helion.pl
KSIĘGARNIA INTERNETOWA · WYDAWNICTWO

IDŹ DO
PRZYKŁADOWY ROZDZIAŁ
SPIS TREŚCI

KATALOG KSIĄŻEK
KATALOG ONLINE
ZAMÓW DRUKOWANY KATALOG

TWÓJ KOSZYK
DODAJ DO KOSZYKA

CENNIK I INFORMACJE
ZAMÓW INFORMACJE O NOWOŚCIACH
ZAMÓW CENNIK

CZYTELNIA
FRAGMENTY KSIĄŻEK ONLINE

HTML i XHTML. Leksykon kieszonkowy



Autor: Jennifer Niederst Robbins
Tłumaczenie: Grzegorz Werner
ISBN: 83-246-0633-5
Tytuł oryginału: HTML and XHTML Pocket Reference (Pocket Reference)
Format: B6, stron: 128

Język HTML, mimo że powstał niedawno, przeszedł już długą drogę rozwoju. Początkowo wykorzystywany jedynie do formatowania dokumentów publikowanych w sieci łączącej ośrodki naukowe, dziś jest zestawem znaczników stosowanych do tworzenia stron WWW. Czuwające nad jego rozwojem konsorcjum W3C co pewien czas publikuje specyfikację określającą składnię znaczników i ich atrybutów. Najnowsza specyfikacja nosi numer 4.01. Język XHTML to zawierająca dodatkowe elementy specyfikacja HTML, określona zgodnie z wymaganiami XML.

„HTML i XHTML. Leksykon kieszonkowy" to zestawienie znaczników HTML — wraz z atrybutami i ich dopuszczalnymi wartościami – które jest niezwykle przydatne każdemu projektantowi witryn WWW. Opisuje najnowsze specyfikacje HTML i XHTML, wskazuje podobieństwa i różnice pomiędzy nimi, a także zawiera praktyczne informacje, dzięki którym Twoja witryna WWW zostanie zaprojektowana zgodnie z rekomendacjami W3C.

W książce znajduje się:

- Struktura dokumentów HTML i XHTML
- Alfabetyczny spis znaczników HTML i XHTML
- Atrybuty i ich wartości
- Przykłady zastosowania znaczników
- Deklaracje DOCTYPE
- Encje znakowe
- Kody kolorów

Twórz witryny WWW według najnowszych standardów

Wydawnictwo Helion
ul. Kościuszki 1c
44-100 Gliwice
tel. 032 230 98 63
e-mail: helion@helion.pl

Spis treści

Podstawy HTML i XHTML 5
Czym XHTML różni się od HTML? 6
Trzy wersje (X)HTML 8
Minimalna struktura dokumentu 9
Deklaracje DOCTYPE dla dostępnych definicji typów dokumentu 10

Alfabetyczna lista elementów 11
Często używane atrybuty i zdarzenia 12
Elementy (X)HTML 15

Encje znakowe 104
Zestaw znaków ASCII 105
Encje niestandardowe (‚ – Ÿ) 107
Latin-1 (ISO-8859-1) 107
Latin Extended-A 112
Latin Extended-B 113
Znaki modyfikujące odstępy 113
Litery greckie 113
Interpunkcja 117
Symbole pseudoliterowe 118
Strzałki 119
Operatory matematyczne 119
Różne symbole techniczne 121
Kształty geometryczne 122
Symbole różne 122

Określanie koloru 122
Wartości RGB 123
Standardowe nazwy kolorów 124

Skorowidz 125

HTML i XHTML
Leksykon kieszonkowy

Niniejszy leksykon kieszonkowy zawiera zwięzły, ale wyczerpujący spis elementów i atrybutów zdefiniowanych w rekomendacjach HTML 4.01 oraz XHTML 1.0. W tekście użyto skrótu „X(HTML)" na oznaczenie pojęć dotyczących obu standardów znakowania.

Książka dzieli się na następujące części:

- podstawy HTML i XHTML;
- alfabetyczna lista elementów;
- encje znakowe;
- określanie koloru.

Podstawy HTML i XHTML

HTML (*HyperText Markup Language*, hipertekstowy język znakowania) to język używany do przekształcania dokumentów w strony WWW. Zasadniczym celem HTML jako języka znakowania jest zapewnienie **semantycznego** opisu (znaczenia) treści oraz ustanowienie **struktury** dokumentu (hierarchii elementów). HTML nie uwzględnia **prezentacji**, czyli tego, jak dokument będzie wyglądał w przeglądarce. Prezentacja jest zadaniem języka CSS

(*Cascading Style Sheets*, kaskadowe arkusze stylów), którego opis wykracza poza ramy niniejszej książki.

