

Projektowanie języków XML 2016/2017

Laboratorium 5

13 marca 2017

Dziś przypomnimy sobie sposób określania struktury dokumentów, DTD, oraz dowiemy się, czym są encje. W dalszej części poznamy koncept przestrzeni nazw w XML.

Zadanie wprowadzające

W przypomnieniu pomocną mogą okazać się [materiały na temat DTD](#). Na podstronie wspomnianego kursu znajdują się [informacje dotyczące encji](#). Encję definiuje się podobnie jak inne elementy DTD.

```
<!ENTITY nazwa-encji "zawartosc-encji">
```

Użycie encji sprowadza się do użycia w kodzie XML poniższej konstrukcji.

```
&nazwa-encji;
```

Do rozwiązywania dalszych zadań przydatne będą materiały o [przestrzeniach nazw w XML](#).

Zadanie 1 (1 pkt)

Encje służą definiowaniu pewnych skrótów, dzięki którym możemy zakodować znaki specjalne lub umówione wartości.

Przygotuj przykładowy plik XML, w którym wykorzystasz dwie encje. Jedna powinna być zdefiniowana w samym pliku XML, druga – w odrębnym pliku DTD.

Za pomocą komendy `xmllint --valid plik.xml` sprawdź, czy utworzony plik jest poprawny.

Zadanie 2 (1 pkt)

Przygotuj plik DTD opisujący szablon zamówienia towarów w sklepie internetowym. Poprawny względem DTD plik XML powinien zawierać niepustą listę zamówionych towarów. Każdy towar musi zawierać atrybut ID, element opisujący cenę oraz element opisujący ilość (wraz z jednostką) albo liczbę sztuk zamówionego towaru. Zdefiniuj element opisujący produkt, ale na razie będzie mógł przechowywać on dowolne dane. Rozwiniemy go w kolejnych zadaniach.

Zadanie 3 (2 pkt)

W pliku XML z poprzedniego zadania zostawiliśmy niezdefiniowany element `Produkt`. Szablon zamówienia nie będzie się zmieniał, ale typy produktów owszem. Nie możemy więc raz na zawsze zdefiniować struktury całego zamówienia w DTD, gdyż nie wiemy, co w nim się znajdzie w przyszłości.

Na stronie laboratorium znajduje się plik `ksiazka.xsd`. Opisuje on tzw. schemę dla książek, czyli strukturę elementu `Książka` (szczegóły dotyczące narzędzia, jakim jest XML Schema, poznamy na kolejnych zajęciach).

Jako że `Produkt` może zawierać dowolne dane (CDATA), a chcemy mieć dla nich zdefiniowaną

strukturę, posłużymy się definicją zaproponowaną w `ksiazka.xsd`. Definicję stanowi dokument XML, którego elementy należą do pewnej dobrze określonej przestrzeni nazw, tzn. nazwy elementów i atrybutów zostały w tej przestrzeni zdefiniowane. Plik `ksiazka.xsd` jest także definicją pewnej przestrzeni nazw – jej elementami są np. `Książka` i `RokWydania`. Przestrzenie nazw pozwalają na ograniczenie elementów, jakie mogą się znaleźć w dokumencie XML.

W głównym elemencie dokumentu XML z zamówieniem należy umieścić poniższe atrybuty:

`xmlns:k="https://www.inf.ug.edu.pl/xsd/ksiazka"` – definiuje prefiks dla elementów z przestrzeni nazw `https://www.inf.ug.edu.pl/xsd/ksiazka`;

`xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"` – definiuje prefiks dla elementów z przestrzeni nazw `http://www.w3.org/2001/XMLSchema-instance` (potrzebny tylko do zdefiniowania atrybutu poniżej);

`xsi:schemaLocation="https://www.inf.ug.edu.pl/xsd/ksiazka
ksiazka.xsd"` – wskazuje położenie pliku ze schemą, który będzie przydatny do walidacji.

Dzięki zastosowaniu prefiksów możemy logicznie oddzielić elementy i atrybuty należące do określonej przestrzeni nazw (np. dla książek).

Utwórz tymczasowy plik z książką, dodaj powyższe atrybuty i wykonaj walidację poniższą komendą.

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

Zawartość pliku skopiuj do elementu `Produkt` w zamówieniu. Zauważ, że część elementów posiada prefiks `k:`, a część nie. Elementy nienależące do przestrzeni nazw z `ksiazka.xsd` należą do domyślnej przestrzeni nazw, na razie niezdefiniowanej.

Na kolejnych zajęciach przygotujemy schemę dla całego pliku XML z zamówieniem i wtedy wykonamy walidację względem tej schemy.

Zadanie 4 (2 pkt)

Na podstawie schemy `ksiazka.xsd` stwórz analogiczną schemę dla audiobooków na CD (wystarczy zmienić nazwę głównego elementu). Do pliku XML z poprzednich zadań dopisz dodatkowy towar do zamówienia – będzie to audiobook. Przeprowadź walidację dla audiobooka i umieść go w odpowiednim miejscu zamówienia. Zdefiniuj prefiks dla nowego produktu.

Zwróć uwagę, że nazwy niektórych elementów powtarzają się w dwóch przestrzeniach nazw. Nie należy się tym martwić, przestrzenie nazw zapewniają nam separację w takich przypadkach. Węzeł `k:Autor` jest innym węzłem niż `a:Autor`.