

Wykład 1 – TINT

Wprowadzenie do XHTML

Zofia Kruczkiewicz

Sieć WWW (*World Wide Web*)

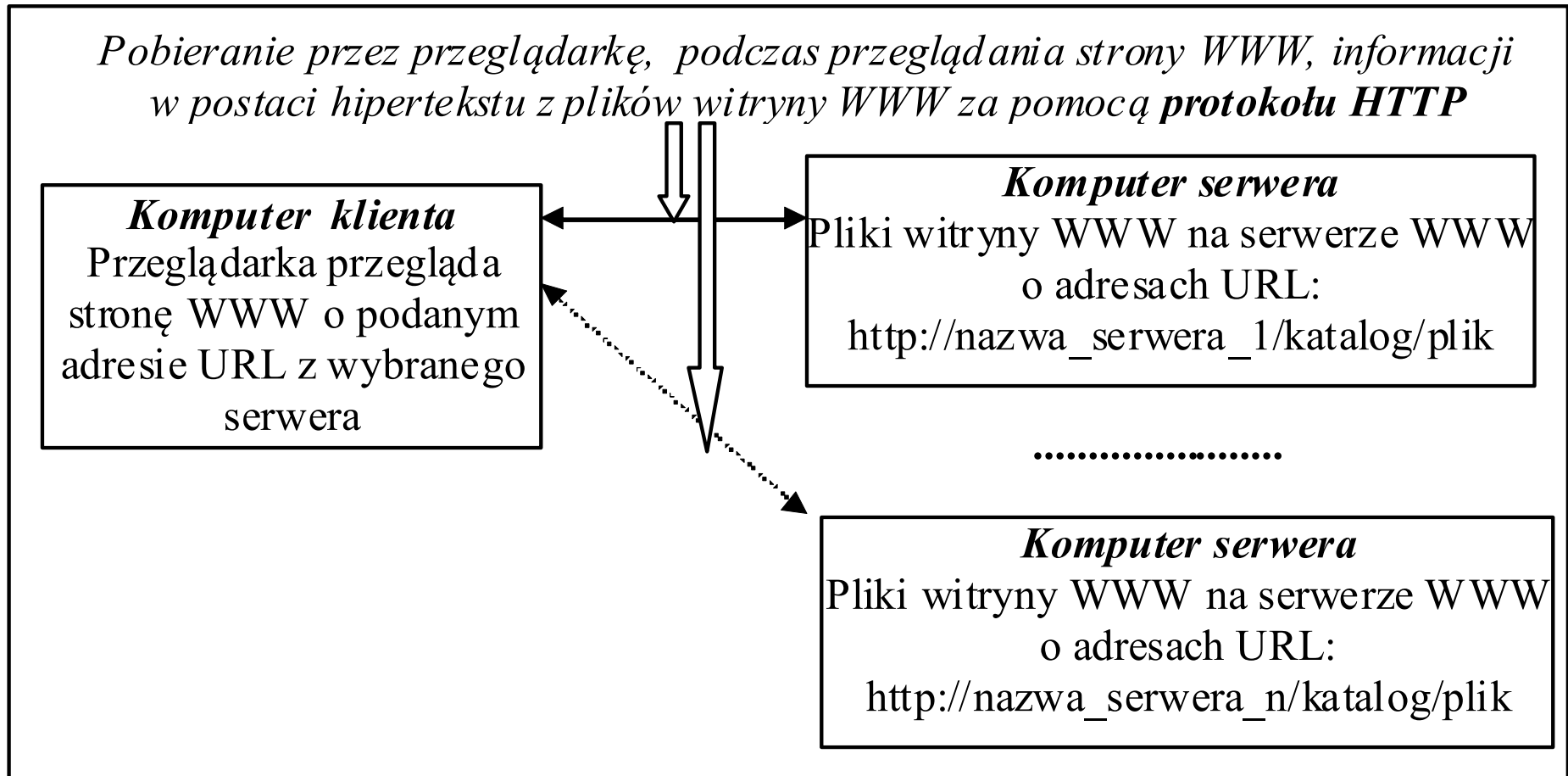
Sieć WWW jest internetowym systemem informacyjnym o charakterze:

- hipertekstowym,
- globalnym,
- interaktywnym (np. zastosowanie formularzy)
- dynamicznym (DHTML)
- wieloplatformowym
- rozproszonym
- graficznym.

Główne grupy organizacji wywierające wpływ na wygląd sieci WWW

- **World Wide Web (W3) Consortium** w Massachusetts Institute of Technology (określają i definiują języki i protokoły w sieci WWW)
- <http://www.w3.org/>

Sieć WWW jako internetowy system informacyjny

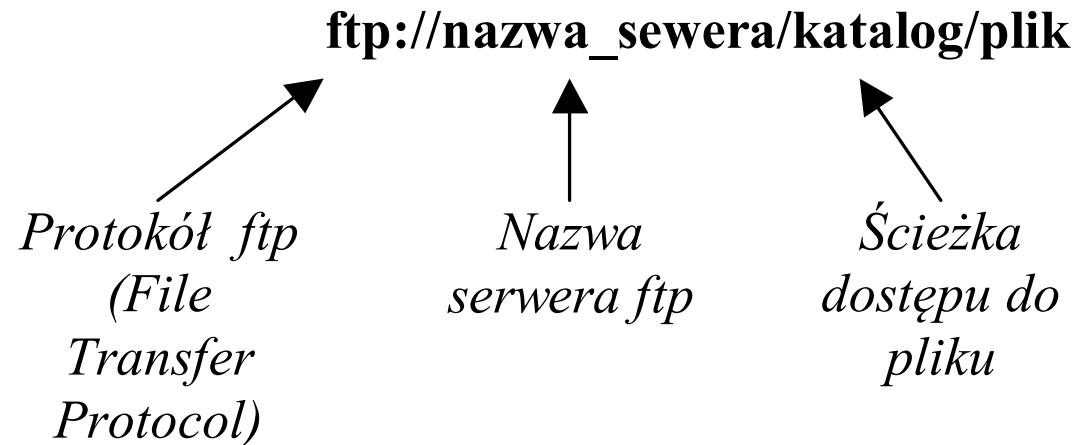
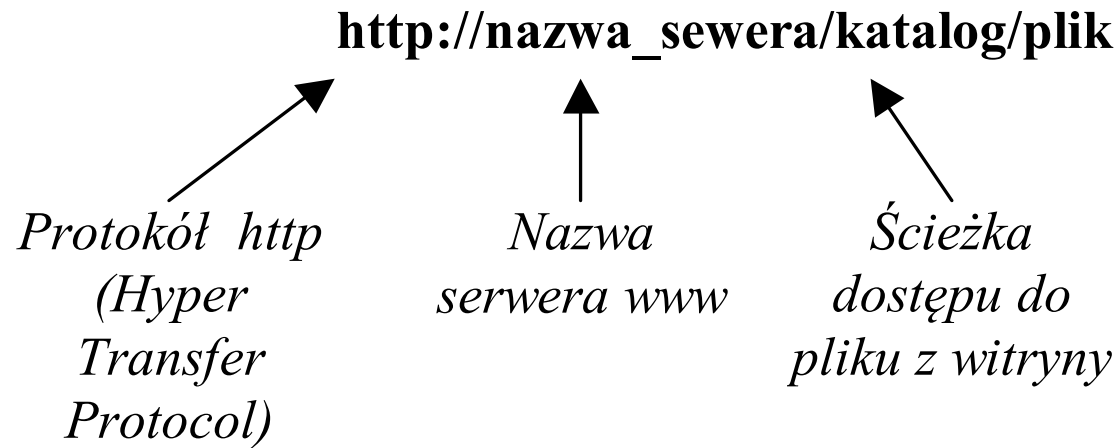


Rys.1. Podstawowy schemat czytania strony WWW w rozproszonej, globalnej sieci WWW

Pobieranie stron html przez przeglądarki

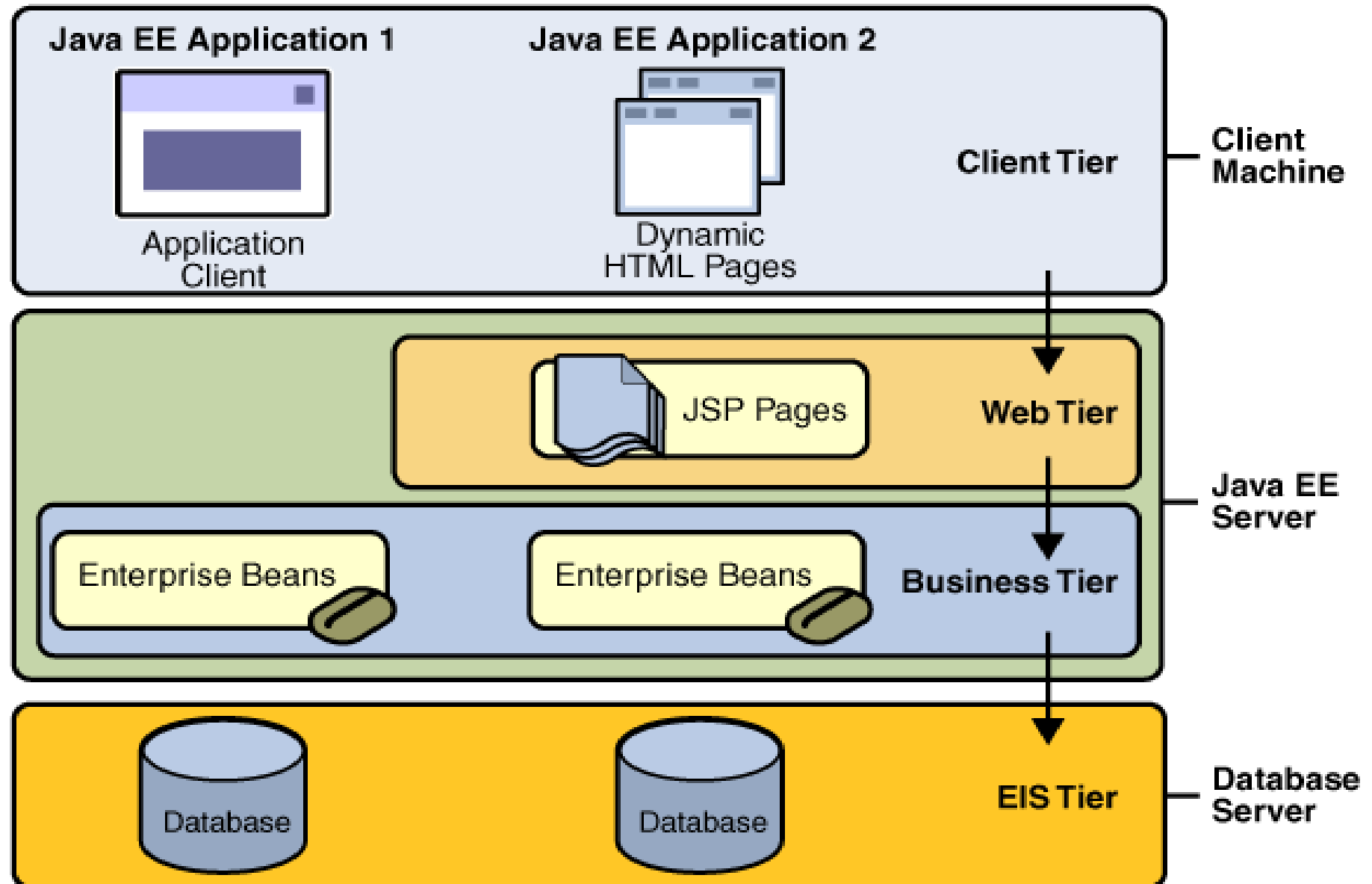
- Witryna WWW znajduje się na serwerze (lub serwerach) WWW i stanowi miejsce, gdzie publikowane są pewne informacje w postaci plików.
- Pliki te są pobierane podczas przeglądania strony WWW przez:
 - dowolną przeglądarkę HTML
 - z dowolnego komputera klienta
 - w standardowej postaci hipertekstu.

Sieć WWW jako sieć globalna rozproszona

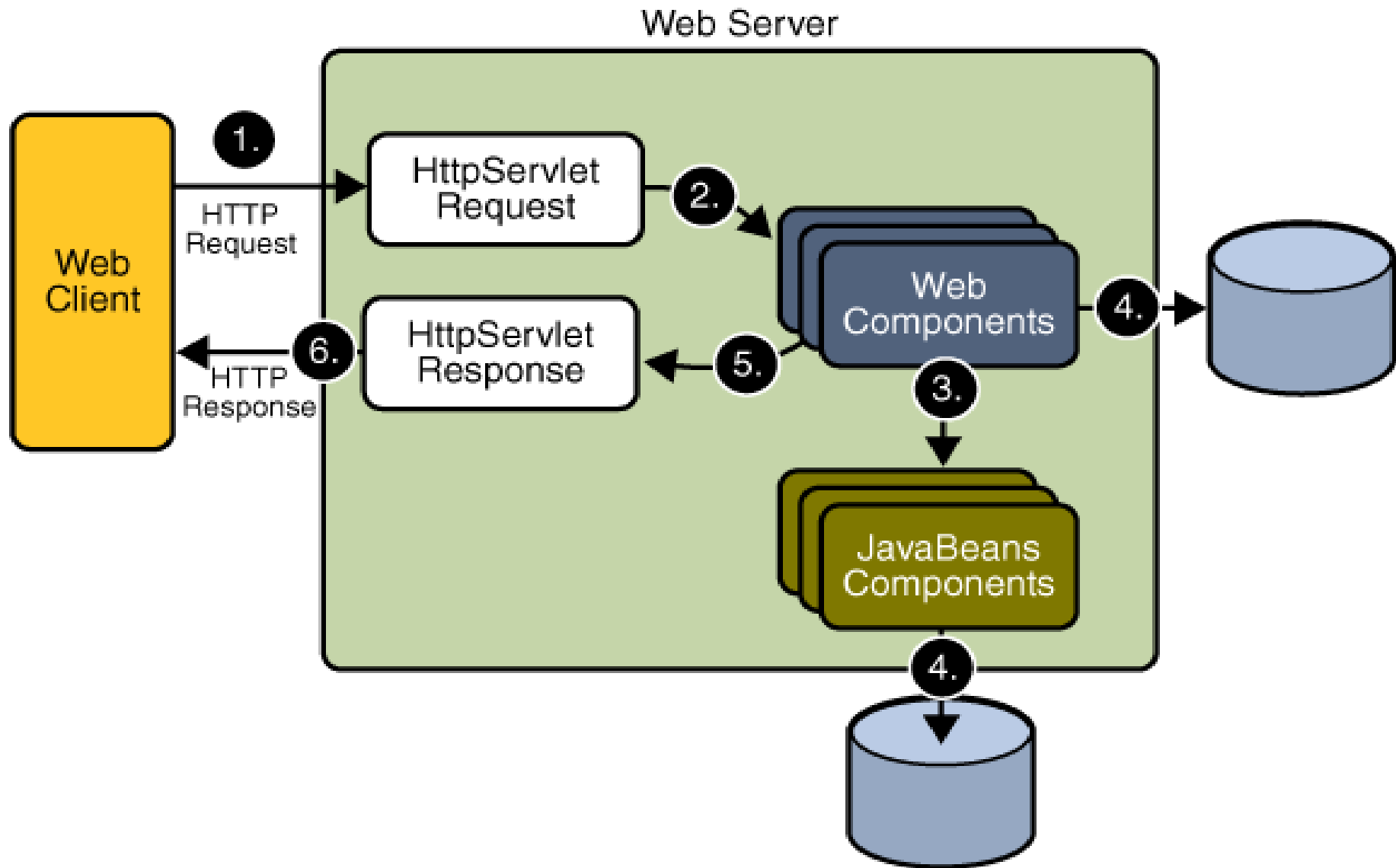


Różne typy adresów URL w postaci podstawowej

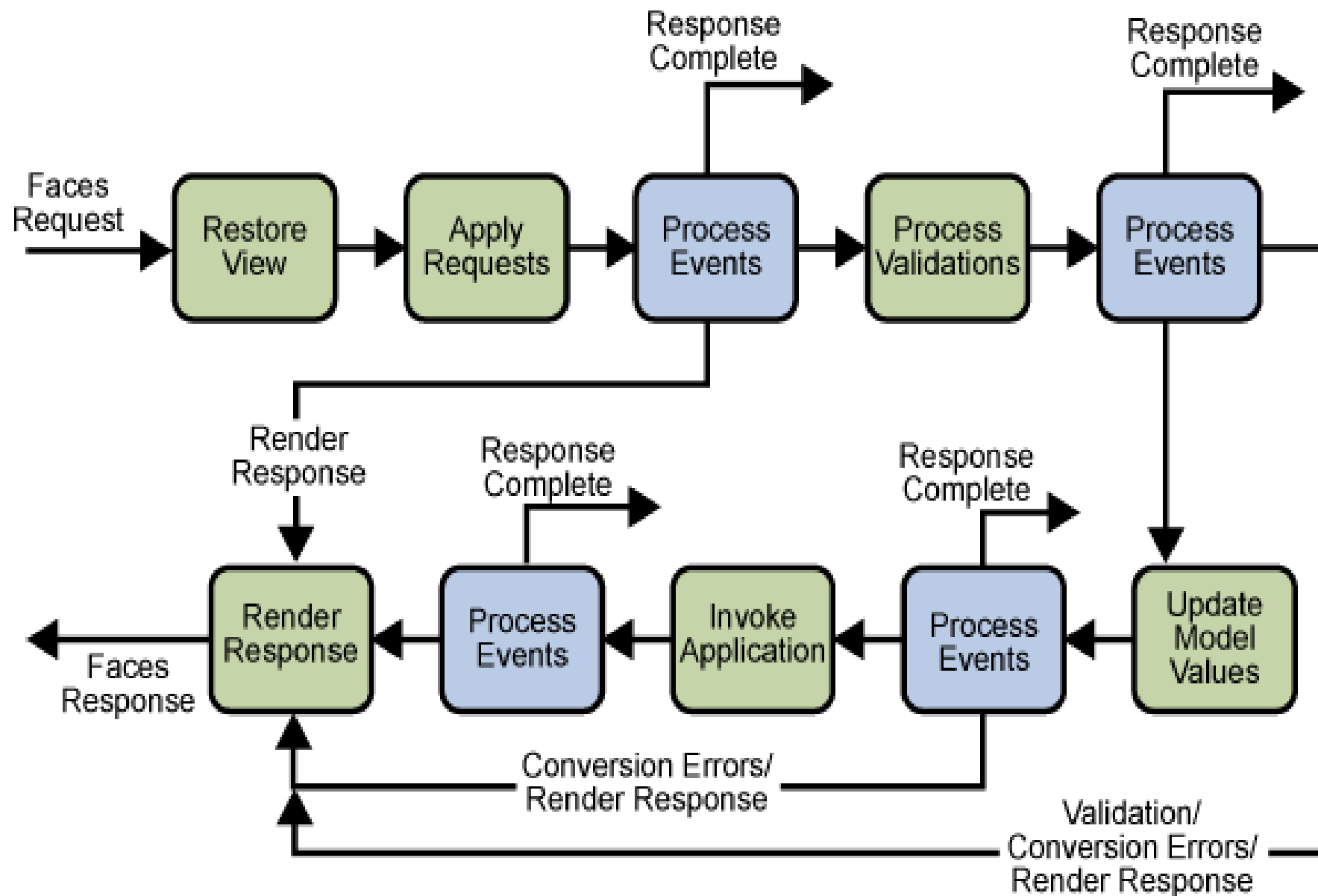
Wielowarstwowe aplikacje w JavaEE5



Podstawowa struktura działania aplikacji internetowej



Standard cyklu życia „Request-Response” dla JavaServer Faces



Charakter hipertekstowy sieci WWW

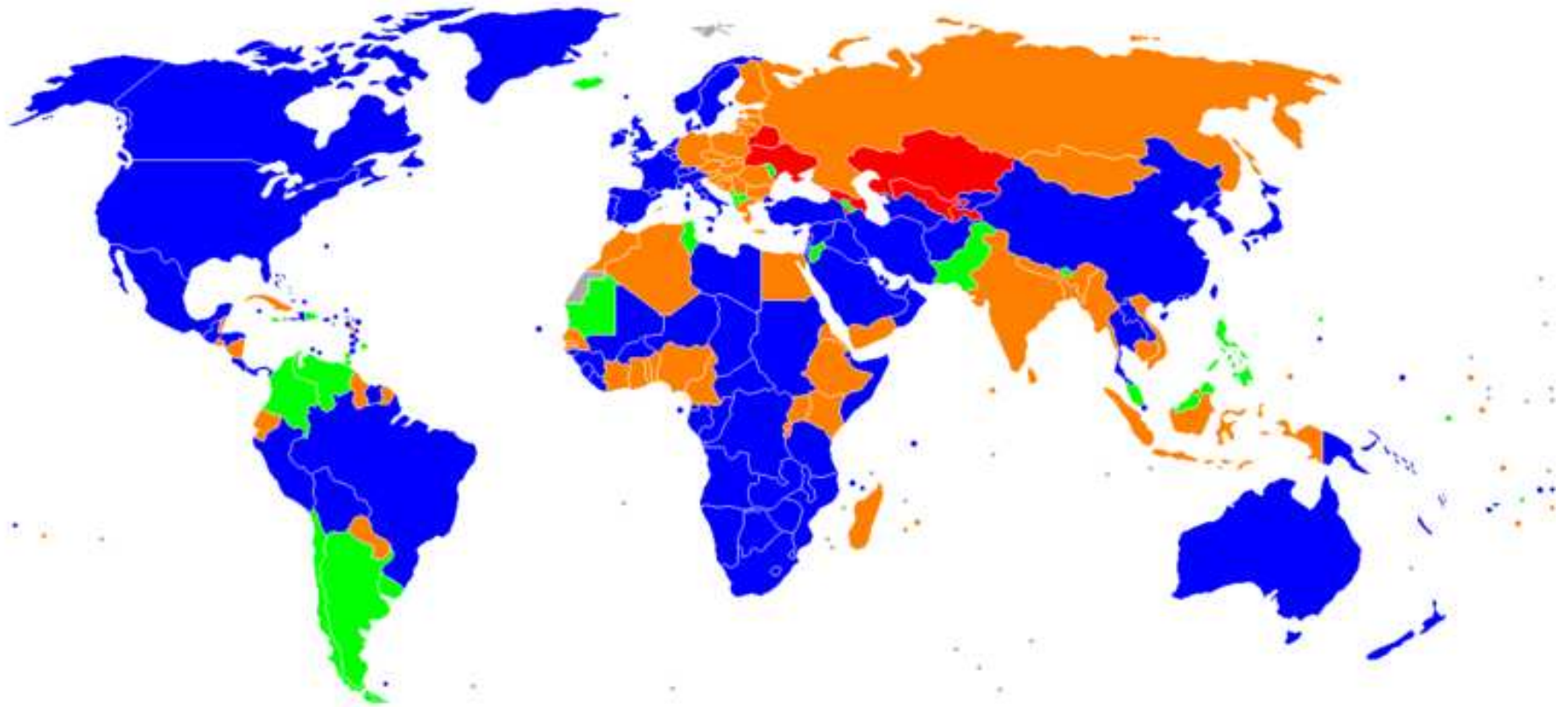
- Serwer WWW jest programem pracującym na komputerze w sieci Internet (np. Apache), umożliwiającym przesyłanie plików na żądanie przeglądarek klientów.
- Serwer WWW powinien zapewnić przez użyciu odpowiednich programów przekształcenie publikowanych w witrynach informacji na standardową postać hipertekstową.
- Hipertekst umożliwia przetwarzanie tekstu na stronach WWW (czytanie, poruszanie się, zapisywanie) w sposób niesekwencyjny w oparciu o dodatkowe informacje dołączone do tekstu.
- **Obecnie hipertekst ma charakter hipermedialny, ponieważ łączy tekst z grafiką, dźwiękiem oraz wideo.**
- Do przetwarzania stron utworzonych za pomocą hipertekstu (lub przekształconych do postaci hipertekstu) służą przeglądarki:
Netscape
 - Microsoft Internet Explorer
 - Mozilla Firefox
 - Safari
 - Opera
 - Przeglądarki mobilne

Sieć WWW jako sieć wieloplatformowa

- Dostęp do Internetu umożliwia dostęp do sieci WWW z dowolnego komputera, który ma zainstalowany dowolny system operacyjny (Windows, Unix, Dos), ponieważ WWW jest siecią **wieloplatformową**.
- Obecnie idea ta jest ograniczana nowymi technologiami internetowymi. Nowe technologie zastosowane do budowy strony WWW mogą uniemożliwić odczytanie jej przez przeglądarkę z dowolnej platformy w przypadku **braku wymaganych narzędzi** (np. dodatków typu plug-ins).

