

Sprawozdanie nr 4

Sieci komputerowe i bazy danych

Badanie protokołów WWW

Arkadiusz Marta 27.03.2019 r.

1. Wyszukanie zasobów RSC 1945 oraz RSC 2616

RSC 1945 i RSC 2616 to protokoły przesyłania danych hipertekstowych – HTTP . Za pomocą protokołu HTTP przesyła się żądania udostępnienia dokumentów WWW i informacje o kliknięciu odnośnika oraz informacje z formularzy. Obecnie używany jest protokół RSC 2616.

Network Working Group
Request for Comments: 1945
Return Informational
Regions: Informational
Residence
UC Irvine
H. Frystyk
MIT/LCS
May 1996

Network Working Group
Request for Comments: 1945
Category: Informational

MIT/LCS
UC Irvine
H. Frystyk
MIT/LCS
MAY 1996

Hypertext Transfer Protocol -- HTTP/1.0

Status of This Memo

This memo provides information for the Internet community. This memo does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

IESG Note:

The IESG has concerns about this protocol, and expects this document to be replaced relatively soon by a standards track document.

Abstract

The Hypertext Transfer Protocol (HTTP) is an application-level protocol with the lightness and speed necessary for distributed, collaborative, hypermedia information systems. It is a generic, stateless, object-oriented protocol which can be used for many tasks, such as name servers and distributed object management systems, through extension of its request methods (commands). A feature of HTTP is the typing of data representation, allowing systems to be built independently of the data being transferred.

HTTP has been in use by the World-Wide Web global information initiative since 1990. This specification reflects common usage of the protocol referred to as "HTTP/1.0".

Hypertext Transfer Protocol -- HTTP/1.0

Status of This Memo

This memo provides information for the Internet community. This memo does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

IESG Note:

The IESG has concerns about this protocol, and expects this document to be replaced relatively soon by a standards track document.

Abstract

The Hypertext Transfer Protocol (HTTP) is an application-level protocol with the lightness and speed necessary for distributed, collaborative, hypermedia information systems. It is a generic, stateless, object-oriented protocol which can be used for many tasks, such as name servers and distributed object management systems, through extension of its request methods (commands). A feature of HTTP is the typing of data representation, allowing systems to be built independently of the data being transferred.

HTTP has been in use by the World-Wide Web global information initiative since 1990. This specification reflects common usage of the protocol referred to as "HTTP/1.0".

2. Badanie działania protokołu http

Zgodnie z poleceniem uruchomiono program **putty,** gdzie ustawiono rodzaj połączenia (*connection type*) na **raw**. Jako nazwę hosta (*host name*) podano następujący adres: **mts.wibor.agh.edu.pl**. Wybrano port **80** oraz zmieniono opcję *close window on exit* na **never**.

Category:						
Session Logging Terminal Keyboard Bell Features Window Appearance Behaviour Translation Colours Connection Data Proxy Telnet Rlogin SSH Serial		Basic options for your PuTTY session				
		pecify the destination you want to connect to Host Name (or IP address) Port mts.wibro.agh.edu.pl Connection type: Raw Telnet Rlogin SSH Serial oad, save or delete a stored session Saved Sessions				
		Default Settings Load Save Delete				
		Close window on exit Always Never Only on clean exit				
About	Help	Open Cancel				

Następnie przy użyciu polecenia **GET** połączono się z serwerem **mts.wibro.agh.edu.pl**.

```
GET /~s292528/ HTTP/1.1
Host: mts.wibro.agh.edu.pl
```

Otrzymano nastepującą odpowiedź serwera.

```
Date: Wed, 27 Mar 2019 09:02:45 GMT
Server: Apache/2.4.18 (Ubuntu)
Last-Modified: Wed, 20 Mar 2019 11:11:10 GMT
ETag: "20e-58484adcd0f80"
Accept-Ranges: bytes
Content-Length: 526
Content-Type: text/html
     <head>
            <title> Strona próbna!! </title>
                                                  <body>Witaj świecie!!! <BR>
img src="https://www.catster.com/wp-content/uploads/2018/07/Savannah-cat-long-bo
dy-shot.jpg" border="2" height=150>
                                       <MARQUEE>Strona testowa</MARQUEE>
                                                                        </div
Arkadiusz Marta
               </div>
                         Polskie znaki : Ćma śmieje się z chrząszcza
                                                                    </div>
                      </div>
                                        cokolwiek</t
d> mozna
                    tu napisac
                                                        </body>
```

W odpowiedzi otrzymaliśmy zawartość pliku **index.html** utworzonego na pierwszych zajęciach laboratoryjnych.

