

Rozdział 2.

Internet, strona WWW i znaczniki HTML

World Wide Web i strona WWW

Co to jest World Wide Web?

World Wide Web (w skrócie WWW) to sieć komputerów nazywana popularnie Internetem. Tworzą ją komputery z całego świata. Komputery te porozumiewają się ze sobą za pośrednictwem protokołu HTTP.

Jak to działa?

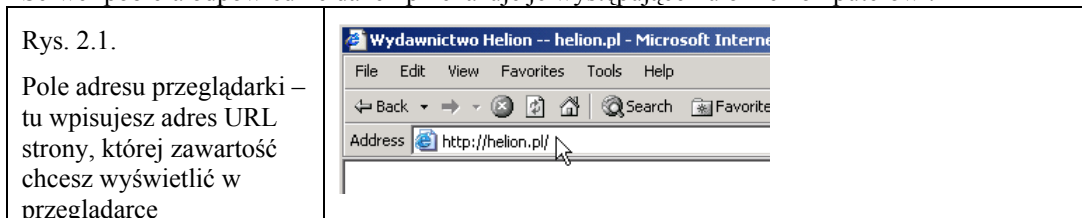
Wszelkie informacje umieszczone w sieci zapisane są w dokumentach zwanymi stronami WWW. Strona WWW to dokument utworzony za pomocą języka HTML. Gotowe strony WWW zapisywane są na specyficznym komputerze, który nazywamy serwerem. Serwer to właściwie program, ale przyjęło się nazywać tym mianem także komputer, na którym go zainstalowano. Serwer jest odpowiedzialny za realizację żądania przesłania dokumentów do tego komputera, który wystąpił z takim żądaniem. Jest także konieczny do opublikowania dokumentów w Internecie. Internet z kolei stanowi sieć komputerów połączonych ze sobą. Oczywiście nie fizycznie. Komputery mogą jednak komunikować się ze sobą i przysyłać sobie dane, czyli informacje. Strony WWW stanowią jedno ze źródeł informacji.

Co musisz zrobić, żeby przeglądać strony WWW?

Aby przeglądać strony WWW, konieczne jest nawiązanie połączenia z Internetem – za pośrednictwem modemu lub łącza stałego – i uruchomienie specjalnego programu, który nazywa się **przeglądarką** – najpopularniejsze przeglądarki to Netscape Navigator i Internet Explorer. Przeglądarkę należy poinformować, gdzie ma szukać danych. Informacja ta zawarta jest w adresie **URL**.

Co to jest adres URL?

URL – to skrót od **U**niform **R**esource **L**ocator, co można przetłumaczyć jako „uniwersalny identyfikator zasobów”. Jego postać jest mniej więcej taka: `http://www.dobrastrona.com/`. Adres ten wpisujesz najczęściej w polu adresu przeglądarki (patrz rysunek 2.1). Przeglądarka wysyła wiadomość do komputera, którego adres podałeś (serwera), zawierającą żądanie strony. Serwer pobiera odpowiednie dane i przekazuje je występującemu o nie komputerowi.



Co przeglądarka robi z danymi, które otrzyma od serwera?

Dane, które trafiły do Twojej przeglądarki, są zakodowane w specjalny sposób – są zapisane w języku HTML. Przeglądarce to nie przeszkadza – potrafi ona interpretować ten kod, więc wyświetla pobraną stronę na Twoim ekranie. Język HTML informuje przeglądarkę jak ma obsługiwać pobrane dane.

Tak więc sam widzisz, jak ważny jest język HTML. Dotrzemy teraz do jego istoty – zajmiemy się znacznikami HTML.

Znaczniki

Jeśli wyświetlisz kod źródłowy strony WWW, zobaczysz, że zawiera on szereg poleceń języka HTML umieszczonych w nawiasach kątowych. Twory te noszą nazwę **znaczników**.

Jak wyglądają znaczniki?

