

Temat: XML Schema cz.1. - typy proste i złożone

Czego dotyczą zajęcia?

DTD jest prostym narzędziem do tworzenia szablonu dokumentu XML i sprawdza się w webowych zastosowaniach (np. HTML), gdzie zawartość elementów nie musi być szczegółowo ustrukturyzowana. Jednak w wielu przypadkach jest zbyt proste i ograniczone - co było szczegółowo omówione na wykładzie. Jednym z głównych ograniczeń jest brak mechanizmu typowania, który jest kluczowy przy projektowaniu samoopisujących się dokumentów. Stąd do tworzenia modeli dokumentów znacznie częściej wykorzystuje się model XML Schema (schematy dokumentów XML), który daje możliwość tworzenia dużo bardziej zaawansowanych i szczegółowych szablonów. Pewien podzbiór modeli zaprojektowanych w XML Schema: https://en.wikipedia.org/wiki/List_of_types_of_XML_schemas

Zadanie 1

Zapoznaj się z materiałami na stronie:

- http://www.w3schools.com/xml/schema_intro.asp (samouczek XSD Schema, XSD Complex i XSD Data: należy zescrollować menu po lewej i przejrzeć te trzy działy)

Pobierz plik osoby.xsd (<https://inf.ug.edu.pl/~maggod/zasoby/osoby.xsd>) , który jest bardzo prostym schematem dla dokumentów XML.

a) Stwórz dokument XML realizujący ten schemat (wypełniony przykładowymi danymi – minimum dwie osoby) i dokonaj jego walidacji. Spróbuj zrozumieć zawartość pliku XSD.

Komenda:

```
xmllint --schema schemat.xsd plik.xml
```

lub bez wyświetlania pliku w konsoli komendą:

```
xmllint --noout --schema schemat.xsd plik.xml
```

lub inny walidator (np. online).

b) Zauważ, że w schemacie osoby.xsd definicje elementów są zagnieżdżone w definicjach innych elementów. Drugą częścią zadania jest pozbycie się zagnieżdżeń w pliku osoby.xsd. Pomoc znajduje się tutaj: http://www.w3schools.com/xml/schema_example.asp

Jak widać są dwa sposoby pozbycia się zagnieżdżeń:

- za pomocą referencji do elementów (rozdział Divide the Schema)
 - za pomocą definiowania typów (rozdział Using Named Types – uwaga jeśli korzysta się z wbudowanych typów prostych bez żadnych restrikcji, to nie trzeba ich oddzielnie definiować jako simpleType – XML Schema jest i tak już wystarczająco przegadana :))
- Zapoznaj się z obiema metodami, a następnie wybierz jedną z nich i zastosuj.

Zadanie 2

Pobierz plik listastud.xml (<https://inf.ug.edu.pl/~maggod/zasoby/listastud.xml>). Dla tego dokumentu napisz schemat XSD walidujący go. Schemat napisz jedną z dwóch poznanych w zadaniu 1 metod unikania zagnieżdżeń – **wybierz inną niż użyłeś w zadaniu 1.**

Wymagania co do zadania:

- Nazwisko i imię to łańcuchy znaków.
- Nr indeksu to łańcuch znaków składający się z 6 cyfr (przykład na stronie, tam gdzie Using Named Types).
- Data urodzenia musi mieć format daty YYYY-MM-DD – użyj odpowiedniego typu wbudowanego

- Średnia ocen to liczba zmiennoprzecinkowa.
- Strona domowa to adres URI – tutaj też odszukaj odpowiedni typ.
- Dokument może zawierać wielu studentów (minimum to 1)

Sprawdzenie

Zadanie 1: zaprezentuj walidację osoby.xsd i osoby.xml. Pokaż plik xml. Jak zostały usunięte zagnieżdżenia?

Zadanie 2: zaprezentuj walidację listastud.xsd i listastud.xml. Zaprezentuj definicję typów w XSD.

Podsumowanie

Na dzisiejszy zajęciach poznaliście, jak opisywać pliki XML za pomocą schematów XSD, a także podstawowe typy (łańcuchy znaków, liczby, daty, itp.) oraz sposób tworzenia typów złożonych przez sekwencję. Poznaliście też różne metody tworzenia schematu (zagnieżdżony, referencyjny, własne typy).

Na kolejnych zajęciach...

Na kolejnych zajęciach poznacie bardziej zaawansowane typy i konstrukcje.