Następnie sprawdzono odpowiedź serwera na błędnie podany nagłówek zamiast polecenia **GET** podając hasło **GT**.

```
GT /~s292528/ HTTP/1.1
Host:mts.wibro.agh.edu.pl
```

Otrzymano następująca odpowiedź serwera.

```
HTTP/1.1 403 Forbidden
Date: Wed, 27 Mar 2019 09:13:16 GMT
Server: Apache/2.4.18 (Ubuntu)
Content-Length: 304
Content-Type: text/html; charset=iso-8859-1
<!DOCTYPE HTML PUBLIC "-//IETF//DTD HTML 2.0//EN">
                                                  <html><head>
                                                              <title>403 Forbidd
en</title>
          </head><body>
                       <h1>Forbidden</h1>
                                         You don't have permission to access
~s292528/
         on this server.<br />
                               <hr>
                                       <address>Apache/2.4.18 (Ubuntu) Server at
mts.wibro.agh.edu.pl Port 80</address>
                                      </body></html>
```

W odpowiedzi serwer zwraca **błąd 403**, który oznacza, że serwer zabrania dostępu do danego zasobu. Otrzymujemy hasło zwrotne: You don't have permission to access on this server.

Następnie sprawdzono odpowiedź serwera na niepoprawny adres strony podając zamiast **mts.wibro.agh.edu.pl** hasło **mts.wibro.agh.edu.us**.

```
GET /~s292528/ HTTP/1.1
Host: mts.wibro.agh.edu.us
```

Otrzymano następująca odpowiedź serwera.

W odpowiedzi serwer zwraca zawartość pliku **index.html** . Dzieje się tak dlatego, iż użytkownik jest już zalogowany na tym serwerze, zatem zwraca on poprawną zawartość

Następnie sprawdzono odpowiedź serwera na niepoprawny adres hosta, podając zamiast numeru albumu **s292528** przypadkowy ciąg znaków **s5o116c**.

```
GET /~s5o116c/ HTTP/1.1
Host: mts.wibro.agh.edu.pl
```

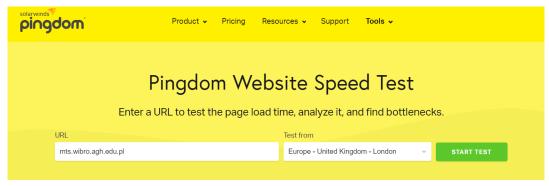
Otrzymano następująca odpowiedź serwera.

```
HTTP/1.1 404 Not Found
Date: Wed, 27 Mar 2019 09:27:36 GMT
Server: Apache/2.4.18 (Ubuntu)
Content-Length: 293
Content-Type: text/html; charset=iso-8859-1
!DOCTYPE HTML PUBLIC "-//IETF//DTD HTML 2.0//EN">
                                                  <html><head>
                                                              <title>404 Not Fou
nd</title>
         </head><body>
                      <h1>Not Found</h1>
                                         The requested URL /~s5o116c/ was not
found on this server.
                          <hr>
                              <address>Apache/2.4.18 (Ubuntu) Server at mts.wibr
 .agh.edu.pl Port 80</address>
                              </body></html>
```

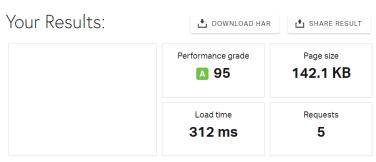
Serwer w odpowiedzi zwraca **błąd 404** – Not Found – który oznacza że nastąpiło połączenie z serwerem, jednak nie został odnaleziony podany zasób.

3. Testowanie szybkości pobierania stron

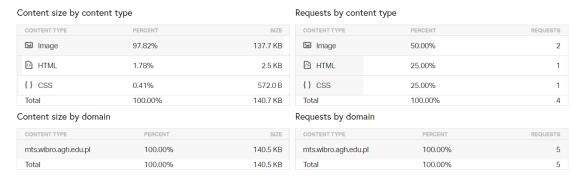
Korzystając ze strony **http://tools.pingdom.com/** sprawdzono szybkość pobierania stron www.



a) Przetestowano szybkość pobierania strony mts.wibroagh.edu.pl.

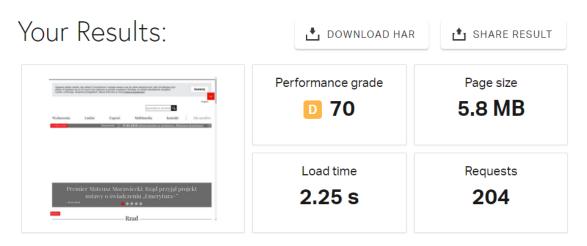


W rezultacie jako informację zwrotną otrzymano rozmiar strony – *page size* – oraz czas pobierania strony – *load time* – który w tym wypadku wyniósł **312 ms**.

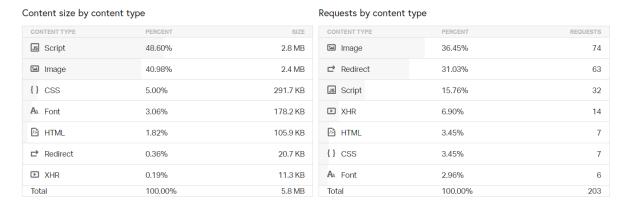


Otrzymano również informację o ilości żądań ze względu na rodzaj zawartości (*obrazy, pliki html i css*) oraz ich udział w ogólnym rozmiarze strony.

b) Przetestowano szybkość pobierania strony https://www.premier.gov.pl/.

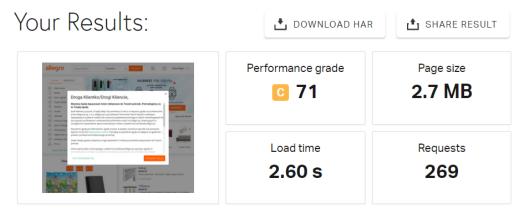


W rezultacie jako informację zwrotną otrzymano rozmiar strony – *page size* – oraz czas pobierania strony – *load time* – który w tym wypadku wyniósł **2.25 s**.

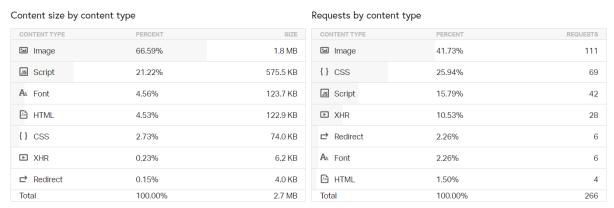


Otrzymano również informację o ilości żądań ze względu na rodzaj zawartości (*obrazy, pliki html i css, itd.*) oraz ich udział w ogólnym rozmiarze strony.

c) Przetestowano szybkość pobierania strony https://allegro.pl/.

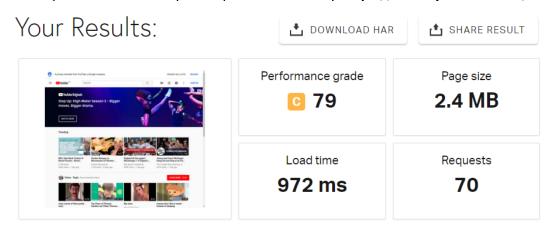


W rezultacie jako informację zwrotną otrzymano rozmiar strony – *page size* – oraz czas pobierania strony – *load time* – który w tym wypadku wyniósł **2.6 s**.



Otrzymano również informację o ilości żądań ze względu na rodzaj zawartości (*obrazy, pliki html i css, itd.*) oraz ich udział w ogólnym rozmiarze strony.

d) Przetestowano szybkość pobierania strony https://www.youtube.com/.