Znaczniki stosowane są zazwyczaj w parach (zwróć uwagę na ukośnik, /, który odróżnia znacznik zamykający od otwierającego):

`<znacznik otwierający>`, `</znacznik zamykający>`, np. `<html>`, `</html>`, ale są wyjątki. Oto niektóre z nich:

- `<p>` – znacznik akapitu (w zasadzie istnieje znacznik zamykający, `</p>`, ale pominięcie go nie spowoduje błędu w działaniu strony; stosuj ten znacznik, aby rozpocząć nowy akapit),
- `
` – znacznik nowego wiersza
- `<hr>` – znacznik umożliwiający dodanie na stronie linii poziomej.

Parę znaczników (lub znacznik pojedynczy, jeśli nie ma on znacznika zamykającego) nazywa się także elementem. Element pozbawiony znacznika zamykającego nazywa się pustym. Może on zawierać jedynie atrybuty.

Jaka jest funkcja znaczników?

Znaczniki informują przeglądarkę, jak wyświetlić stronę na ekranie.

Wielkie czy małe litery w znacznikach?

Znaczniki nie są wrażliwe na wielkość liter. Oznacza to, że <body> to to samo co <BODY>. Warto jednak zacząć przyzwyczajając się do małych liter. Są one zalecane w rekomendacji W3C specyfikacji HTML 4, a w specyfikacji XHTML są wymagane.

Co to są atrybuty znaczników i jaka jest ich funkcja?

Atrybuty znaczników dostarczają dodatkowych informacji o elementach strony definiowanych poprzez znaczniki. Jeśli na przykład chciałbyś, aby Twoja strona miała tło w kolorze zielonym, wystarczy, że poinformujesz o tym przeglądarkę dodając do znacznika <body> odpowiedni atrybut: <body bgcolor="green">. Atrybut definiowany jest za pomocą pary: atrybut="wartość" i umieszczany zawsze w znaczniku otwierającym danego elementu HTML. Wartości należy podawać w cudzysłowie – pojedynczym, ' ', lub podwójnym, " ".

Elementy dokumentu html

Pewne znaczniki pełnią szczególne funkcje. Tworzą bowiem podstawowy szkielet dokumentu html.

Jaka jest ta podstawowa postać kodu HTML dokumentu html?

Otwórz Notatnik lub edytor Webber i wpisz w nim następujący tekst (Wydruk 2.1) – jest to podstawowy, najprostszy szkielet dokumentu html:

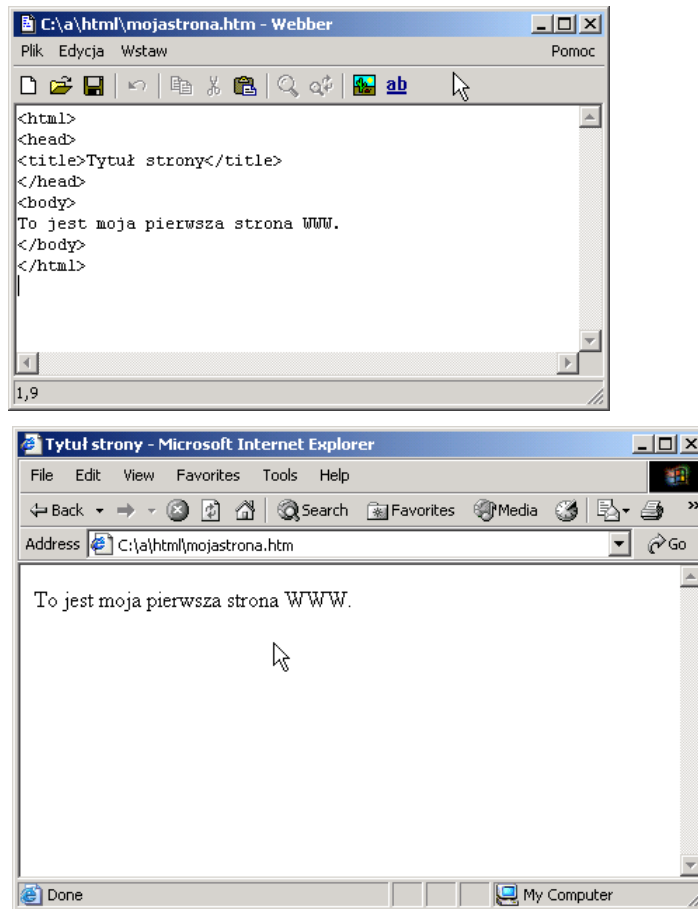
Wydruk 2.1. Podstawowy szkielet dokumentu html

```
<html>
<head>
<title>Tytuł strony</title>
</head>
<body>
To jest moja pierwsza strona WWW.
</body>
</html>
```

