD.