

Laboratorium 1 – Projektowanie języków XML

Standard XML

Co poznamy na zajęciach i do czego się to przyda?

Na laboratoriach przygotujemy środowisko pracy i poznamy standard XML. Jest to bardzo popularny i szeroko stosowany sposób (metajęzyk) do opisu danych. Tak jak wspomniano na wykładzie: ma prostą i przejrzystą składnię, jest elastyczny i rozszerzalny (można definiować własne znaczniki) i niezależny od platform. Każdy programista powinien znać język XML, bo wiele ze środowisk wykorzystuje lub działa w oparciu o XML.

Zadanie 1

Pobierz ze strony (<https://inf.ug.edu.pl/~gmadejsk/xml/docs/listastudentow.xml>) plik listastudentow.xml, najlepiej korzystając z opcji „Zapisz jako”, a nie kopiując kod. Otwórz plik w wybranym przez Ciebie edytorze tekstowym np.

- Gedit,
- Jedit,
- Notepad++

lub innym. Zwróć uwagę czy edytor tekstu ma następujące opcje, znacznie ułatwiające edycję plików XML:

- kolorowanie składni,
- dostępny podgląd i zmiana kodowania (UTF-8, ANSI),
- możliwość zwijania kodu (patrz rysunek)
- Automatyczne wcięcia (auto-indent)

```
1 <?xml version="1.0" encoding="utf-8"?>
2 <listaStudentow>
3 <student grupa="I" pozycjaWGrupie="10">
9 <student grupa="I" pozycjaWGrupie="11">
10 <nazwisko>Lipińska</nazwisko>
11 <imie>Ewa</imie>
12 <nrIndeksu>124</nrIndeksu>
13 <wydzialKod>MFI</wydzialKod>
14 </student>
15 </listaStudentow>
```

Zapoznaj się również z przydatnymi linkami dotyczącymi standardu XML. Odsyłam w szczególności do organizacji W3C, opiekującej się standardem XML:

- <https://www.w3.org/XML/> (wstęp)
- <https://www.w3.org/TR/REC-xml/> (specyfikacja, 5th edition)
- <http://www.w3schools.com/xml/> (przystępne przykłady, zalecam przejrzanie zakładek od „XML Home” do „XML Attributes”)

W Internecie znajdziesz wiele ciekawych tutoriali i źródeł uzupełniających wiedzę na temat przedmiotu. Zachęcam do googlania.

Zadanie 2

Mając pojęcie o składni XML z pewnością łatwo odpowiesz na następujące pytania dotyczące pliku listastudentow.xml:

- Podaj przykład elementu w tym pliku XML. Gdzie przechowywana jest nazwa elementu? Co jest treścią/zawartością elementu?
- Podaj przykład znacznika (tagu) początkowego i końcowego. Czym się różnią?
- Podaj przykład atrybutu: jego nazwę i jego wartość. Jak są pokolorowane w Twoim edytorze tekstowym?

Dodaj jeszcze jednego studenta do listy (wymyślone, ale sensowne dane) i zapisz plik XML.

Zadanie 3

Jak sprawdzić czy plik XML ma poprawną składnię czyli jest „well-formed” (więcej pod linkiem: http://www.w3schools.com/xml/xml_syntax.asp) ? Edytor tekstowy raczej nie ma możliwości wyłapania błędów w składni. Na szczęście łatwo to zrobić w inny sposób. Przetestuj poniższe opcje:

- Parser Online ze strony W3C: http://www.w3schools.com/xml/xml_validator.asp
- Otwarcie pliku w przeglądarce FireFox.
- Otwarcie pliku w innych przeglądarkach Opera / Chrome.
- Jeśli korzystasz z Linuxa to warto sprawdzić działanie komendy xmllint (<http://xmlsoft.org/xmllint.html>):
`xmllint ścieżka-do-pliku/listastudentow.xml`
lub
`xmllint --noout ścieżka-do-pliku/listastudentow.xml`

Zrób błąd w xml (np. robiąc literówkę w tagu domykającym) i przetestuj jak powyższe narzędzia radzą sobie z wykrywaniem tego błędu.

Zadanie 4

W pliku ksiazki.xml (<https://inf.ug.edu.pl/~gmadejsk/xml/docs/ksiazki.xml>) popełniono kilka błędów. Za pomocą wybranego parsera znajdź owe błędy, wskaż jakiego typu to błędy (http://www.w3schools.com/xml/xml_syntax.asp) i popraw je w sensowny sposób, tak by dokument XML nie zatracił sensu.

Informacje o błędach (komunikaty) oraz ich rodzaju możesz zapisać w zwykłym pliku tekstowym.

Zadanie 5

Stwórz plik sprzet.xml, zawierający dane o dwóch lub trzech urządzeniach biurowych (wybierz urządzenia jednego typu np. komputery, drukarki, monitory, maszyny ksero, laptopy). Dla przykładu dla komputera mogą to być dane o procesorze, karcie graficznej, itp. Niektóre dane będą liczbowe, inne tekstowe. Postaraj się, by dokument miał

sensowną strukturę, wypełnij go również sensownymi, niepustymi danymi.

Dodatkowe wymagania:

- Drzewo dokumentu (http://www.w3schools.com/xml/xml_tree.asp) musi być 4 poziomowe tzn. korzeń ma dzieci, wnuki i prawnuki.
- W dokumencie występują atrybuty.
- W dokumencie występują komentarze.

Zapisz i sprawdź składnię pliku za pomocą wybranego parsera.

Sprawdzenie

Po wykonaniu zadań poproś prowadzącego o sprawdzenie ich. Zwróć uwagę na następujące punkty:

- Jaki edytor tekstowy wybrałem? Czy znajdują się w nim wspomniane w zadaniu 1 funkcjonalności? Pokaż je.
- Odpowiedz na pytania z zadania 2.
- Który parser z zadania 3 przypadł Ci najbardziej do gustu? Wyjaśnij dlaczego.
- Pokaż jakie błędy znalazłeś po w pliku z zadania 4? Jak wygląda poprawiony plik?
- Zadeemonstruj plik z zadania 5.

Podsumowanie

Po dzisiejszych zajęciach powinieneś docenić dodatkowe opcje oferowane przez zaawansowane edytory tekstu (niektórzy zbyt przywiązują się do windowsowego notatnika!). Powinieneś też mieć ogólne pojęcie o strukturze plików XML, standardowych błędach składniowych jakie można popełnić oraz o narzędziach parsujących pliki XML.

Na kolejnych zajęciach...

Na kolejnych zajęciach zajmiemy się CSS (kaskadowe arkusze stylów). Osobom, które nie miały okazji poznać tej technologii polecam zapoznanie się z linkiem:

<http://www.w3schools.com/css/>