W rezultacie jako informację zwrotną otrzymano rozmiar strony – *page size* – oraz czas pobierania strony – *load time* – który w tym wypadku wyniósł **972 ms**.

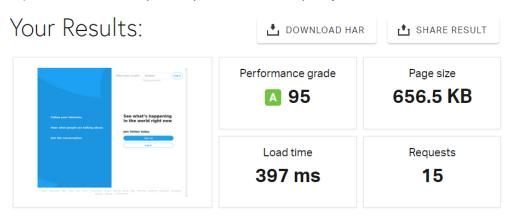
Content size by content type

Requests by content type

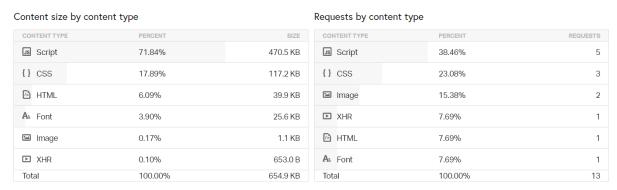
CONTENT TYPE	PERCENT	SIZE	CONTENT TYPE	PERCENT	REQUESTS
Js Script	46.79%	1.1 MB	■ Image	56.92%	37
Image	26.83%	641.8 KB	Js Script	13.85%	9
	21.06%	503.9 KB	{} css	9.23%	6
A Font	2.78%	66.5 KB	■ XHR	7.69%	5
{} css	2.34%	56.0 KB	A _A Font	6.15%	4
▶ XHR	0.16%	3.8 KB		4.62%	3
□ Redirect	0.04%	901.0 B	→ Redirect	1.54%	1
Total	100.00%	2.4 MB	Total	100.00%	65

Otrzymano również informację o ilości żądań ze względu na rodzaj zawartości (*obrazy, pliki html i css, itd.*) oraz ich udział w ogólnym rozmiarze strony.

e) Przetestowano szybkość pobierania strony https://twitter.com/.



W rezultacie jako informację zwrotną otrzymano rozmiar strony – *page size* – oraz czas pobierania strony – *load time* – który w tym wypadku wyniósł **397 ms**.



Otrzymano również informację o ilości żądań ze względu na rodzaj zawartości (*obrazy, pliki html i css, itd.*) oraz ich udział w ogólnym rozmiarze strony.

Wnioski

Na podstawie powyższych testów można wyciągnąć wniosek iż strony, które zawierają więcej zdjęć – *posiadają one największy rozmiar* – są pobierane najwolniej.

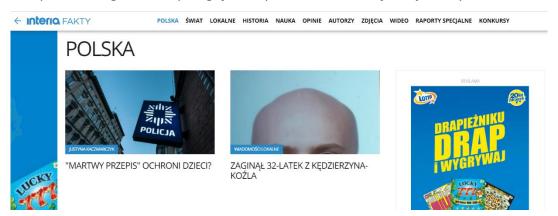
4. Definicja i zastosowanie URL

URL (adres URL) – ujednolicony format adresowania i identyfikowania zasobów internetu, w tym najczęściej stron WWW. Adres URL (ang. Uniform Resource Locator) składa się z dwóch elementów: rodzaju zasobu oraz części zależnej. Rodzaj zasobu to w przypadku strony WWW protokoły **HTTP** lub **HTTPS**, czyli szyfrowana wersja protokołu HTTP. Natomiast część zależna to nazwa domeny, rozszerzenie, ścieżka dostępu.

a) Odniesienie do grafiki umieszczonej na stronie Odniesienie następuje poprzez podanie ścieżki do danej grafiki. Poniżej podano ścieżkę do logo interii: https://fakty.interia.pl/i/interia-logo-2-0.svg. Po podaniu tego adresu przeglądarka zwraca żądany obraz.



b) Odniesienie do elementu (zakładki) na stronie
Odniesienie do zakładki następuje poprzez podanie ścieżki do danego elementu.
Podano ścieżkę do zakładki 'Polska' na stronie interii: https://fakty.interia.pl/polska.
Po podaniu tego adresu przeglądarka przenosi nas do żądanej strony.