Udział przeglądarek na rynkach światowych (dane z maja [2011](#) wg serwisu StatCounter)

http://en.wikipedia.org/wiki/File:Countries_by_most_used_web_browser.png



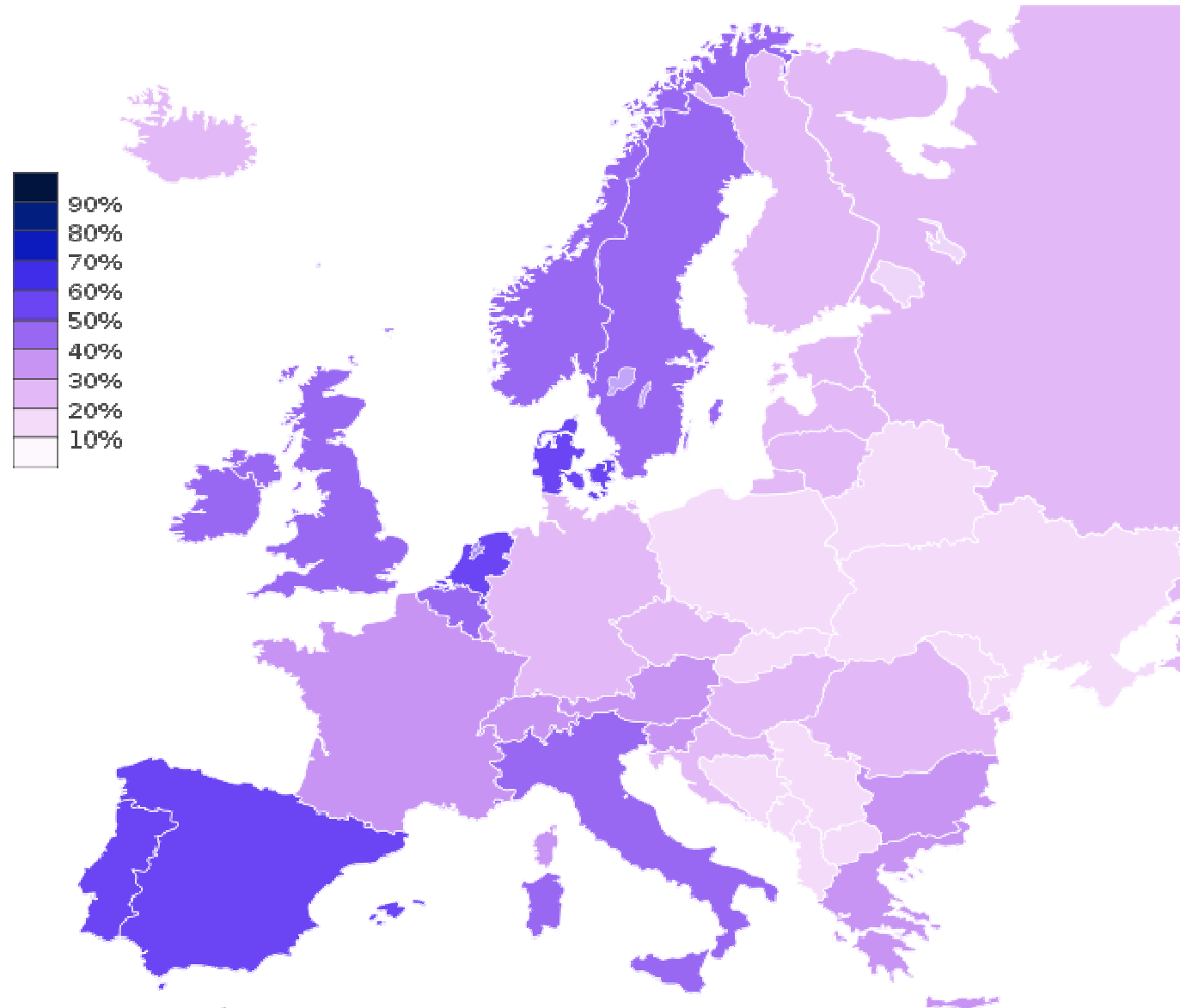
Niebieski: [MS Internet Explorer](#)

Pomarańczowy: [Mozilla Firefox](#)

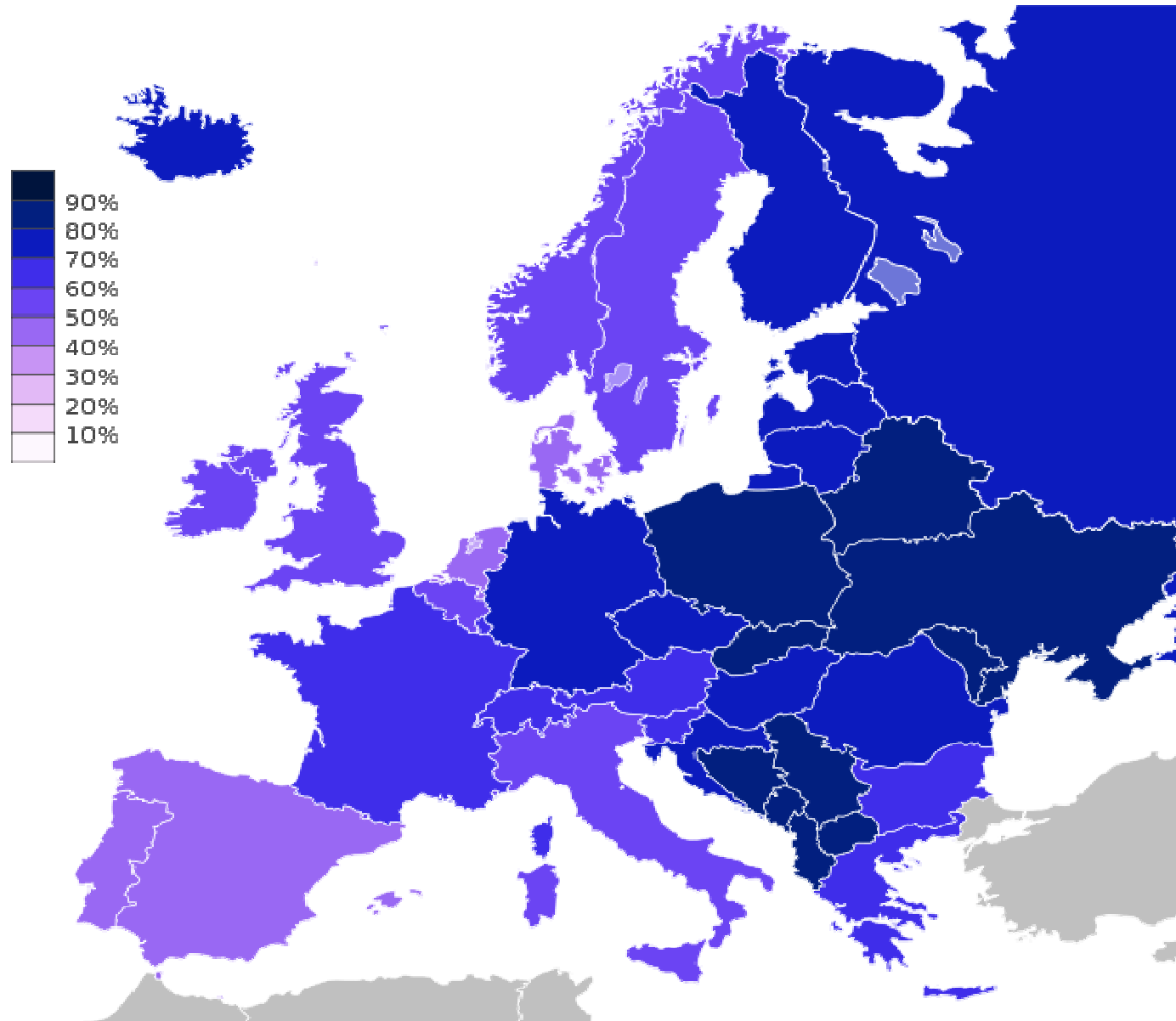
Zielony: [Google Chrome](#)

Czerwony: [Opera](#)

Pozycja Internet Explorera na rynkach europejskich (dane z kwietnia 2011 wg serwisu StatCounter)



Udział alternatywnych przeglądarek na rynkach europejskich (dane z kwietnia 2011 wg serwisu StatCounter)



