

Rozdział 1. Informacje podstawowe

Książka poświęcona jest podstawom języka HTML – HyperText Mark-up Language. Utworzą one solidny fundament, który pozwoli Ci budować nawet skomplikowane strony WWW. Jeśli absolutnie nie masz pojęcia o znacznikach i wyrażeniach przez nie tworzonych, tutaj dowiesz się tego wszystkiego, co w niezbędne.

Zaawansowani twórcy stron WWW także powinni znaleźć w książce swego pomocnika – omawia ona bowiem poszczególne znaczniki (a pamięć bywa zawodna!) oraz prezentuje szereg przydatnych przykładów ich zastosowania.

Zacznijmy od A

Co będzie Ci potrzebne?

- ❑ Komputer – to oczywiste.
- ❑ Przeglądarka – na przykład Internet Explorer lub Netscape Navigator (lub inna, najlepiej jednak graficzna, czyli taka, która obsługuje obrazy).
- ❑ Edytor plików tekstowych – wystarczy nam prosty edytor typu Notepad lub Wordpad, polecam jednak bezpłatny edytor Webber, który ma tę zaletę, że udostępnia numerację wierszy (a jak się przekonasz, jest to rzecz niezwykle przydatna) oraz automatycznie dokonuje konwersji kodów polskich znaków diakrytycznych ze standardu Windows na standard używany w Internecie. Edytor ten znajdziesz na stronie WWW Radka Sokoła <http://www.magsoft.com.pl/~rsokol/>.

Brak połączenia do Internetu, w żaden sposób nie ogranicza Twoich możliwości tworzenia kodu HTML. Fazy pisania i wstępnego testowania kodu można przeprowadzać bez łączenia się z Internetem.

Czy HTML to język programowania?

HTML – HyperText Mark-up Language nie jest językiem programowania, takim jak C++ czy Pascal. A czym jest? **Hyper** oznacza, że HTML nie jest językiem liniowym. Języki programowania mają zazwyczaj strukturę liniową – innymi słowy struktura programu określa kolejność jego wykonania. W przypadku strony WWW użytkownik ma możliwość skorzystania z dowolnej opcji strony w dowolnym czasie. Drugi człon nazwy, **Text**, to narzędzie, którym się posługujesz, a jest nim zwykły (no, może nie całkiem zwykły) tekst. „Niezwykłość” tekstu bierze się stąd, że umieszczony jest on w specjalnych znacznikach, „<>”, „</>”, i one właśnie są odpowiedzialne za kolejny człon nazwy: **Mark-up**. Dzięki znacznikom przeglądarka wie, że ma do czynienia z kodem strony WWW, a nie ze zwykłym tekstem. Znaczniki nadają specjalne atrybuty i instruują przeglądarkę jak interpretować wygląd tekstu. Format znacznika jest następujący:

`<ZNACZNIK>To jest zawartość</ZNACZNIK>` (jest parę wyjątków, na przykład przy wstawianiu obrazu niezbędny jest tylko znacznik otwierający).

Ostatnia litera w skrócie HTML to L – od Language, co oznacza „język”. HTML jest bowiem językiem publikowania w sieci WWW, a co najważniejsze, językiem niezależnym od stosowanej platformy sprzętowej, czyli rodzaju komputera, z jakiego korzystasz.

Co HTML umożliwia?

HTML pozwala:

- ❑ Publikować dokumenty zawierające nagłówki, tekst, tabele, listy, zdjęcia, formularze, itp.
- ❑ Pobierać za pośrednictwem „połączeń hipertekstowych” informacje z Internetu (mówiąc w skrócie polega to na klikaniu „połączeń”, co uaktywnia pobieranie danych).
- ❑ Projektować formularze pozwalające na korzystanie ze zdalnych usług, takich jak na przykład wyszukiwanie informacji, dokonywanie rezerwacji oraz zamawianie produktów.
- ❑ Umieszczać w dokumentach arkusze kalkulacyjne, klipy wideo oraz inne aplikacje.

Jak HTML działa?

Aby lepiej zrozumieć funkcjonowanie języka HTML, spróbuj potraktować znaczniki HTML jako przymiotniki. Kolor obiektu, na przykład piłki, można opisać korzystając z takich przymiotników jak *niebieska*, *czerwona* czy *żółta*. Relację względem innego obiektu można z kolei wyrazić za pomocą słów takich jak: *z lewej*, *z prawej*, *w centrum*. Podobnie HTML pomaga opisać wygląd i ulokowanie tekstu, obrazów i innych obiektów na Twojej stronie WWW.