c) Odniesienie do programu pocztowego

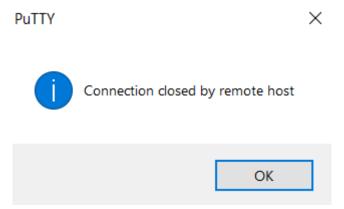
Odniesienie do programu pocztowego następuje poprzez podanie do niego ścieżki. Podano ścieżkę do usługi pocztowej serwisu interia: https://poczta.interia.pl/?referer=https%3A%2F%2Ffakty.interia.pl%2Fpolska&crc=2693211881. Po podaniu adresu zostajemy przeniesieni na stronę poczty.



5. Protokół http

Protokół HTTP jest protokołem bezstanowym, co oznacza, że serwer WWW rozpatruje każde żądanie niezależnie od innych nie szukając żadnych powiązań między nimi, np. czy były wysłane przez tego samego nadawcę.

W przypadku gdy opcja **close window on exit** jest ustawiona na **never** otrzymujemy powiadomienie, że połączenie zostało przerwane przez innego hosta, a okno nie ulega zamknięciu.



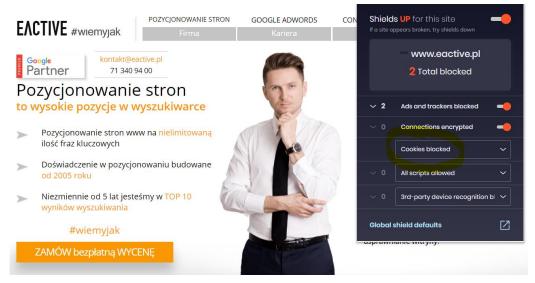
W przypadku gdy opcja **close window on exit** jest ustawiona na **always**, to po uzyskaniu połączenia następuje zamknięcie okna bez komunikatu zwrotnego. Bezstanowość objawia się tym, że inny host jest w stanie przerwać ustalone przez nas połączenie, ponieważ protokół http nie zwraca uwagi na nadawcę, gdy wybrana jest opcja **always** po przerwaniu połączenia następuje bezzwłoczne zamknięcie okna, nie wiemy dlaczego połączenie uległo przerwaniu.

6. Pliki cookie

Pliki cookie - czyli "ciasteczka" to niewielkie informacje wysyłane przez serwis internetowy, który odwiedzamy i zapisywane na urządzeniu końcowym (komputerze, laptopie, smartfonie), z którego korzystamy podczas przeglądania stron internetowych. W cookies, składających się z szeregu liter i cyfr, znajdują się różne informacje niezbędne do prawidłowego funkcjonowania serwisów internetowych, np. tych wymagających autoryzacji – m.in. podczas logowania do konta pocztowego czy sklepu internetowego.

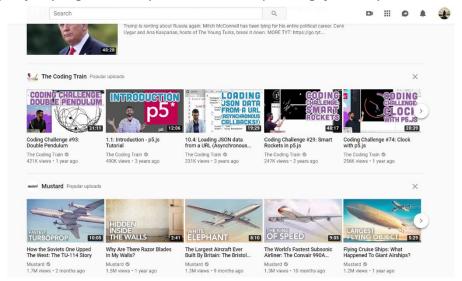
a) Sprawdzono działanie plików cookie na stronie eactive.pl.

Używamy cookies i podobnych technologii. Szczegóły znajdziesz w Polityce Cookies. Po zablokowaniu plików cookie w przeglądarce **Brave** nie zauważono żadnych zmian w funkcjonowaniu strony.

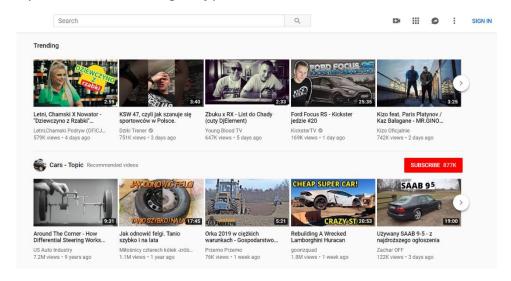


b) Sprawdzono działanie plików cookie na stronie **youtube.com**.

