

XML i nowoczesne technologie zarządzania treścią 2008/09 – Zadanie 1

Zadaniem jest opracowanie struktury dokumentów, w których będą przechowywane dane sklepu internetowego, i zapisanie jej w postaci schematu XML Schema. Ponadto należy dołączyć dokumentację i przykładowy dokument zgodny ze schematem.

Główną częścią dokumentu – "bazy danych" jest katalog towarów. Dokument powinien zawierać także kategorie towarów oraz wszystkie pojęcia, do których odwołują się towary.

Najważniejsze pojęcia

Towary

Towar charakteryzuje się następującymi cechami:

- kod unikalny w skali bazy danych, należy ograniczyć format kodów wyrażeniem regularnym (wedle własnego uznania)
- nazwa
- cena
- stawka VAT

Ponadto opcjonalnie mogą występować:

- opis (szczegóły dalej)
- specyfikacja techniczna (szczegóły dalej)
- producent
- zdjęcia lub grafiki:
 - lokalne – można albo umożliwić umieszczanie ich bezpośrednio w dokumencie (zakodowanych jako tekst), albo założyć, że są w zewnętrznych plikach w „domyślnym repozytorium”
 - zdalne – wskazane przez URL,
- informacje o gwarancji,
- cena promocyjna wraz z datami obowiązywania (od – do)

Kategorie

Kategoria to zbiór towarów, które chcemy jakoś ze sobą skojarzyć, np. w celu pogrupowanej prezentacji na stronach sklepu.

Kategoria posiada swoją krótką nazwę oraz opcjonalnie dłuższy opis. Kategorie mogą posiadać podkategorie.

Każdy towar może należeć do wielu kategorii, każda kategoria może zawierać wiele towarów, towary w obrębie kategorii nie powtarzają się.

Stawka VAT

Towar może mieć stawkę VAT podaną liczbowo (przy czym dopuszczalne wartości to 0, 3, 7 i 22 procent) lub być zwolniony z VAT.

Producent

Z każdym producentem, oprócz jego nazwy, może być związany opis i adres strony internetowej.

Dodatkowe wyjaśnienia

Opisy

Opis towaru, kategorii lub producenta powinien być tekstem z możliwością zanurzenia w nim dodatkowych elementów:

- odnośników do innych towarów oraz producentów i kategorii opisanych w tej samej "bazie danych",
- odnośników do miejsc w internecie,
- obrazków (te same zasady co dla obrazków bezpośrednio przypisanych do towarów – patrz wyżej),
- znaczników wpływających na format, co najmniej akapitów, list i wyróżnień.

Ponadto powinna być możliwość wyłączenia powtarzających się fragmentów opisów (dla różnych, ale podobnych towarów) na zewnątrz opisów i odwoływania się do nich za pomocą znaczników.

Tam, gdzie to adekwatne, można posłużyć się już istniejącym standardem, np. użyć elementów XHTML jako znaczników formatujących.

Specyfikacje techniczne

Towar może posiadać specyfikację techniczną, co ma pozwolić na zautomatyzowane wyszukiwanie i porównywanie towarów w sklepie. Różnego rodzaju towary mogą mieć różny zestaw cech opisywanych w specyfikacjach technicznych. Należy zaproponować elastyczne rozwiązanie.

Dodatkowe wymagania

Klucze i referencje

Korzystając z kluczy i referencji w XML Schema należy zapewnić, aby wszystkie wewnętrzne referencje w dokumencie były zamknięte. W uzasadnionych przypadkach można wprowadzić do różnego rodzaju pojęć sztuczne identyfikatory.

Przestrzenie nazw

Schemat powinien mieć określoną docelową przestrzeń nazw. W przypadku odwoływania się do elementów, typów itp. z istniejących standardów, należy korzystać ze standardu przestrzeni nazw.

Definiowane w schemacie elementy powinny mieć formę kwalifikowaną, a atrybuty niekwalifikowaną.

Struktura dokumentów

Dane powinny być zapisane atomowo, w sposób ułatwiający przetwarzanie przez programy, arkusze XSLT czy zapytania XQuery. Należy możliwie ściśle kontrolować typy danych korzystając z właściwych typów prostych i, w razie potrzeby, dodatkowo je zawężając.

W dokumentach należy unikać redundancji danych, np. opis producenta powinien być podany raz, a z towarów powinny być jedynie referencje do producentów. Nie należy jednak popadać w przesadę w odwzorowywaniu struktury relacyjnych baz danych, gdyż w XML w naturalny sposób można wewnątrz elementów umieszczać sekwencje podelementów i należy z tego korzystać.

Modularność schematu

Należy zastosować takie sposoby modularyzacji schematu (niekoniecznie wszystkie) jak typy nazwane, rozszerzanie i zawężanie typów, grupy elementów i atrybutów, aby takie same obiekty (elementy, atrybuty, modele zawartości) nie były definiowane wielokrotnie a schemat nadawał się do dalszych rozszerzeń i modyfikacji.

Dokumentacja i przykłady

Do schematu należy dołączyć dokumentację zawierającą zwięzły opis wszystkich elementów, atrybutów i typów wprowadzanych przez schemat (przede wszystkim ich znaczenie) oraz uzasadnienie podjętych nietrywialnych decyzji. Dopuszczalne formaty to PDF, (X)HTML, TXT.

Opis poszczególnych elementów, atrybutów, typów itp. może także być zawarty w samym schemacie w elementach `annotation`. W takim przypadku nie trzeba (ale można) kopiować tych opisów do zewnętrznej dokumentacji.

Należy także dołączyć przykładowy dokument ilustrujący zastosowanie poszczególnych konstrukcji (co najmniej 5 towarów, 3 kategorie w tym podkategoria, po jednym egzemplarzu każdego wprowadzonego elementu i atrybutu).

Informacje organizacyjne

W razie wątpliwości można pytać mailowo autora zadania: czarnik@mimuw.edu.pl

Proszę sprawdzać [liste pytań i odpowiedzi](#).

Rozwiązania (w postaci archiwum ZIP o nazwie równej loginowi studenckiemu) należy wysyłać do **24 listopada 2008** włącznie na adres: czarnik@mimuw.edu.pl z konta mailowego na `students`.