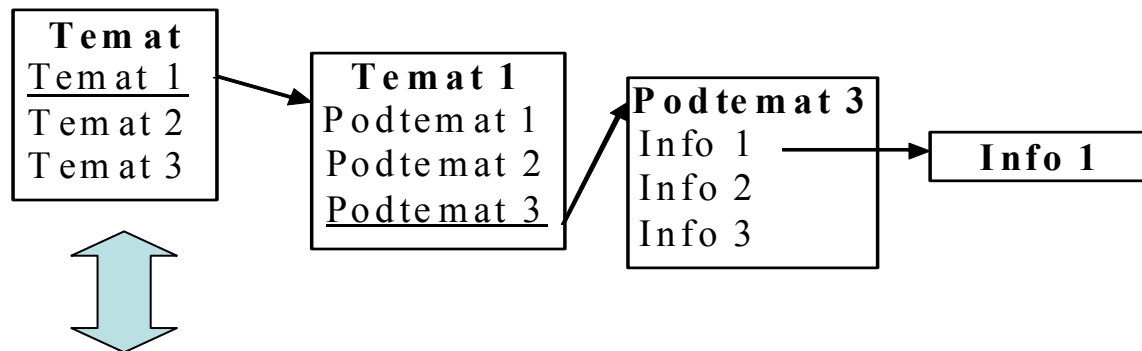
Podstawowe pojęcia

- Serwer WWW – oprogramowanie zarządzające plikami witryn oraz danymi
- Prezentacja WWW - zbiór jednej lub wielu stron WWW
- Witryna WWW - system zawierający jedną lub wiele prezentacji
- Strony WWW- pojedynczy element prezentacji, ograniczony do jednego pliku
- Strony główne – strona inicjująca działanie prezentacji

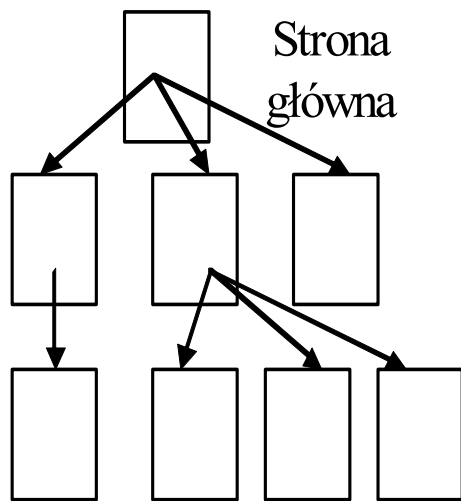
Zawartość prezentacji WWW

- informacja osobista
- hobby i zainteresowania
- publikacja
- prezentacje firm
- dokumentacja elektroniczna
- katalogi towarów
- wirtualne sklepy
- sondaże opinii publicznej i głosowania
- edukacja
- inne

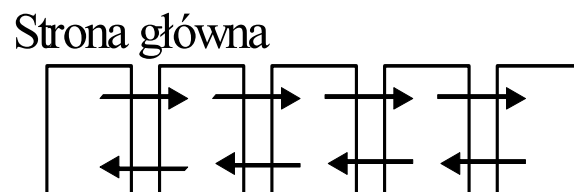
Struktury dokumentów, Hipertekst w systemie pomocy kontekstowej na stronach w sieci WWW



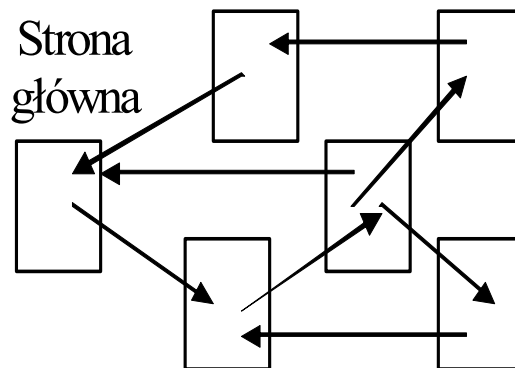
- hierarchiczna
- liniowa
- liniowa z alternatywą
- sieciowa
- mieszana



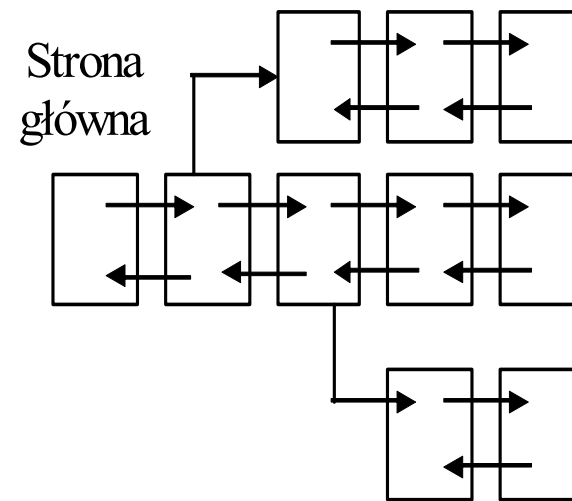
Struktura hierarchiczna



Struktura liniowa



Struktura sieciowa (pajęczyna)



Struktura liniowa z alternatywą

HTML

- HTML (*Hyper Text Markup Language*) jest implementacją języka SGML (*Standard Generalized Markup Language*). Język SGML służy do tworzenia dokumentów, opisujących strukturę dokumentu.
- HTML jest językiem znaczników służącym do:
 - opisu struktury dokumentu
 - formatowania tekstu
 - łączenia grafiki, dźwięku, video itd.
 - oddzielenia struktury dokumentów od prezentacji (stosuje się dwa typy plików: ze strukturą dokumentów i arkusze stylów, które pozwalają tę samą strukturę dokumentów prezentować na wiele sposobów).
- Metaznaczniki stanowią podstawę budowy dokumentu HTML i opisują strukturę dokumentu i format dokumentu.

Standardy HTML

- **Standardy HTML**
- HTML 2.0
- HTML 3.2 (1996)
- HTML 4.0 (1997)
- HTML 4.01 (1999) -
<http://www.w3.org/TR/html401/>
- XHTML (na bazie XML i HTML 4.0)
<http://www.w3.org/TR/xhtml1/>

Zadanie edytorów HTML

- **Sprawdzanie poprawności pisowni**
- **Edycja informacji nagłówkowych.**
- **Projekty.**
- **Szablony.**
- **Biblioteki wstawek.**
- **Obsługa języków.** języka HTML, JavaScript, CSS
- **Kolorowanie kodu.** .
- **Wprowadzanie znaków specjalnych.** spoza alfabetu łacińskiego
- **Mapowanie odsyłaczy.**
- **Podgląd dokumentu.**
- **Kontrola poprawności składniowej.**
- **Kontrola spójności.**
- **Weryfikacja odsyłaczy.**
- **Wysyłanie na serwer**

Przykłady edytorów HTML

- **Edytory znaczników:** płatne edytory (Pajaczek NxG, Tiger98 i TigerII Milenium) i darmowe edytory (edHTML, ezHTML, MiniPad, Snape, TigerII, WebPager Xpress, Website Pro, Zajaczek PHP)
- **Edytory WYSIWYG** (*What You See Is What You Get*) np. Microsoft FrontPage, Macromedia Dreamweaver, NetObjects Fusion
- **Programy konwertujące** pliki utworzone za pomocą znanych edytorów do formatu HTML

Dokumenty XHTML

- Dokument XHTML jest dokumentem tekstowym, który zawiera znaki o ustalonym kodowaniu, składającym się z elementów (tagów/znaczników) pisanych małymi literami.
- Element składa się z trzech części (większość elementów):
 - znacznika początku (w formie **<znacznik>**)
 - zawartości elementu (tekst i/lub inne elementy)
 - znacznika końca (w formie **</znacznik>**)
- Elementy mogą mieć przypisane atrybuty określające znaczenie danego elementu. Atrybuty wraz z ich wartościami umieszczonymi w cudzysłowach wpisuje się wewnątrz znacznika początkowego
<znacznik atrybut="wartość">.

W języku XHTML każdy znacznik musi posiadać znacznik końca, co nie obowiązywało w wielu znacznikach w HTML.

Typy dokumentów XHTML

- W dokumentach XHTML stosuje się specjalny prolog, który identyfikuje poziom używanego języka np. XHTML 1.0.
- Jest on wstawiany jako pierwszy element dokumentu, **jeszcze przed** otwarciem szkieletu strony `<html>`.
- Prolog jest m.in. wykorzystywany jako oznaczenie poziomu w procesie weryfikacji poprawności składni za pomocą tzw. parserów np. [W3C HTML Validation Service](http://validator.w3.org/)
<http://validator.w3.org/>.

Specyfikacja XHTML 1.0 przewiduje trzy obowiązkowe wersje prologu.

- **Typ "Strict"**

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

Wersja **Strict** definicji typu dokumentu (DTD) wyklucza wszelkie elementy prezentacyjne. Wersja **Strict** jest okrojonym HTML 4, przedkładającym strukturę nad prezentację.

