Introdução às tecnologias Web - ITW

Aula 2 – Introdução à World Wide Web (WWW)

Sumário

Breve história da Internet

O protocolo http

A linguagem html – versões iniciais Cabeçalhos, parágrafos, áreas, hiperligações, tabelas. ...

Vannevar Bush, 1890-1974, foi um engenheiro, inventor e político americano, conhecido pelo seu papel político no desenvolvimento da bomba atómica e pela ideia do memex — visto como um conceito pioneiro, precursor da world wide web.

Memex (**mem**ory + ind**ex**) – máquina visionária imaginada para auxiliar a memória e guardar conhecimentos. (As We May Think, 1945).

Constatação: o conhecimento aumenta a um ritmo prodigioso, e não encontra contrapartida na evolução dos meios de armazenamento e acesso aos dados

Observando o funcionamento da mente humana, operando sempre por meio de associações, Bush imaginou e descreveu de maneira detalhada uma máquina capaz de armazenar grandes quantidades de informação, para posterior e rápida recuperação.

Tal engenho, concebido para suprir as "falhas da memória humana" através de recursos mecânicos, é considerado o precursor do conceito de <u>hipertexto</u>.

The Memex - Dissected



Douglas C. Engelbart (1925-2013) foi pioneiro na interação humano computador.

Baseou-se no conceito imaginado em "As we may think" de Vannevar Bush para desenvolver o hipertexto, computadores em rede e os precursores de interfaces gráficas.

Foi responsavel, nos finais dos anos 60, pela primeira troca de mensagens entre duas aplicações (em salas distintas), a primeira aplicação com o conceito de "windows" e ainda pela invenção de diversos dispositivos de interação com o computador – com especial destaque para mouse / rato dos computadores.







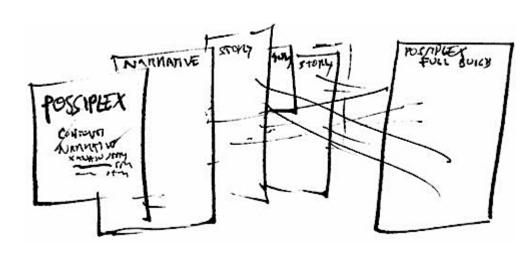


24 de janeiro de 1984

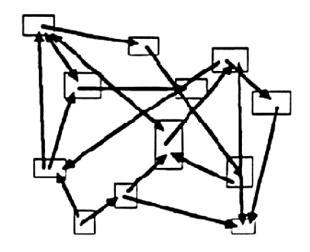
Ted Nelson, EUA, foi o "inventor" das palavras <u>hipertexto</u> e <u>hipermédia</u>.

Fundou o Projeto Xanadu em 1960 com o objetivo de criar uma rede de computadores de interface simples.

O trabalho está documentado no seu livro Computer Lib/Dream Machines, de 1974, e no livro Literary Machines, de 1981.



ORDINARY HYPERTEXT



6 de agosto de 1991 – **Tim Berners-Lee**, a trabalhar no CERN, Suíça, publica um resumo do projeto da World Wide Web no **alt.hypertext** newsgroup. Essa data marca o nascimento da Web como um serviço público da Internet;

1993 – **Marc Andreessen**, no NCSA, University of Illinois, EUA, apresenta o primeiro navegador gráfico, o Mosaic.

Depois da graduação Marc Andreessen associou-se a um anterior chefe da Silicon Graphics e iniciaram a comercialização do Mosaic que passou a chamar-se Netscape em Abril 1994.





A **Netscape Communications Corporation** foi responsável pelo desenvolvimento de:

- Secure Sockets Layer Protocol (SSL) protocolo para transmissão segura de dados, ainda hoje largamente utilizado (https, por exemplo);
- JavaScript a linguagem de programação mais utilizada na criação de programas para páginas web, do lado do cliente, ainda hoje utilizada (até nesta disciplina! ⁽¹⁾).