Zapisz plik pod nazwą `mojastrona.html`, uruchom przeglądarkę i otwórz w niej swój plik (wybierz w menu File lub Plik – w zależności od wersji językowej – pozycję Open, Open File lub Otwórz). W przeglądarce wyświetlone zostanie **tylko** zdanie *To jest moja pierwsza strona WWW*. Reszta tekstu nie pojawi się – to kod dla przeglądarki informujący ją, co ma z dokumentem zrobić (patrz rysunek 2.2).

Rys. 2.2.

Zapisz dokument HTML i otwórz go w przeglądarce – oto Twoja pierwsza strona WWW



Jakie znaczenie mają poszczególne wyrażenia kodu z wydruku 2.1?

Pierwszym znacznikiem w Twoim dokumencie HTML jest znacznik `<html>`. Informuje on przeglądarkę, że jest to **początek dokumentu HTML**. Ostatnim znacznikiem kodu jest znacznik `</html>` – jak już wiesz, jest to znacznik zamykający i informuje on przeglądarkę, że na nim **kończy się dokument HTML**.

Znaczniki `<html>` i `</html>` informują, że wszystko, co znajduje się między nimi stanowi dokument HTML. Nową stronę będziesz zawsze rozpoczynał od znacznika `<html>` i kończył ją znacznikiem `</html>`.

Znaczniki `<head>` i `</head>` wyznaczają element strony, który nosi nazwę **nagłówek**. Zawartość nagłówka nie jest wyświetlana na stronie WWW. Zwróć jednak uwagę na pasek tytułu okna przeglądarki, w którym wyświetliłeś swoją pierwszą stronę WWW. Tekst, który wpisałeś między znacznikami `<title></title>`, pojawia się na pasku tytułu tego okna. Strona może mieć tylko jeden tytuł, a ten z kolei może zawierać tylko tekst – żadne inne znaczniki nie są tu dozwolone. Tytuł nie może być zbyt długi, bo nie zmieści się na pasku tytułu, ale nie przesadz też

w odwrotną stronę, aby dać szansę przypadkowym czytelnikom zorientować się jaka jest tematyka strony.

Pierwszym elementem strony jest **nagłówek**. Jego ramy wyznacza para `<head>`, `</head>`. Między tymi znacznikami umieszczany jest kod definiujący tytuł strony. W żadnym wypadku nie należy tu umieszczać treści strony!

Strona WWW ma oczywiście swoją zawartość. W przykładzie umieściliśmy ją między znacznikami `<body>`, `</body>`. One właśnie wyznaczają drugi element strony WWW – obszar treści.

Para `<body>`, `</body>` wyznacza podstawową część strony – tu twórca strony WWW umieszcza wszystko: treść, obrazy, połączenia, itd.

Elementy strony WWW często nazywa się od angielskich nazw znaczników sekcją *head* i sekcją *body*. Potrafisz już je zdefiniować, potrafisz wyświetlić na ekranie przeglądarki tekst. Pora, abyś poznał nowe znaczniki i nauczył się kontrolować wygląd wyświetlanego na ekranie tekstu oraz nadawać mu określoną strukturę.

Co to jest element meta?