- **Typ "Transitional"**

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

Często stosowaną wersją prologu jest tzw. wersja przejściowa **Transitional** dla dokumentów zawierających elementy i atrybuty HTML,

- **Typ "Frameset"**

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

Szerszą odmianą **Transitional** jest prolog **Frameset** dla strony zawierającej ramki .

Metaznaczniki

<meta> - **opcjonalne polecenie** usprawniające funkcjonowanie witryny w sieci, które bardziej szczegółowo informuje o zawartości dokumentu, np. dla celów katalogowych czy indeksów, wykorzystywanych przez odpowiednie programy, np. wyszukiwarki sieciowe lub przeglądarki internetowe

Zawiera ono trzy atrybuty: **http-equiv**, **name** i **content** w kombinacji pierwszy z trzecim lub drugi z trzecim.

http-equiv – definiuje zmienne systemowe, natomiast

name – definiuje zmienne użytkownika.

content – wartość zmiennych

Przykłady:

1) **<meta http-equiv="content-type" content="text/html; charset=windows-1250 "/>**

Strona kodowa, opis strony, wyrazy kluczowe, autor strony.

Standard wyświetlania automatycznie przyjęty przez wyszukiwarke i koniecznie zastosowany podczas edycji strony w edytorze.

2) **<meta name ="description" content=" treść..."/>**

Polecenie opisuje zawartość strony. Znaleziona strona będzie zatytułowana za pomocą **title** i opisana za pomocą **content** (150-200 znaków, lub 20-25 wyrazów).

3) **<meta name="keywords" content="wyrazy kluczowe..."/>**

Polecenie informuje o wyrazach kluczowych dokumentu, zwracając przykładową wartość keywords: HTML, znaczniki itp. (do 1000 znaków). Ułatwia to pracę sieciowym programom indeksująco-wyszukiwawczym i zwiększa szansę znalezienia strony przez innych użytkowników.

4) **<meta name="Author" content="imię i nazwisko"/>**

Polecenie informuje o autorze strony.

5) **<meta http-equiv="content-language" content="pl"/>**

Polecenie informuje o języku strony. Niektóre przeglądarki korzystają z niego przy precyzowaniu zapytań. Dla języka angielskiego użyjemy **en**, dla amerykańskiego angielskiego **en-us**, dla niemieckiego **de**, dla francuskiego **fr**, dla rosyjskiego "ru" itd.

6) **<meta name="generator" content="nazwa edytora"/>**

Polecenie informuje o użytym narzędziu do tworzenia stron HTML.

7) **< meta http-equiv="expires" content="data i czas"/>**

Polecenie wskazuje przeglądarce, kiedy dokument traci "ważność" i trzeba go wczytać na nowo z Sieci. Data musi być podawana w formacie zdefiniowanym przez dokument RFC850.

<meta http-equiv="expires" content="Wed, 26 Apr 2001 08:21:57 GMT"/>

8) **<meta http-equiv="pragma" content="no-cache"/>**

Polecenie określa, czy strona ma być cache'owana na lokalnym dysku. Użycie **no-cache** zakazuje przeglądarce tej czynności i za każdym razem strona jest wczytywana z Sieci.

9) **<meta http-equiv="refresh" content="t"/>**

Polecenie spowoduje regularne odświeżanie strony co *t* sekund. Może to mieć praktyczne zastosowanie w przypadku, gdy strona zawiera bardzo często aktualizowane informacje (wiele razy dziennie).

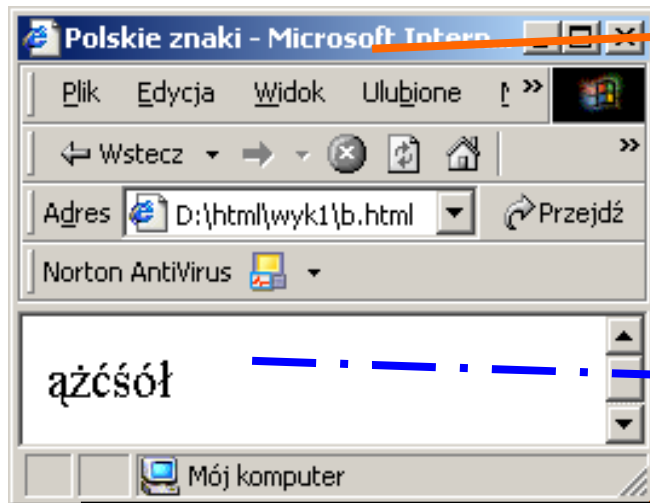
10) **<meta http-equiv="refresh" content="x; url=http://.../strona.html"/>**

Polecenie spowoduje automatyczne przekierowanie, czyli wczytanie nowej strony po *x* sekundach.

11) **<meta http-equiv="creation-date" content="Wed, 21 Sep 2005 21:29:02 GMT"/>**

Polecenie informuje o dacie utworzenia dokumentu, zwracając wartość Creation-Date: Wed, 21 Sep 2005 21:29:02 GMT.

(1) Struktura dokumentu XHTML



Tytuł

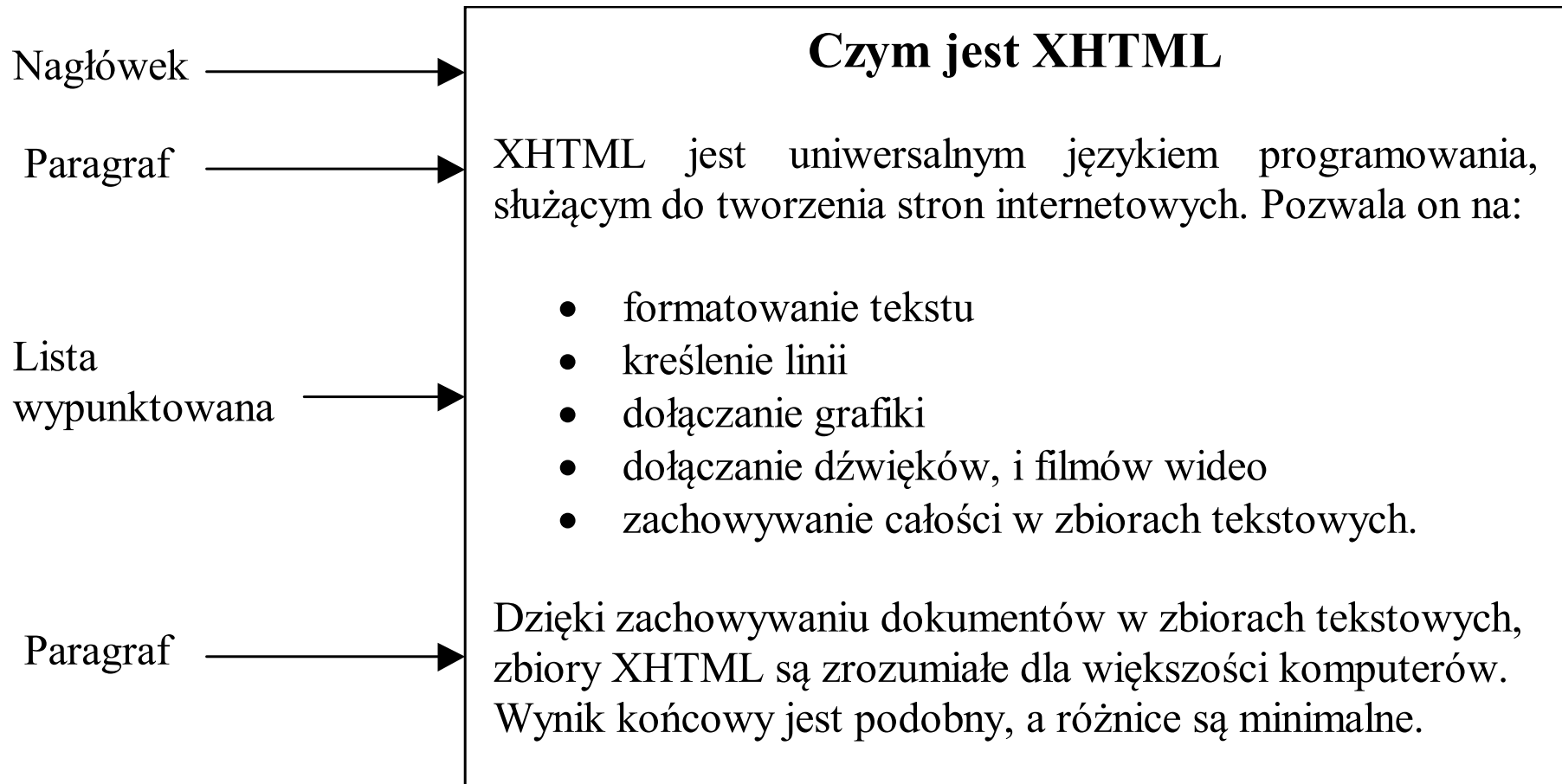
Kodowanie znaków
polskich w standardzie
UNICODE