W przypadku gdy przeglądarka zezwala na użycie plików cookie na stronie tytułowej otrzymujemy zagadnienia spersonalizowane pod względem użytkownika.



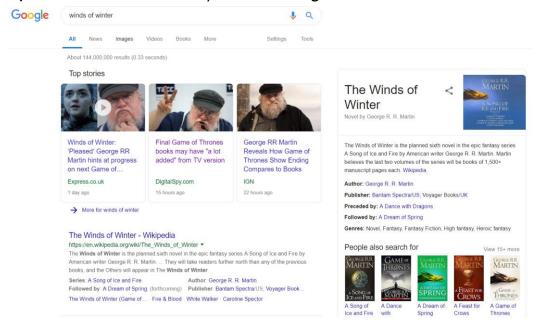
W przypadku zablokowania plików cookie strona zwraca zagadnienia popularne lub domyślnie kierowane do ogólnej publiczności.



7. Testowanie wyszukiwarek internetowych

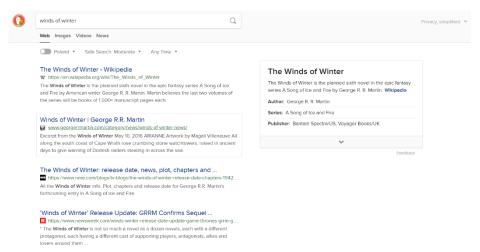
Przetestowano działanie pięciu przeglądarek pod względem zwracanych rezultatów, ilości trafień oraz szybkości działania. Sprawdzono również możliwość wykorzystania dodatkowych narzędzi np. kalkulator, kurs walut. W przeglądarkach wyszukiwano informacji na temat szóstej książki z serii Pieśń Lodu i Ognia p.t. "Winds of Winter" autorstwa George'a R. R. Martina.

a) Przetestowano działanie wyszukiwarki Google.



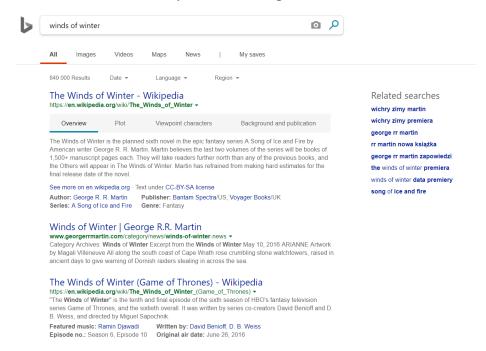
Wyszukiwarka zwraca najnowsze informacje związane z książką, najważniejsze informacje na jej temat z serwisu **Wikipedia**. Otrzymano łącznie **144,000,000** wyników na temat wyszukanej frazy w przeciągu **0.33** sekundy.

b) Przetestowano działanie wyszukiwarki **DuckDuckGo**.



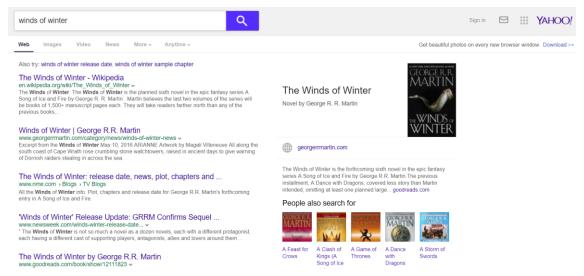
Wyszukiwarka ta zwraca informacje na temat książki zawarte w serwisie **Wikipedia** oraz linki do strony, na której znajduje się blog autora, oraz do popularnych wyników. Nie otrzymujemy tak jak w przypadku **Google** najnowszych informacji prasowych oraz informacji o poprzednich tytułach z serii. Nie otrzymujemy informacji o liczbie znalezionych wyników oraz o czasie wyszukiwania.

c) Przetestowano działanie wyszukiwarki Bing.