XHTML (*eXtensible HyperText Markup Language*, rozszerzalny hipertekstowy język znakowania) jest reformulacją HTML 4.01 według ściślejszych reguł składniowych XML (*eXtensible Markup Language*, rozszerzalny język znakowania). Elementy są te same, ale standard narzuca dodatkowe ograniczenia na znakowanie dokumentu (wymienione w następnym podrozdziale). Standard XHTML składa się z rodziny rekomendacji, w tym XHTML 1.0, XHTML 1.1, Modularization of XHTML oraz XHTML Basic; w przygotowaniu są także inne wersje. Niniejsza książka skupia się na rekomendacjach XHTML 1.0. Więcej informacji o wszystkich wersjach HTML i XHTML można znaleźć w witrynie World Wide Web Consortium pod adresem *www.w3.org/MarkUp*.

Czym XHTML różni się od HTML?

Ponieważ XHTML jest językiem XML, jego składnia jest bardziej ścisła i różni się od HTML pod następującymi kluczowymi względami:

- Wszystkie elementy i atrybuty muszą być pisane małymi literami. Na przykład `<a href="example.com">...</a>`.
- Wszystkie elementy muszą być zakończone, tzn. zawierać znacznik końcowy. Na przykład `<p>...</p>`.
- Elementy puste również muszą być zakończone. Polega to na dołączeniu ukośnika na końcu znacznika. Przed ukośnikiem często dodaje się spację, aby zapewnić zgodność ze starszymi przeglądarkami. Na przykład `<hr />`, `<img />`, `<meta />`.

- Wszystkie wartości atrybutów muszą być umieszczone w cudzysłowie (podwójnym lub pojedynczym). Na przykład `<td colspan="2">`.
- Wszystkie wartości atrybutów muszą być jawne i nie mogą być zminimalizowane do jednego słowa, co jest dozwolone w HTML. Na przykład:
  - `checked="checked"`;
  - `selected="selected"`;
  - `multiple="multiple"`.
- Ściślej egzekwowane są ograniczenia dotyczące zagnieżdżania. Standard określa następujące ograniczenia:
  - element `a` nie może zawierać innego elementu `a`;
  - element `pre` nie może zawierać elementów `img`, `object`, `applet`, `big`, `small`, `sub`, `sup`, `font` ani `basefont`;
  - element `form` nie może zawierać innych elementów `form`;
  - element `button` nie może zawierać elementów `a`, `form`, `input`, `select`, `textarea`, `label`, `button`, `iframe` ani `isindex`.
- Znaki specjalne `<`, `>`, `&`, `'` oraz `"` muszą zawsze być reprezentowane przez encje znakowe (nawet jeśli pojawiają się w wartościach atrybutów), odpowiednio: `<`, `>`, `&`, `'` oraz `"`.
- W HTML można używać atrybutu `name` w elementach `a`, `applet`, `form`, `frame`, `iframe`, `img` oraz `map`. Atrybuty `name` oraz `id` służą do identyfikowania fragmentów dokumentu. W dokumentach XHTML we wspomnianych wyżej elementach

zamiast `name` należy używać `id`. Zastosowanie atrybutu `name` w tych elementach uznano za przestarzałe w specyfikacji XHTML 1.0.

Trzy wersje (X)HTML

Zarówno rekomendacje HTML 4.01, jak i XHTML obejmują trzy nieco inne dokumenty specyfikacyjne określane mianem **definicji typu dokumentu** (*Document Type Definition*, DTD). Dokumenty DTD definiują każdy element, atrybut i encję, a także reguły ich stosowania. Trzy wersje DTD to:

Strict DTD (wersja ścisła)
: Ta wersja pomija wszystkie elementy i atrybuty uznane za **przestarzałe** (takie jak `font` i `align`), aby podkreślić odrębność struktury i prezentacji dokumentu.

Transitional DTD (wersja przejściowa)
: Ta wersja obejmuje wszystkie przestarzałe elementy i atrybuty, aby zachować zgodność z większością tradycyjnych przeglądarek. Przestarzałe elementy i atrybuty są dozwolone, ale należy ich unikać.