Od historii do teraźniejszości

Kto zaczął i gdzie to zmierza?

Twórcą języka HTML jest Tim Berners-Lee. Język pomyślany jako narzędzie ułatwiające komunikowanie się naukowców w sieci szybko zyskał ogromną popularność. Przyniósł ją w latach

90-tych szalony rozwój Internetu. Wraz z rozwojem sieci globalnej zmieniał się też sam język. W 1995 roku zatwierdzono specyfikację HTML 2, natomiast specyfikacja HTML 3, także z tego roku, nie doczekała się rekomendacji. W 1996 roku powstała specyfikacja HTML 3.2, która uzyskała rekomendację w 1997 roku. W kolejnych specyfikacjach wprowadzano zmiany, których zadaniem było poprawić przenośność dokumentów HTML między różnymi platformami i przeglądarkami. Zdano sobie bowiem sprawę, że jedynie pełna kompatybilność zagwarantuje prawidłowy rozwój Internetu, a jej brak będzie oznaczał lawinę formatów, która utrudni lub wręcz uniemożliwi powszechne wykorzystanie sieci. Wprowadzane modyfikacje rozszerzały potencjał języka nie naruszając jego dotychczasowych możliwości. Chodziło o to, aby już utworzone strony były nadal dostępne.

Inne ważne kwestie, które zadecydowały o kierunku rozwoju to potrzeba umożliwienia korzystania z informacji sieciowych za pomocą różnych urządzeń, takich jak telefony komórkowe czy komputerowe urządzenia służące do komunikacji głosowej, oraz uniezależnienia się od konfiguracji sprzętowej (a więc na przykład od ustawień takich jak rozdzielczość czy głębokość barw) i parametrów łączy sieciowych.

Wersja 4 (z 1998 r.) rozszerzyła możliwości języka HTML o arkusze stylów, skrypty, ramki i osadzanie obiektów. Poprawiono obsługę tekstu, rozbudowano opcje tabel i formularzy wprowadzając ułatwienia dla osób niepełnosprawnych.

W wersji 4.01, wprowadzonej w 1999 roku, poprawiono błędy i wprowadzono drobne zmiany.

Kto dba o standaryzację?

Definiowaniem standardu języka HTML zajmuje się World Wide Web Consortium (W3C). Konsorcjum W3C utworzone zostało w 1994 roku przez 500 organizacji. Dba ono o rozwój technologii, wyznacza kierunki rozwoju, tworzy i zatwierdza specyfikacje, udostępnia oprogramowanie i narzędzia, a wszystko to w celu wyzwolenia pełnych możliwości Internetu jako forum wymiany informacji i usług oraz centrum komunikacji między ludźmi.

Czy W3C przewiduje opublikowanie kolejnej specyfikacji, HTML 5.0? I tak, i nie. Na początku 2000 roku pojawiła się specyfikacja XHTML 1.0 (eXtensible Hypertext Markup Language) – oznacza się ją także jako xHTML, Xhtml, XML/HTML lub właśnie jako HTML 5.0. XHTML, wyznaczony na „oficjalnego” następcę języka HTML 4.0, to hybryda językowa wyrastająca z języka XML, posiadająca jednocześnie wiele cech wspólnych z językiem HTML.

A co to takiego ten XML?

XML – eXtensible Markup Language – to rodzina technologii, służąca do definiowania formatu i struktury dokumentów. Stanowi ona uproszczoną wersję języka SGML, a dokładniej jego podzbiór. No tak, z deszczu pod rynnę. Wyjaśnijmy sobie więc najpierw skrót SGML. SGML – Standard Generalized Markup Language, to przemysłowy standard obróbki dokumentów elektronicznych zgodny z ISO 8879. Jest on nadrzędny zarówno w stosunku do XML, jak i języka HTML. Oba te języki mają swoje miejsce pod ogromnym parasolem SGML – XML jest aplikacją SGML, a XHTML aplikacją XML. Za pomocą znaczników oraz deklaracji typu dokumentów aplikacje te ustalają jednolite formatowanie danych uwzględniające ich specyfikę i zakres zastosowań. Sam SGML nie zawiera natomiast żadnych instrukcji formatujących.