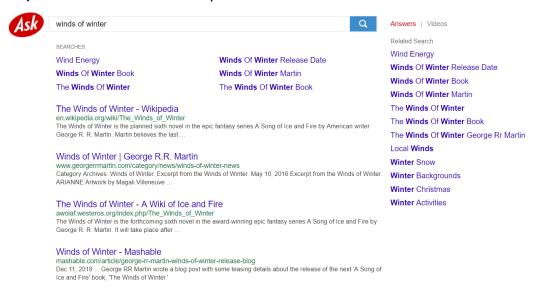
Wyszukiwarki zwraca nam informacje o tytule zawarte w serwisie **Wikipedia** oraz powiązane hasła. Otrzymujemy informację o liczbie wyników, która wynosi **849000**, jednak nie wiemy ile czasu zajęło wyszukiwanie.

d) Przetestowano działanie wyszukiwarki Yahoo.



Wyszukiwarka, tak jak w powyższych przykładkach, zwraca informację o książce z serwisu **Wikipedia**. Dodatkowo informuje nas o innych tytułach z serii. Nie otrzymujemy informacji o ilości otrzymanych wyników oraz ile trwało ich wyszukanie.

e) Przetestowano działanie wyszukiwarki Ask.



Wyszukiwarka zwraca informację o powiązanych tematach oraz hasła posiadające frazy podane wyszukiwarce. Nie otrzymujemy informacji o czasie jaki był potrzebny do uzyskania otrzymanych wyników oraz ilości tych wyników.

8. Serwery proxy

Serwer proxy, czyli serwer pośredniczący– daje możliwość przeglądania stron internetowych, lub prowadzenia innych operacji z wykorzystaniem (poprzez) tych że serwerów. Łącząc się za pomocą proxy najpierw łączymy się z odpowiednim serwerem a dopiero później oglądamy zawartość serwisu, ale udostępnioną nam przez proxy. Zapewnia to użytkownikowi szybkość– w przypadku gdy połączenie z serwerem proxy jest szybsze niż z stroną wywoływaną i anonimowość.

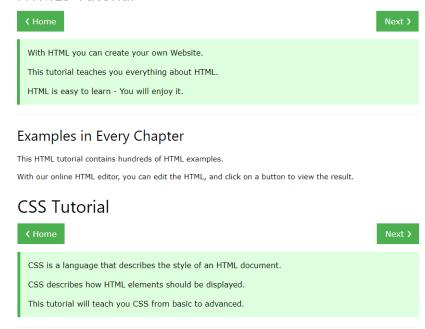
9. Porównanie HTTP i HTTPS

HTTPS jest to szyfrowaną wersją protokołu http, która szyfruje połączenie z wykorzystaniem protokołu SSL. Obecnie wykorzystuje się również protokół TLS będący rozwinięciem SSL. HTTPS został opracowany, aby w bezpieczny sposób umożliwić transakcje internetowe i uniemożliwić dostęp nieautoryzowanym użytkownikom. Dane przesyłane za pomocą protokołu HTTPS korzystające z protokołu TLS wykorzystują trzy ochronne warstwy: Szyfrowanie – wymiana danych, aby były bezpieczne. Integralność danych – dane nie mogą być zamienione lub uszkodzone podczas transferu. Uwierzytelnienie – potwierdzenie zadeklarowanej tożsamości użytkownika.

10. Specyfikacja HTTP oraz CSS

Zapoznano się ze specyfikacją HTTP i CSS umieszczonymi na stronie http://www.w3schools.com.

HTML5 Tutorial



Examples in Each Chapter

This CSS tutorial contains hundreds of CSS examples.

With our online editor, you can edit the CSS, and click on a button to view the result.

11. Modyfikacja strony internetowej

Poniżej zamieszczono podgląd strony przed modyfikacją.



Arkadiusz Marta

Polskie znaki: Ćma śmieje się z chrząszcza

Button cokolwiek mozna tu napisac

Strona testowa

Poniżej zamieszczono podgląd strony po modyfikacji.





Poniżej zamieszczono kod strony.

Strona testowa

Przy użyciu **CSS** zmodyfikowano rozmiary tabeli oraz zamieszczonej na stronie fotografii. Zmieniono również kolor i charakter obramowania tabeli oraz rozmiar czcionki napisów wewnątrz tabeli.