Frameset DTD (wersja z obsługą ramek)
: Ta wersja obejmuje te same elementy co Transitional DTD, a ponadto elementy przeznaczone do tworzenia ramek (`frameset`, `frame` oraz `noframe`). Wersję Frameset DTD oddzielono od pozostałych, ponieważ struktura dokumentu podzielonego na ramki (w którym element `frameset` zastępuje `body`) jest zasadniczo różna od zwykłych dokumentów HTML.

Uwaga

Rekomendacja XMTL 1.1 obejmuje tylko jedną definicję DTD, która przypomina wersję Strict w tym sensie, że nie uwzględnia przestarzałych elementów i atrybutów.

W następnym podrozdziale wyjaśniono, jak wskazuje się definicje DTD w dokumentach (X)HTML.

Minimalna struktura dokumentu

Poniższa próbka znakowania pokazuje minimalną strukturę dokumentu HTML 4.01. W przykładzie zastosowano ścisłą definicję typu dokumentu HTML (Strict DTD):

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

  <body>
    Treść dokumentu...
  </body>
</html>
```

Poniższa próbka znakowania pokazuje minimalną strukturę dokumentu XHTML 1.0 określoną w rekomendacji XHTML 1.0. Dokument napisano z wykorzystaniem przejściowej definicji typu dokumentu XHMTL (Transitional DTD):

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD XHTML 1.0
Transitional//EN"
   "http://www.w3.org/TR/xhtml1/DTD/xhtml1-
   transitional.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en"
lang="en">
  <head>
    <title>Tytuł dokumentu</title>
  </head>

  <body>
    Treść dokumentu...
  </body>
</html>
```

W obu przykładach pierwszym wierszem jest **deklaracja typu dokumentu** (*Document Type Declaration*), zwana też deklaracją DOCTYPE, która określa wersję DTD używaną w dokumencie. Służy ona do sprawdzania poprawności dokumentu. Niektóre przeglądarki na podstawie kompletnej deklaracji DOCTYPE przełączają się w tryb renderowania zgodny ze standardami. Zauważmy, że dokument XHTML używa innej deklaracji DOCTYPE niż HTML i zawiera identyfikację przestrzeni nazw XML oraz języka w głównym elemencie `html`.

Dokumenty XHMTL mogą opcjonalnie zawierać deklarację XML przed deklaracją DOCTYPE, jak pokazano w poniższym przykładzie:

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE HTML PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
   "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
```

Deklaracja XML nie jest wymagana, jeśli kodowaniem znaków jest UTF-8. Ponieważ obecnie deklaracje XML sprawiają problemy nawet przeglądarkom respektującym standardy, zwykle się je pomija.

Deklaracje DOCTYPE dla dostępnych definicji typów dokumentu

Deklaracja `<!DOCTYPE>` (typu dokumentu) obejmuje dwie metody wskazywania informacji DTD: jednym jest publiczny identyfikator dokumentu, a drugim konkretny adres URL na użytek przeglądarek, które nie rozpoznają identyfikatora publicznego. Poniżej wymieniono dokładne deklaracje DOCTYPE dla każdej wersji HTML i XHMTL:

HTML 4.01 Strict

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"
 "http://www.w3.org/TR/HTML4.01/strict.dtd">
```

HTML 4.01 Transitional

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01
Transitional//EN"
 "http://www.w3.org/TR/HTML4.01/loose.dtd">
```

HTML 4.01 Frameset

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"
 "http://www.w3.org/TR/HTML4.01/frameset.dtd">
```

XHTML 1.0 Strict

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
 "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
```

XHTML 1.0 Transitional

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD XHTML 1.0
Transtitional//EN"
 "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
```