Element `head` zawiera ogólne informacje na temat dokumentu. Umieszczane są one między innymi w znaczniku `<meta>`. Najczęściej są to informacje przydatne przeglądarkom lub słowa kluczowe wykorzystywane przez serwisy wyszukiwujące (coraz rzadziej jednak). W znaczniku `<meta>` można na przykład umieścić adres `url`, pod który przeglądarka ma się skierować, aby wyświetlić nową wersję strony WWW, lub opis dokumentu.

Do czego służą atrybuty znacznika `<meta>`?

Niektóre serwisy wyszukiwujące korzystają z atrybutów znacznika `<meta>` przy indeksowaniu dokumentów.

Atrybuty opisują dokument i dostarczają słów kluczowych. I tak atrybut `name` określa sposób opisu dokumentu dostarczony w atrybucie `content`.

Oto przykład elementu `meta`, który definiuje opis strony (`name="description"`):

```
<meta name="description" content="Strona poświęcona kotom i innym futrzakom">
```

Oto przykład elementu `meta`, który zawiera słowa kluczowe, charakteryzujące zawartość strony WWW (`name="keywords"`):

```
<meta name="keywords" content="HTML, DHTML, CSS, XML, XHTML, JavaScript, VBScript">
```

Oto przykład dokumentu `html`, w którym za pomocą elementu `meta` umieszczono informacje o autorze strony, dacie ostatniej modyfikacji dokumentu oraz o stosowanym oprogramowaniu (patrz wydruk 2.2).

Wydruk 2.2. Atrybuty znacznika `<meta>`

```
<html>
```

```
<head>
```

```
<meta name="author"
content="Jan Kowalski">

<meta name="revised"
content="Jan Kowalski,6/10/02">

<meta name="generator"
content="Microsoft FrontPage 4.0">
</head>
<body>
<p>
Atrybuty meta podają w tym dokumencie dane autora, datę wprowadzenia
poprawek oraz typ stosowanego oprogramowania.
</p>
</body>
</html>
```

Jak wykorzystać element meta, aby skierować użytkownika pod inny adres url?

Element `meta` będzie bardzo przydatny, jeśli zmienił się adres `url`, pod którym można było oglądać Twoją stronę WWW. Na wydruku 2.3 znajdziesz kod `html`, który pozwoli poinformować użytkownika, gdzie powinien Cię teraz szukać, a nawet sam skieruje go we właściwe miejsce.

Wydruk 2.3. Kierowanie użytkownika pod nowy adres url

```
<html>
<head>
<meta http-equiv="Refresh"
content="5;url=http://www.adres.com">
</head>
<body>
<p>
Przepraszam! Już mnie tu nie ma! Nowy URL jest: <a
href="http://www.adres.com">http://www.adres.com</a>
</p>
<p>
Za kilka sekund zostaniesz skierowany pod nowy adres.
</p>
<p>
```

Jeśli komunikat wyświetlany jest dłużej, niż 5 sekund, kliknij to połączenie!

</p>

</body>

</html>

Kodowanie polskich znaków

Jak wykorzystać znacznik <meta>, aby zdefiniować stronę kodową?

Poprawne zdefiniowanie kodowania znaków jest w dokumencie HTML bardzo pożądane. Zaleca się aby polskie strony były kodowane w ISO-8859-2, bo tylko wtedy będą poprawnie wyświetlane. Korzystając z edytora Webber nie musisz się martwić o poprawność kodowania samych liter – to jego zmartwienie - musisz jednak pamiętać o wstawieniu odpowiedniego nagłówka <meta>.

Aby zapewnić czytelność dokumentu HTML, możesz:

- ❑ Zrezygnować ze stosowania polskich znaków diakrytycznych – takie rozwiązanie posiada wiele zalet: niezależnie od systemu zawsze widać znaki – może słowa nie są poprawnie ortograficznie, ale względnie czytelne. Wyrazy nie są podzielone dziwnymi znakami. Zdarza się że potencjalni czytelnicy nie posiadają przeglądarki czytającej po polsku. Większość przeglądarek ma polską stronę kodową – większość w Polsce, na pewno tak, ale niekoniecznie tak będzie gdzieś w świecie. Dlatego też kodowanie bez polskich znaków jest pewnym rozwiązaniem i nie wymaga żadnych specjalnych zabiegów.
- ❑ Korzystać z kodowania Windows-1250 – kodowanie Windows zapewnia poprawne oglądanie strony użytkownikom systemu Windows, ale są przecież inne systemy i inni użytkownicy. Niewątpliwie zaletą tego sposobu jest to, że użytkownicy polskiej wersji systemu Windows mogą bez żadnych problemów pisać kod swoich stron korzystając z Notatnika. Gdy z jakich powodów zechcesz zdefiniować w dokumencie HTML stronę kodową Windows-1250, wpisz odpowiednią informację meta: <meta http-equiv="content-type" content="text/html; charset=windows-1250">. Takie rozwiązanie należy jednak odradzać.
- ❑ **Zdefiniować stronę kodową ISO-8859-2** – aby zdefiniować w dokumencie HTML tę stronę kodową, wpisz w elemencie meta:

```
<meta http-equiv="content-type" content="text/html; charset=ISO-8859-2">.
```

Wymaga to korzystania z edytora, który potrafi prawidłowo kodować polskie litery – takim edytorem jest Webber (<http://www.magsoft.com.pl/~rsokol/>). To rozwiązanie gorąco polecam.

Najważniejsze informacje

- ❑ World Wide Web (w skrócie WWW) to sieć komputerów nazywana popularnie Internetem.
- ❑ Informacje dostępne są w Internecie w postaci dokumentów HTML zwanych stronami WWW.
- ❑ Strony WWW umieszczane są na specyficznych komputerach zwanych serwerami.
- ❑ Aby przeglądać strony WWW, konieczne jest nawiązanie połączenia z Internetem – za pośrednictwem modemu lub łącza stałego – i uruchomienie specjalnego programu, który nazywa się **przeglądarką**.
- ❑ URL – to skrót od **Uniform Resource Locator**, co można przetłumaczyć jako „uniwersalny identyfikator zasobów”. Jego postać jest mniej więcej taka:
`http://www.dobrastrona.com/`. Adres ten określa ulokowanie strony WWW i jest podawany w polu adresu przeglądarki.
- ❑ Dane, które trafiły do Twojej przeglądarki, są zakodowane w specjalny sposób – są zapisane w języku HTML. Język HTML informuje przeglądarkę jak ma obsłużyć pobrane dane – informacje te są zawarte w znacznikach .
- ❑ Znaczniki to polecenia języka HTML umieszczone w nawiasach kątowych. Ich postać jest następująca: <znacznik otwierający>, </znacznik zamykający>.
- ❑ Znaczniki nie są wrażliwe na wielkość liter.
- ❑ Dokument HTML otwiera znacznik <html>, a zamyka znacznik </html>. Elementy strony umieszczane są między tymi znacznikami.
- ❑ Podstawowe elementy strony WWW to obszar nagłówka i obszar treści (sekcja *head* i sekcją *body*).
- ❑ Obszar nagłówka wyznaczają znaczniki <head> i </head> .
- ❑ Zawartość strony WWW – tekst, obrazy, połączenia, itp. – jest umieszczana między znacznikami <body>, </body>. One właśnie wyznaczają drugi element strony WWW – obszar treści.
- ❑ Element *head* zawiera ogólne informacje na temat dokumentu. Umieszczane są one między innymi w znaczniku <meta>. Najczęściej są to informacje przydatne przeglądarkom lub słowa kluczowe wykorzystywane przez serwisy wyszukiwujące. W znaczniku <meta> można na przykład umieścić adres `url`, pod który przeglądarka ma się skierować, aby wyświetlić nową wersję strony WWW, lub opis dokumentu. Można także umieścić nowy adres `url`, pod który użytkownik zostanie automatycznie skierowany po określonym czasie.
- ❑ Stronę kodową ISO-8859-2 (polskie litery!) definiuje się w elemencie `meta`:

```
<meta http-equiv="content-type" content="text/html; charset=ISO-8859-2">
```