Paragraf

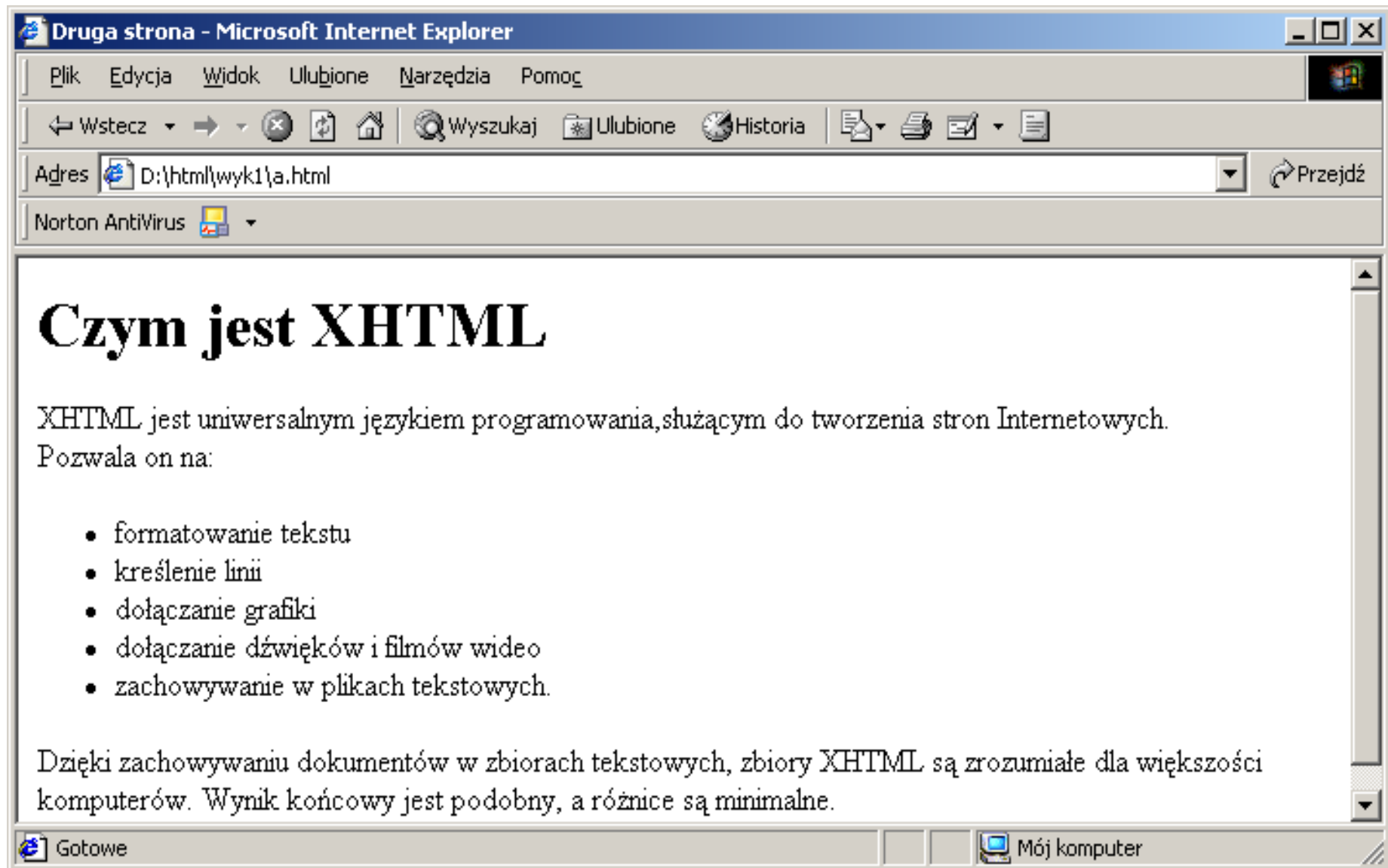
Deklarowanie
języka polskiego

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="pl" lang="pl">
<head>
  <meta http-equiv="Content-Type" content="text/html; charset=utf-8"/>
  <title>Polskie znaki</title>
</head>
<body>
  <p>ąźćśół</p>
</body>
</html>
```

(2) Elementy dokumentu XHTML (wg L.Lemay: *HTML 4*) - projekt



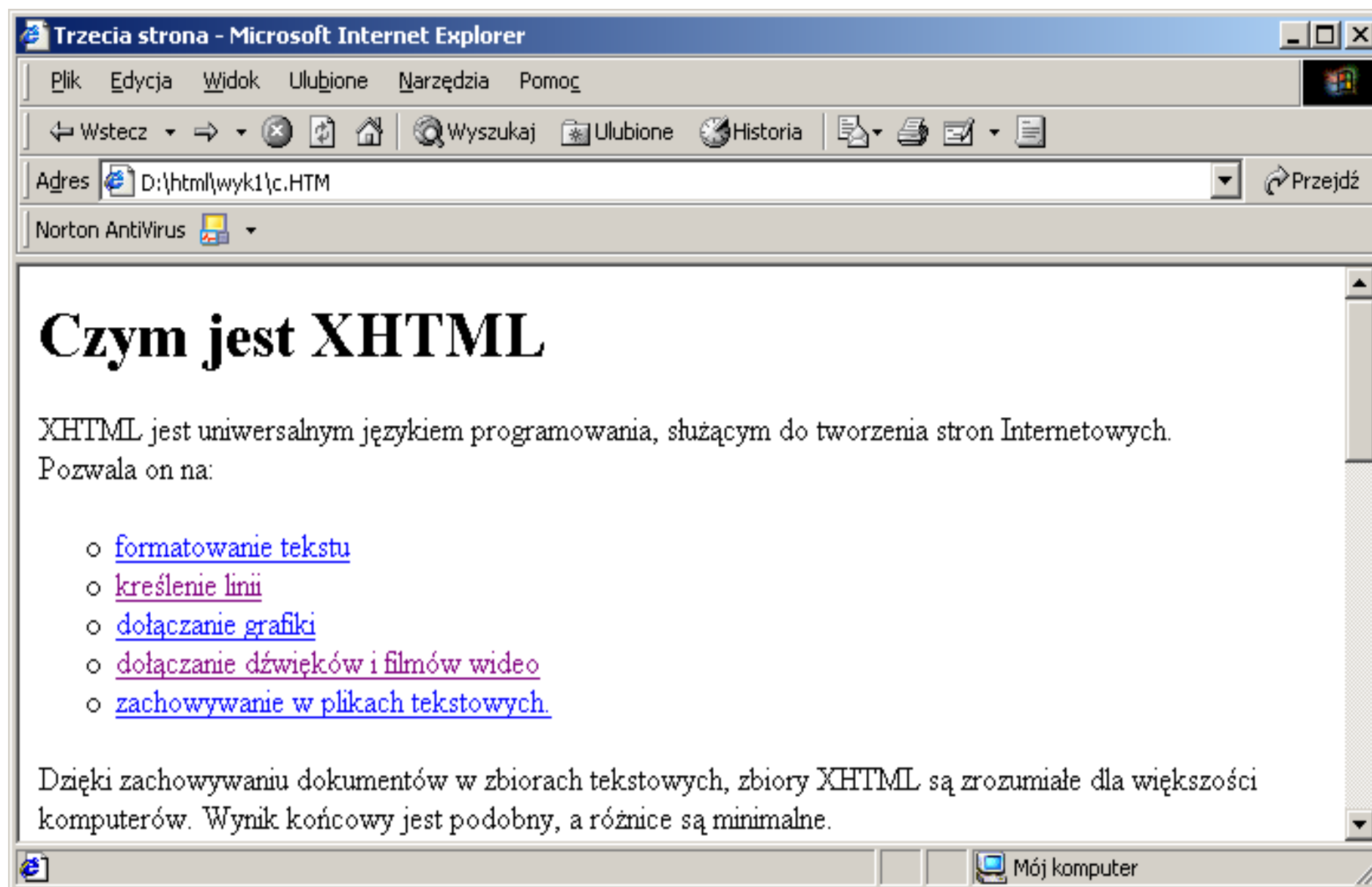
(2a) Widok strony w przeglądarce Microsoft Internet Explorer



(2b) Kod źródłowy XHTML/HTML