XHTML 1.0 Frameset

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD XHTML 1.0 Frameset//EN"
 "http://www.w3.org/TR/xhtml1/DTD/xhtml1-frameset.dtd">
```

XHTML 1.1

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

Alfabetyczna lista elementów

Ten rozdział zawiera spis wszystkich elementów i atrybutów wymienionych w rekomendacjach HTML 4.01 oraz XHTML, a także kilku elementów niestandardowych.

Czytelnicy powinni zwracać uwagę na poniższe oznaczenia elementów i atrybutów:

Przestarzały

Elementy oznaczone słowem „**przestarzały**" są obecnie eliminowane z HTML — zwykle na korzyść CSS — i nie należy ich używać. Wszystkie przestarzałe elementy i atrybuty zostały usunięte z wersji Strict HTML i XHMTL.

Wymagany
: Atrybuty oznaczone słowem „**wymagany**" muszą być dołączone do elementu, aby znakowanie było prawidłowe.

Tylko XHTML
: Atrybuty oznaczone słowami „**tylko XHMTL**" dotyczą tylko dokumentów oznakowanych w XHTML. W podobny sposób oznaczone są inne drobne różnice między HMTL a XHTML.

Niestandardowy
: **Niestandardowe** elementy i atrybuty nie są uwzględnione w żadnej wersji rekomendacji HTML i XHTML, ale są obsługiwane przez większość przeglądarek.

Często używane atrybuty i zdarzenia

Niektórych atrybutów HTML 4.01 i XHTML 1.0 można używać w niemal wszystkich elementach. Aby zaoszczędzić miejsce, w niniejszym leksykonie skrócono je, podobnie jak w oficjalnych rekomendacjach. W niniejszym podrozdziale wyjaśniono skrótową notację atrybutów, która będzie używana w reszcie spisu elementów.

Podstawowe

Jeśli na liście atrybutów znajduje się słowo „**podstawowe**", odnosi się ono do zbioru podstawowych atrybutów, które mogą występować w większości elementów (jak zaznaczono w opisie każdego elementu):

`id`
: Przypisuje elementowi unikatową, identyfikującą go nazwę.

`class`
: Przypisuje elementowi jedną lub więcej klasyfikujących go nazw.

`style`
: Wiąże z elementem informacje o stylu.

`title`
: Podaje tytuł albo dodatkowe informacje o elemencie.

Internacjonalizacja

Jeśli na liście atrybutów znajduje się słowo „**internacjonalizacja**", oznacza to, że element akceptuje atrybuty określające język i kierunek pisma:

`dir`
: Określa kierunek elementu (od lewej do prawej lub od prawej do lewej).

`lang`
: Określa język elementu za pomocą kodu języka.

`xml:lang`
: ***Tylko XHTML***. Określa język elementu w dokumentach XHTML.

Zdarzenia

Jeśli na liście atrybutów znajduje się słowo „**zdarzenia**", oznacza to, że do elementu odnoszą się podstawowe zdarzenia używane przez języki skryptowe. Dodatkowe zdarzenia (spoza zbioru zdarzeń podstawowych) są wymienione oddzielnie:

`onclick`
: Występuje, kiedy nad elementem zostanie kliknięty przycisk urządzenia wskazującego.

`ondblclick`
: Występuje, kiedy nad elementem zostanie dwukrotnie kliknięty przycisk urządzenia wskazującego.

`onkeydown`
: Występuje, kiedy nad elementem zostanie wciśnięty klawisz.

`onkeypress`
: Występuje, kiedy nad elementem zostanie wciśnięty i zwolniony klawisz.

`onkeyup`
: Występuje, kiedy nad elementem zostanie zwolniony klawisz.

`onmousedown`
: Występuje, kiedy nad elementem zostanie wciśnięty przycisk urządzenia wskazującego.

`onmousemove`
: Występuje, kiedy wskaźnik zostanie przesunięty w obrębie elementu.

`onmouseout`
: Występuje, kiedy wskaźnik zostanie odsunięty znad elementu.

`onmouseover`
: Występuje, kiedy wskaźnik zostanie przesunięty nad element.

`onmouseup`
: Występuje, kiedy nad elementem zostanie zwolniony przycisk urządzenia wskazującego.

Fokus

Termin **fokus** odnosi się do stanu, w którym element (taki jak łącze albo element formularza) jest wyróżniony i gotowy do reakcji na czynności użytkownika. Jeśli na liście atrybutów znajduje się słowo „fokus", oznacza to, że do elementu odnoszą się poniższe atrybuty i zdarzenia związane z obsługą fokusu:

`accesskey="znak"`
: Przypisuje łączu klawisz dostępowy (skrót klawiaturowy). Klawiszy dostępowych używa się również w polach formularza. Wartością tego atrybutu jest pojedynczy znak. Użytkownicy mogą uzyskać dostęp do elementu, naciskając klawisze *Alt+znak* (PC) albo *Ctrl+znak* (Mac).

`onblur`
: Występuje, kiedy element utraci fokus wskutek naciśnięcia przycisku na urządzeniu wskazującym albo klawisza tabulacji.

`onfocus`
: Występuje, kiedy element uzyska fokus wskutek naciśnięcia przycisku na urządzeniu wskazującym albo klawisza tabulacji.

`tabindex="liczba"`
: Określa pozycję bieżącego elementu w kolejności tabulacji bieżącego dokumentu. Wartość musi należeć do zakresu od 0 do 32 767. Atrybut ten pozwala wykorzystać klawisz tabulacji do przełączania się między łączami na stronie (albo polami formularza).

Elementy (X)HTML

a

`<a>...</a>`

Definiuje kotwicę w dokumencie. **Kotwica** służy do tworzenia hiperłączy do innych dokumentów albo zasobów internetowych. Może również oznaczać docelowy fragment dokumentu (określa się to mianem kotwicy nazwanej) i pozwalać na przeskakiwanie do tego fragmentu.

Atrybuty

Podstawowe, internacjonalizacja, zdarzenia, fokus

`charset="zestaw znaków"`
: Określa kodowanie znaków docelowego dokumentu.

`coords="współrzędne x, y"`
: Określa współrzędne *x* i *y* aktywnego obszaru mapy obrazkowej. Rekomendacja HTML 4.01 proponuje, aby mapy obrazkowe obsługiwane przez klienta zastąpić elementem `object` zawierającym obraz oraz zbiorem kotwic definiujących „gorące" obszary (za pomocą atrybutów kształtu i współrzędnych). Ten system nie jest jeszcze obsługiwany przez przeglądarki.

`href="URL"`
: Określa adres URL docelowego dokumentu albo zasobu WWW (takiego jak obraz, dźwięk, plik PDF albo inny plik multimedialny).

`hreflang="kod języka"`
: Określa podstawowy język docelowego dokumentu.

`id="tekst"`
: Nadaje łączu unikatową nazwę (podobnie jak element `name`), aby można się było do niego odwoływać z innego łącza, skryptu albo arkusza stylów. W XHTML atrybut `id` jest wymagany do identyfikacji fragmentów dokumentu. W celu zachowania zgodności ze starszymi przeglądarkami autorzy oznaczają fragmenty zarówno atrybutem `name`, jak i `id`.

`name="tekst"`
: ***Tylko HTML***. W XHTML fragmenty dokumentów oznacza się atrybutem `id`. Umieszcza identyfikator fragmentu w dokumencie HTML.

`rel="relacje"`
: Ustanawia jedną lub więcej relacji między dokumentem źródłowym (bieżącym) a docelowym. Do często używanych relacji należą `stylesheet`, `next`, `prev`, `copyright`, `index` oraz `glossary`.

`rev="relacje"`
: Ustanawia jedną lub więcej relacji między dokumentem docelowym a źródłowym (odwrotność atrybutu `rel`).

`shape="rect|circle|poly|default"`
: Definiuje kształt aktywnego obszaru mapy obrazkowej. Rekomendacja HTML 4.01 proponuje, aby zastąpić mapy obrazkowe obsługiwane przez klienta kombinacją elementów `object` oraz `a`. Ten system nie jest jeszcze obsługiwany przez przeglądarki.

`target="tekst"`
: Określa nazwę okna lub ramki, w której powinien zostać wyświetlony dokument docelowy.

`type="typ treści"`
: Określa typ treści (typ MIME) docelowego dokumentu, na przykład `text/html`.

Przykłady

Do lokalnego pliku:

```
<a href="nazwapliku.html">...</a>
```

Do zewnętrznego pliku:

```
<a href="http://serwer/sciezka/plik.html">...</a>
```

Do nazwanej kotwicy:

```
<a href="http://serwer/sciezka/plik.html#fragment">...</a>
```

Do nazwanej kotwicy w bieżącym pliku:

```
<a href="#fragment">...</a>
```

Do wysłania wiadomości e-mail:

```
<a href="mailto:uzytkownik@domena">...</a>
```

Do pliku na serwerze FTP:

```
<a href="ftp://serwer/sciezka/nazwapliku">...</a>
```

Tworzenie nazwanej kotwicy w HTML:

```
<a name="fragment">...</a>
```

Tworzenie nazwanej kotwicy w XHTML (niektórzy autorzy dołączają nadmiarowy element `name`, aby zachować zgodność z przeglądarkami w wersji 4.):

```
<a id="fragment">...</a>
```

abbr

`<abbr>...</abbr>`

Identyfikuje zawarty w nim tekst jako skrót.

Atrybuty

Podstawowe, *internacjonalizacja*, *zdarzenia*

`title="`*`tekst`*`"`

Określa pełne rozwinięcie skrótu. Może to być przydatne w przeglądarkach niewizualnych, syntezatorach mowy, systemach translacji i wyszukiwarkach.

Przykład

```
<abbr title="Massachussetts">Mass.</abbr>
```