

XML – uniwersalny język wymiany informacji w e-świecie

Tomasz Traczyk

ttraczyk@ia.pw.edu.pl
http://www.ia.pw.edu.pl/~ttraczyk/



Politechnika Warszawska
Wydział Elektroniki i Technik Informacyjnych
Instytut Automatyki i Informatyki Stosowanej

Podstawy XML

Extensible Markup Language

- Służy do definiowania języków znakowania
 - do zapisu różnorodnych dokumentów i struktur danych
- Przeznaczony głównie do użycia w WWW

Geneza XML

- Podzbiór SGML (*Standard Generalised Markup Language*)
- Uproszczone przez usunięcie trudnych i niekoniecznych konstrukcji
- Dostosowane do łatwego przetwarzania w WWW
- Nieco podobny do HTML
- Standardy tworzone przez W3C (*World Wide Web Consortium*)



T.Traczyk-IBIS2001

2

Podstawy XML: dokument

Dokument XML

- Plik tekstowy: znaczniki, atrybuty
- Znakowanie znaczeniowe

Definicja składni dokumentu

- DTD (*Document Type Definition*)
- Schemat (*XML Schema*)
 - typy danych
- Przestrzeń nazw
 - unikanie konfliktów nazw
 - sformalizowane budowanie słowników

Poprawność dokumentu XML

- *Well-formed*
- *Valid*



T.Traczyk-IBIS2001

3

Podstawy XML: przykład

```
<?xml version="1.0" encoding="ISO-8859-2"?>
<!DOCTYPE konferencja SYSTEM "konferencja.dtd">
<?xml-stylesheet type="text/xsl" href="konferencja.xsl"?>

<konferencja id="IBIS01" rok="2001">
  <nazwa>Implementation of Business Information Systems</nazwa>
  <miesto>Malmö-Kopenhaga</miesto>
  <referaty>
    <referat id="r1" sesja="1" godzina="13:45">
      Referat <tytul>XML – uniwersalny język wymiany informacji w
      e-świecie</tytul>, którego autorem jest <autor>Tomasz Traczyk</autor>,
      mówi o tym, dlaczego język XML jest dobrym uniwersalnym środkiem do
      wymiany informacji w globalnym społeczeństwie informacyjnym.
    </referat>
    ...
  </referaty>
</konferencja>
```



T.Traczyk-IBIS2001

4

Użycie dokumentów XML

Walidacja

- Wykorzystanie DTD lub schematu przez standardowe analizatory leksykalne

Prezentacja

- Arkusze stylistyczne (*style-sheets*)
 - język XSL (*eXtensible Style Language*)
- Przeglądarki

Transformacje

- Język XSLT (*XSL Transformations*)

Przetwarzanie

- DOM (*Document Object Model*)
 - model obiektowy
 - silny, ale wymaga dużych zasobów
- SAX (*Simple API for XML*)
 - model zdarzeniowy
 - szybki



T.Traczyk-IBIS2001

5

XML w e-świecie

Języki

- Powszechnego użycia, np. WML, SVG, XHTML
- Branżowe/specialistyczne, np. MathML, CML
- Tworzone *ad hoc*

Główne zastosowania

- EDI (*Electronic Data Interchange*), wymiana danych w systemach heterogenicznych
- Gospodarka elektroniczna: B2C, B2B
- Komunikacja w sferze publicznej

Ważne cechy

- Sformalizowany zapis informacji
- Uniwersalność
- Duża siła wyrazu
- Elastyczność – rozszerzalność
- Możliwość zróżnicowanej prezentacji
- Czytelność
- Dostosowanie do specyfiki WWW
- Niewygodne koszty – standardowe oprogramowanie
- Względna prostota



T.Traczyk-IBIS2001

6

Podsumowanie

Zalety

- Liczne cechy odpowiednie dla uniwersalnego środka wymiany informacji
- Bardzo szerokie możliwości zastosowania
- Znaczna popularność
- Standardy tworzone przez organizację o dużym autorytecie
- Wsparcie ze strony przemysłu informatycznego
- Standardowe, często darmowe narzędzia

Problemy

- Szybki rozwój – zmienność standardów
- Nowe wyzwania dla twórców systemów informacyjnych, baz danych i narzędzi informatycznych

XML ma wielkie szanse stać się podstawowym i uniwersalnym środkiem wymiany informacji w e-świecie

