

Języki HTML, XML, XHTML

1. HTML (ang. HyperText Markup Language)

Podstawowym językiem projektowania stron internetowych jest HTML. W kolejnych wersjach język HTML został wzbogacony o mechanizmy: korzystania z zewnętrznych arkuszy stylów, pisanie skryptów, wstawiania obiektów, animowania tekstów. Obecnie najczęściej stosowaną specyfikacją jest HTML w wersji 5.

Podstawowym elementem języka HTML jest znacznik. Znaczniki są poleceniami umieszczonymi w nawiasach ostrokątnych, np.: `<body>`, `<head>`. Nazwy poleceń mogą być pisane małymi lub wielkimi literami, nie ma to znaczenia. Wyróżnia się znaczniki otwierające, np. `<p>`, i zamykające, np. `</p>`.

Prawie wszystkie znaczniki otwierające trzeba zamknąć, stosując znaczniki zamykające.

Wyjątki to np.:

- `<br>` — znacznik łamania wiersza,
- `<hr>` — znacznik linii poziomej,
- `<img>` — znacznik wstawiania obrazka.

Warto jednak pamiętać o tym, że znaczniki, które formalnie nie posiadają oddzielnego elementu domykającego, możemy zapisać w taki sposób, aby to domknięcie się pojawiło. Wymóg taki pojawił się w języku **XHTML**, ale zapis taki został również dopuszczony w HTML5 i warto go stosować. Wspomniany zapis od czasu do czasu jest używany w niniejszym podręczniku. Dla wskazanych powyżej wyjątków może on mieć następującą postać:

- `<br />`,
- `<hr />`,
- `<img />`.

Znaczniki mają dodatkowe elementy, zwane atrybutami, które definiują ich działanie, np.:

`<nazwa_znacznika atrybut=wartość> zawartość </nazwa_znacznika>`

lub

`<nazwa_znacznika atrybut="wartość"> zawartość </nazwa_znacznika>`

W języku HTML znaczniki mogą być zagnieżdżane, czyli umieszczane w innych znacznikach. Jednak nie wszystkie znaczniki mogą zawierać w sobie inne znaczniki.

W dokumencie tworzonym w języku HTML powinna być wstawiona deklaracja typu dokumentu (ang. *Document Type Definition* — DTD). Określa ona, która wersja języka HTML jest stosowana na stronie.

Deklaracja w HTML5:

`<!DOCTYPE html>`

Deklaracja w HTML 4.01:

`<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN" "http://www.w3.org/TR/html4/strict.dtd">`

2. XML (ang. Extensible Markup Language)

XML jest metajęzykiem opartym na zwykłym tekście, przeznaczonym do definiowania języków znakowania. Jest to coś więcej niż tekstowy format do opisywania dokumentów. XML daje dostęp do wielu technologii konstruowania, przetwarzania i przeszukiwania danych. Został stworzony w celu definiowania nowych formatów dokumentów w internecie.

Ze względu na to, że składnia XML jest oparta na zwykłym tekście, dokumenty XML mogą być czytane i edytowane przy użyciu najprostszych edytorów tekstu.

XML nie jest związany z żadnym językiem programowania ani z żadnym systemem operacyjnym. Dzięki temu dokumenty można tworzyć i edytować, korzystając z różnych języków programowania, a XML może współpracować z różnymi systemami operacyjnymi.

Dokument w XML1.0 jest zbudowany z elementów, a każdy element składa się ze znacznika otwierającego (np. `<table>`), znacznika zamykającego (np. `</table>`) oraz z informacji znajdujących się pomiędzy tymi

znacznikami. Elementy mogą być opatrywane atrybutami zawierającymi metadane opisujące element i jego zawartość.

3. XHTML (ang. Extensible HyperText Markup Language)

Język XHTML to język, który dostosowuje zasady języka HTML4 do specyfikacji języka XML. Opiera się na specyfikacji języka HTML 4.01, ale równocześnie wprowadza obostrzenia wynikające ze składni języka XML.

W XHTML jest stosowana taka sama składnia języka jak w HTML, ale zostały zdefiniowane bardziej rygorystyczne zasady, które dla projektantów stron i twórców przeglądarek internetowych stały się obowiązującym standardem:

- znaczniki muszą być zamykane w kolejności odwrotnej do tej, w której były otwierane,
- nazwy znaczników i atrybutów muszą być pisane małymi literami,
- znaczniki zamykające są obowiązkowe,
- wartości atrybutów muszą być ujęte w cudzysłowy,
- nie można pomijać atrybutów logicznych,
- elementy puste muszą mieć znacznik zamykający lub ich znacznik otwierający musi się kończyć sekwencją `</>`, np. `<br />`.

Do dokumentu utworzonego w języku XHTML musi zostać wstawiona deklaracja typu dokumentu (DTD). Określa ona, która wersja języka HTML jest stosowana na stronie.

W wersji XHTML 1.1 stosowana jest tylko jedna wersja deklaracji:

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1//EN" "http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd">
```

Deklaracja XML zarówno dla wersji XHTML 1.0, jak i dla wersji XHTML 1.1 ma postać:

```
<?xml version="1.0" encoding="UTF-8"?>
```

4. HTML5

Język HTML5 jest rozwinięciem języków HTML4 i XHTML1. Jest on kompatybilny z poprzednimi wersjami języka HTML.

Standard HTML5 upraszcza zapis deklaracji kilku elementów.

Deklaracja typu dokumentu ma postać:

```
<!DOCTYPE html>
```

Sposób kodowania znaków ma postać:

```
<meta charset="UTF-8" />
```

Dołączenie arkusza stylów nie wymaga atrybutu type:

```
<link rel="stylesheet" href="file . css">
```

Deklaracja JavaScript również nie wymaga atrybutu type:

```
<script src="file.js"></script>
```

Zgodnie z zasadami obowiązującymi w HTML5 następuje rozdzielenie struktury strony, za którą odpowiada HTML, od jej wizualizacji, za którą to odpowiadają arkusze stylów CSS. Układ stron buduje się za pomocą specjalnych pojemników (znacznik `<div>` w HTML4), tzw. pojemników semantycznych (znaczniki `<header>`, `<nav>`, `<footer>`, `<section>`, `<article>`, `<aside>`), które są formatowane przy użyciu stylów CSS. Nowe elementy nie tylko poprawiają strukturę strony internetowej, powodując, że struktura ta staje się zrozumiała i logiczna.