
Laboratorium 5 – Projektowanie języków XML

XML Schema – typy proste i złożone

Co poznamy na zajęciach i do czego się to przyda?

DTD służyło nam jako pewien rodzaj szablonu dokumentu XML i sprawdza się w wielu webowych zastosowaniach (np. HTML), niestety w wielu innych przypadkach jest zbyt proste i ma za małe możliwości. Ograniczenia DTD były szczegółowo omówione na wykładzie, jednym z nich jest np. brak możliwości zdefiniowania typu zawartości elementu (tekstowe, numeryczne, daty). Na szczęście W3C wspiera drugi standard: XML Schema (schematy dokumentów XML), które są dużo bardziej zaawansowanymi szablonami niż DTD. Znajdują zastosowanie w niektórych technologiach opartych na XML np. microsoftowy BizTalk opisujący i zarządzający informacjami biznesowymi. Ze schematami XSD zapoznamy się na tych zajęciach.

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 zscrollować menu po lewej i przejrzeć te trzy działy)

Spróbuj znaleźć i inne ciekawe tutoriale.

Pobierz plik osoby.xsd (<https://inf.ug.edu.pl/~gmadejsk/xml/docs/osoby.xsd>), który jest schematem dla pewnego pliku XML. Stwórz plik XML do tego schematu (wypełniony przykładowymi danymi – minimum dwie osoby) i dokonaj walidacji komendą:

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

lub bez wyświetlania pliku w konsoli komendą:

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

Zastanów się:

- Jakie elementy występują w pliku XML i gdzie są podane ich nazwy?
- Które elementy są złożone (complexType), a które są prostego typu (np. stringi, liczby)?
- Co oznacza znacznik `<xsd:sequence></xsd:sequence>` w typach złożonych?

Zadanie 2

Utwórz kopię schematu z zadania 1 i zapisz go pod nazwą osoby2.xsd. Twoim zadaniem jest zmodyfikować teraz osoby2.xsd tak, aby schemat nadal robił dokładnie to samo co osoby.xsd, ale miał „ładniejszą” strukturę.

Zauważ, że w schemacie osoby.xsd definicje elementów są pozagnieżdżane w definicjach innych elementów. Twoim zadaniem jest pozbycie się zagnieżdżeń w pliku osoby2.xsd.

Odsyłam tutaj do dobrego omówienia na stronie:

http://www.w3schools.com/xml/schema_example.asp

Jak widać są dwa sposoby na rozbieżności

- za pomocą referencji do elementów (rozdział Divide the Schema)
- za pomocą definiowania typów (rozdział Using Named Types)

Zapoznaj się z obiema metodami, a następnie wybierz tę, która bardziej ci odpowiada i zastosuj ją dla osoby2.xsd.

Ponownie zwaliduj plik osoby.xml względem osoby2.xsd.

Zadanie 3

Pobierz plik listastud.xml (<https://inf.ug.edu.pl/~gmadejsk/xml/docs/listastud.xml>) i zapoznaj się z nim. Dla danego pliku XML napisz schemat XSD walidujący go. Schemat napisz jedną z dwóch poznanych w zadaniu 2 metod (najlepiej użyj tej, której nie wybrałeś w poprzednim zadaniu).

Wymagania co do zadania:

- Nazwisko i imię to łańcuchy znaków.
- Nr indeksu to liczba całkowita.
- Data urodzenia musi mieć format daty YYYY-MM-DD – jak się nazywa wbudowany typ?
- Średnia ocen to liczba zmiennoprzecinkowa.
- Strona domowa to adres URI – tutaj też odszukaj odpowiedni typ.
- Czy plik XML się waliduje? Czy plik XML będzie się walidował, gdy dopiszemy więcej studentów? Jeśli nie, to jak zaradzić temu problemowi? (pytanie dla chętnych)

Sprawdzenie

Po wykonaniu zadań poproś prowadzącego o sprawdzenie plików:

- osoby.xml z pierwszego zadania (podaj odpowiedzi na pytania z tego zadania i dokonaj walidacji)
- osoby2.xsd z drugiego zadania (wyjaśnij stosowaną metodę i dokonaj walidacji)
- schemat z zadania 3 (omów po krótku stosowane typy, metodę konstrukcji schematu oraz dokonaj walidacji)

Podsumowanie

Na dzisiejszych zajęciach poznaliśmy jak opisywać pliki XML za pomocą schematów XSD. Poznaliśmy podstawowe typy (łańcuchy znaków, liczby, daty, itp.) oraz sposób tworzenia typów złożonych przez sekwencję. Znamy też różne metody tworzenia schematu (zagnieżdżone, referencyjne, własne typy).

Na kolejnych zajęciach...

Na kolejnych zajęciach kontynuujemy temat XSD. Poznamy bardziej zaawansowane typy i konstrukcje.