Dokumenty SGML przechowywane są w plikach tekstowych.

Więcej informacji na temat języka XHTML znajdziesz w rozdziale 9.

Definicja typu dokumentu, DTD

Co to jest DTD?

Definicja typu dokumentu, DTD, to specyfikacja towarzysząca każdemu dokumentowi stworzonemu w języku SGML (Standard Generalized Markup Language), której zadaniem jest dostarczenie przeglądarce istotnych informacji o składni dokumentu. Innymi słowy definicje DTD umożliwiają określenie zasad stosowania znaczników.

DTD zawiera przeważnie definicje nazw i zawartości wszystkich rodzajów elementów, jakie mogą wystąpić w danym dokumencie. Może także określać wymaganą kolejność elementów oraz zasady ich zagnieżdżenia, dopuszczalne atrybuty wraz z ich domyślnymi wartościami, warunki rezygnacji ze znaczników końcowych czy nazwy predefiniowanych symboli, do których można się odwoływać w dokumencie.

Szczególnym przypadkiem DTD jest język HTML. Odbiornikiem definicji jest tutaj przeglądarka internetowa, którą zaprojektowano pod kątem pobierania i analizy dokumentów tekstowych sformatowanych przy pomocy znaczników i zasad określonych w specyfikacji HTML.

Definicja typu dokumentu może znajdować się w pliku dokumentu lub w pliku zewnętrznym, wskazanym przez adres URL. Takie zewnętrzne pliki DTD mogą być używane przez różne dokumenty i różne strony WWW.

Oto przykład odwołania do definicji typu dokumentu umieszczonego w obrębie dokumentu HTML:

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

<html>
<head>
...
</head>
<body>
...
</body>
</html>
```

Zapisywanie dokumentów HTML w edytorze i otwieranie w przeglądarce

Jakie edytory?

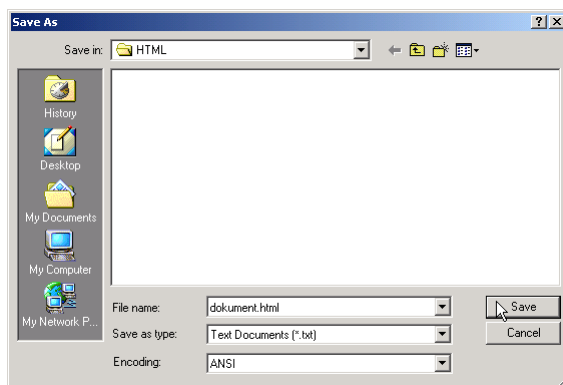
Zanim przystąpimy do tworzenia kodu HTML na serio, kilka uwag na temat korzystania z edytorów, w których kod HTML będzie wprowadzany. **Pamiętaj, że dokument HTML jest dokumentem tekstowym.** Zapisując go, musisz wziąć to pod uwagę.

Nazwa dokumentu składa się z dwóch członów – nazwa.rozszerzenie. Zdefiniuj więc dowolną nazwę i dodaj do niej rozszerzenie. W przypadku dokumentów HTML rozszerzenie ma postać .htm lub .html (bardziej pewne jest to właśnie).

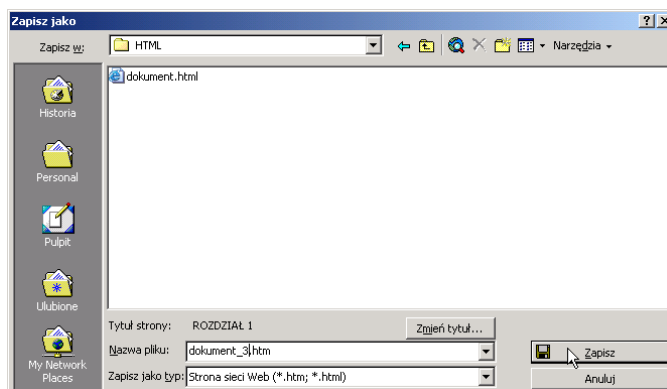
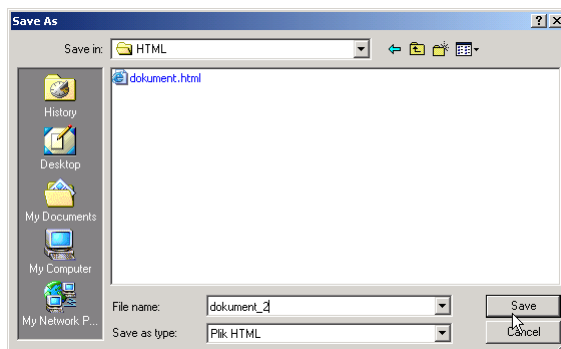
Zapisując dokument w edytorach takich jak NotePad czy WordPad, w których format tekstowy jest jedynym formatem, musisz rozszerzenie wpisać ręcznie. W edytorze Webber format HTML jest formatem domyślnym, nie musisz więc pamiętać o rozszerzeniu.

Jeśli chcesz korzystać z edytora Word, możesz, lecz musisz pamiętać, by w oknie dialogowym Zapisz jako (Save As) wybierać z listy Zapisz jako Typ (Save As Type) typ (*.html) (patrz rysunek 1.1). Jeśli pozostawisz format domyślny (*.doc), Word zapisze znacznie więcej, niż sam tekst na stronie. Zapisze też ustawienia marginesów, tabulatorów, czcionki i całą masę innych ustawień formatowania potrzebnych do poprawnego wyświetlenia strony, lecz absolutnie niepotrzebnych w kodzie HTML. Na wszelki wypadek nie przyzwyczajaj się zbytnio do edytora Word jako narzędzia tworzenia stron WWW – każda rzecz wymaga odpowiednich narzędzi.

Rys. 1.1.
Zapisywanie dokumentu HTML w Notatniku – tutaj musisz ręcznie wpisać rozszerzenie nazwy .html, Webberze – domyślnie edytor przyjmuje rozszerzenie .html, i w edytorze Word – w tym przypadku



format pliku
trzeba
wybrać z
listy
rozwijanej
Zapisz jako
typ

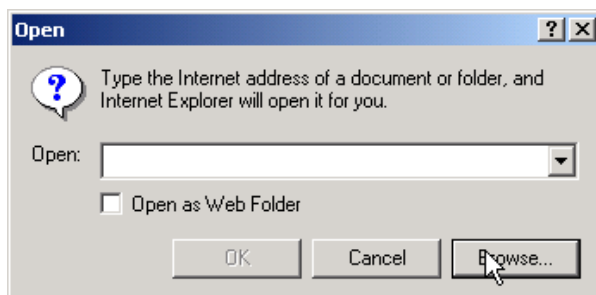


Jak wyświetlić w przeglądarce plik html zapisany na dysku twardym?

Dokument HTML zapisany na dysku twardym lub na dyskietce można otworzyć w przeglądarce. Operacja jest bardzo prosta. Wybierz w menu **File** pozycję **Open**, a następnie kliknij przycisk **Browse** lub **Open File** i odszukaj żądany dokument (patrz rysunek 1.2), a następnie naciśnij **OK**, by pojawił się w oknie przeglądarki.

Rys. 1.2.

Otwieranie dokumentu HTML w przeglądarce – tutaj jest to przeglądarka Internet Explorer



Trochę treningu

Żeby oswoić się z językiem HTML, przyglądnij się postaciom źródłowym stron WWW dostępnych w Internecie.

Jak wyświetlić postać źródłową strony WWW?

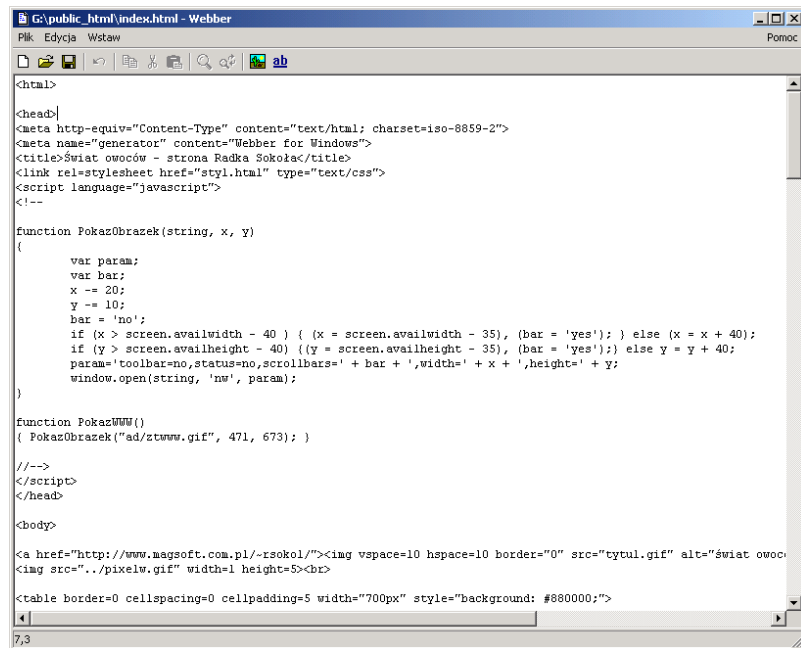
Jak już wiesz, w swojej postaci źródłowej, strona WWW to po prostu plik tekstowy, do którego nazwy dołączone jest przedłużenie .html. Jeśli wyświetlisz taki plik w jakimś edytorze tekstu, zobaczysz tekst wymieszany ze specyficznymi znakami – są to **znaczniki HTML**. Być może będziesz zaskoczony, że piękne strony WWW, z obrazkami i animacjami, wyglądają tak nieciekawie. Aby pojawiły się w całej krasie, ich kod źródłowy musi zostać „odczytany” przez przeglądarkę – ot i cała tajemnica.

Wyświetl kolejno kilka stron WWW w oknie przeglądarki, a najlepiej wyświetl strony w osobnych oknach. Jeśli korzystasz z przeglądarki Netscape Navigator, wybierz w menu View pozycję Page source (użytkownicy programu Internet Explorer powinni wybrać pozycję Source w menu View) (patrz rysunek 1.3).

Rys. 1.3.

Strona WWW i jej postać źródłowa