```
<!--Komentarz-->
<!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="pl" lang="pl">
  <head>
    <meta http-equiv="content-type" content="text/html; charset= utf-8"/>
    <title> Druga strona</title>
  </head>
  <body>
    <h1>Czym jest XHTML</h1>
    <p>XHTML jest uniwersalnym językiem programowania,
      służącym do tworzenia stron Internetowych.<br/> Pozwala on na:</p>
    <ul>
      <li>formatowanie tekstu</li>
      <li>kreślenie linii</li>
      <li>dołączanie grafiki</li>
      <li>dołączanie dźwięków i filmów wideo</li>
      <li>zachowywanie w plikach tekstowych.</li>
    </ul>
    <p>Dzięki zachowywaniu dokumentów w zbiorach tekstowych, zbiory XHTML są
      zrozumiałe dla większości komputerów. Wynik końcowy jest podobny, a różnice są
      minimalne.</p>
  </body>
</html>
```

(3) Strona html



(3a) Połączenia i odnośniki

```
<!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="pl" lang="pl">
<head>
  <meta http-equiv="content-type" content="text/html; charset=utf-8"/>
  <title> Trzecia strona</title>
</head>
<body>
<h1>Czym jest XHTML</h1>
<p> XHTML jest uniwersalnym językiem programowania, służącym do tworzenia stron
  Internetowych. <br/> Pozwala on na: </p>
<ul type="circle">
  <li><a href="c1.htm">formatowanie tekstu</a></li>
  <li><a href="http://sprocket.ict.pwr.wroc.pl/~zkruczkiewicz"> kreślenie linii</a></li>
  <li><a href="c3.htm#Miejsce1">dołączanie grafiki</a></li>
  <li><a href="#Miejsce2">dołączanie dźwięków i filmów wideo</a></li>
  <li><a href="c5.htm">zachowywanie w plikach tekstowych.</a></li>
</ul>
<p><a name="Miejsce2" id="Miejsce2" >Dzięki zachowywaniu dokumentów w zbiorach
  tekstowych, zbiory XHTML są zrozumiałe dla większości komputerów. Wynik
  końcowy jest podobny, a różnice są minimalne.</a></p>
<p>ccccccc</p>
  <!--tutaj powinna być zawartość strony przynajmniej o rozmiarze okna-->
</body>
</html>
```


(3a cd) Plik html z połączeniami do plików w bieżącym katalogu strony (pozycje listy: 1,5), do pliku na podanej stronie (pozycja 2 listy), do miejsca wskazanego w pliku w bieżącym katalogu strony (pozycja listy 3) oraz do miejsca wskazanego w tym samym pliku (pozycja listy 4 i odnośnik Miejsce2)

(3b) Plik c1.htm z połączeniem do pliku głównego c.htm
Powrót do pliku głównego.

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
    "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="pl" lang="pl">
<head>
  <meta http-equiv="content-type" content="text/html; charset= utf-8"/>
  <title> Połączenie 1</title>
</head>
<body>
  <h2>Zachowywanie w plikach tekstowych</h2>
  <p>eeeeeeeeeeeeeeee</p>
  <p><a href="c.htm">Powrót do pliku głównego</a></p>
</body>
</html>
```

(3c) Plik c3.htm

Plik c3.htm z odnośnikiem

** Dołączanie grafiki**

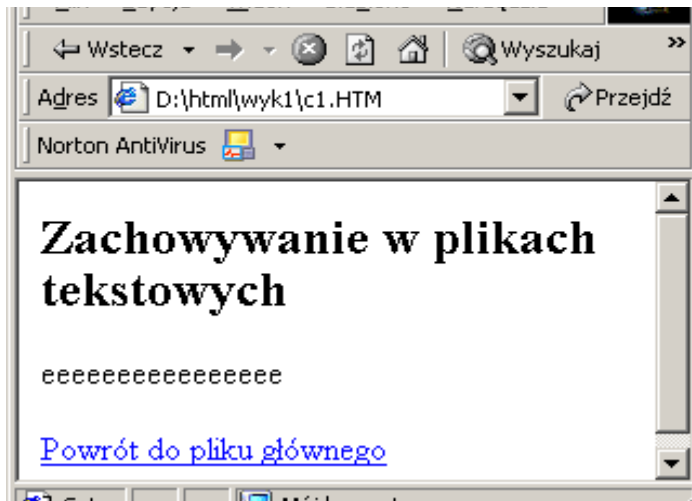
oraz z połączeniem do pliku głównego c.htm

Powrót do pliku głównego.

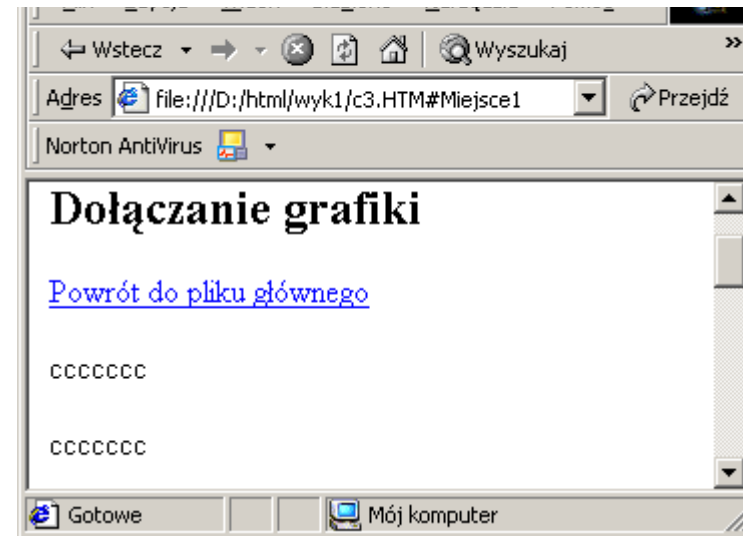
```
<!DOCTYPE html
  PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
  "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="pl" lang="pl">
<head>
<meta http-equiv="content-type" content="text/html; charset= utf-8"/>
<meta name="keywords" content="Miejsce1"/>
<title>Połączenie 3</title>
</head>
<body>
<p>ccccccc</p>
<h2><a name="Miejsce1" id ="Miejsce1" >Dołączanie grafiki</a></h2>
<p><a href="c.htm">Powrót do pliku głównego</a></p>
<p>ccccccc</p>
  <!--tutaj powinna być zawartość strony przynajmniej o rozmiarze okna-->
</body>
</html>
```

(3d) Zawartość okna przeglądarki po uruchomieniu połączeń 1, 3, 2 i 4.

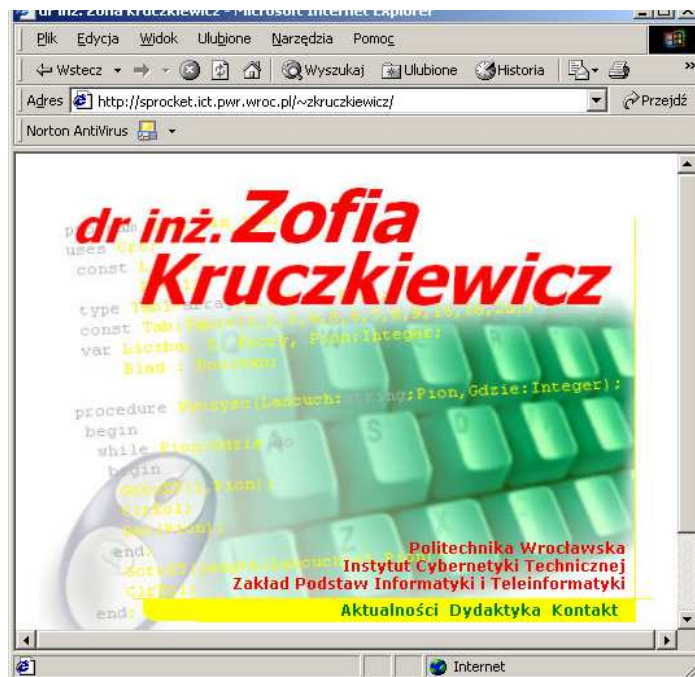
Po wyborze linku 1 „formatowanie tekstu”



Po wyborze linku 3 „dołączanie grafiki”



Po wyborze linku 2 „kreślenie linii”



Po wyborze linku 4 „dołączanie dźwięków i filmów wideo”