```
G:\public_html\index.html - Webber
Plik Edycja Wstaw
<html>
<head>
<meta http-equiv="Content-Type" content="text/html; charset=iso-8859-2">
<meta name="generator" content="Webber for Windows">
<title>Świat owoców - strona Radka Sokoła</title>
<link rel="stylesheet" href="styl.html" type="text/css">
<script language="javascript">
<!--
function PokazObrazek(string, x, y)
{
    var param;
    var bar;
    x -= 20;
    y -= 10;
    bar = 'no';
    if (x > screen.availwidth - 40) { (x = screen.availwidth - 35), (bar = 'yes'); } else (x = x + 40);
    if (y > screen.availheight - 40) {(y = screen.availheight - 35), (bar = 'yes');} else y = y + 40;
    param='toolbar=no,status=no,scrollbars=' + bar + ',width=' + x + ',height=' + y;
    window.open(string, 'nw', param);
}

function PokazWWW()
{ PokazObrazek("ad/ztwww.gif", 471, 673); }

//-->
</script>
</head>
<body>
<a href="http://www.magsoft.com.pl/~rsokol/"><br>
<table border=0 cellspacing=0 cellpadding=5 width="700px" style="background: #880000;">
7,3
```

Przyglądaj się postaciom źródłowym stron zwracając uwagę na powtarzające się elementy. Podglądanie gotowych dokumentów HTML to doskonała metoda nauki języka.

Najważniejsze informacje

- ❑ HTML to język hipertekstowego znakowania (Hypertext Markup Language), który stanowi ustanowioną przez konsorcjum World Wide Web Consortium (W3C) specyfikację nakreślającą postać dokumentów prezentowanych w Internecie. Twórcy przeglądarek sieciowych sprawdzają za pośrednictwem języka HTML jak ich produkty będą wyświetlać strony WWW na ekranie Twojego komputera. Twórcą języka HTML jest Tim Berners-Lee.
- ❑ HTML pozwala publikować w Internecie dokumenty zawierające nagłówki, tekst, tabele, listy, zdjęcia, formularze, itd., pobierać za pośrednictwem „połączeń hipertekstowych” informacje z Internetu, projektować formularze oraz umieszczać w dokumentach arkusze kalkulacyjne, klipy wideo oraz inne aplikacje.
- ❑ Dokument HTML jest dokumentem tekstowym. Jego nazwa składa się z dwóch członów – nazwa.rozszerzenie. W przypadku dokumentów HTML rozszerzenie ma postać .htm lub .html (bardziej pewne jest to właśnie).
- ❑ Dokumenty html można tworzyć przy użyciu najprostszych edytorów tekstowych, takich jak Notatnik czy Webber.