A Netscape foi comprada em 1999 por 4.2MM\$ pela AOL, perto do pico do que na altura de denominou como "dot-com bubble"

The Deal is Done: AOL buys
Netscape Communications

Paul Thurrott | Nov 23, 1998

The mercurial rise and fall of Netscape Communications completed Tuesday when online giant America Online announced that it was buying the browser company for \$4.21 billion. The differences between the two companies are almost astounding, despite the fact that both are involved with the Internet online experience: AOL, with 14 million users, is seen as the home for "newbies" on the Net, while people who like to think of themselves as more Web-savvy tend to download Netscape Navigator. Most unclear is what AOL hopes to gain







Antes da venda da Netscape Communications Corporation à AOL, foi disponibilizado o código fonte (*open source*) e criada a **Mozilla Organization** para permitir o desenvolvimento futuro do produto em diferentes direções.

A Mozilla Organization reescreveu todo o código do browser e denominou esse desenvolvimento **Gecko engine**;

Todas as versões do Netscape até ao seu encerramento foram baseadas no Gecko que foi posteriormente utilizado no browser **Firefox** – da Mozilla Foundation.





Evolução histórico-gráfica da marca firefox



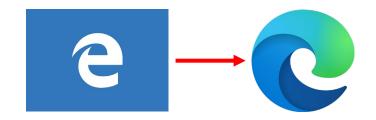


Windows **Internet Explorer** foi o browser da Microsoft desde 23 de agosto de 1995.

Em 2 de setembro de 2008 saiu a primeira versão beta do browser **Chrome** da Google. Em 11 de dezembro de 2008 foi lançada uma versão estável ao público em geral.

Em 17 de Março de 2015, a Microsoft anunciou que o Microsoft **Edge** substituirá o Internet Explorer no Windows 10.

Assim, o Internet Explorer 11 foi a última versão deste browser.



Em 8 de abril de 2019, a Microsoft anuncia que o Microsoft Edge passa a ser

baseado em Chromium.



Não confundir com ...



O <u>Chromium</u> é um projeto browser de código aberto que é também a base do browser da web Chrome.

A maior diferença entre os dois browsers é que, embora o Chrome seja baseado no Chromium, a Google também adiciona ao Chrome vários recursos proprietários, como atualizações automáticas e suporte para formatos de vídeo adicionais

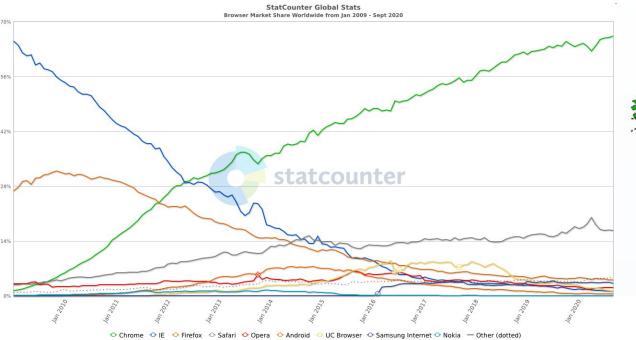






O IE foi durante muitos anos hegemónico vindo, no entanto a perder cota de mercado para o Google Chrome da Google.

O Chrome assumiu a liderança em maio de 2012, de acordo com o site StatsCounter.





12

História de Internet / WWW Os pioneiros - resumos

Tim Berners Lee

- Criador e <u>ainda</u> atual mentor da WWW
- Uniform Resource Locator (URL),
- Hypertext Transfer Protocol (HTTP),
- Hypertext Markup Language (HTML)

Marc Andreessen

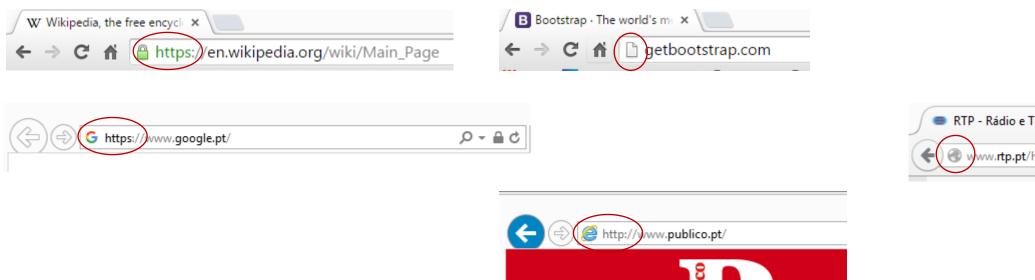
- Criador do Mosaic
- Fundador da Netscape Communications Corporation

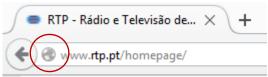
HTTP

Hypertext Transfer Protocol

Hypertext Transfer Protocol (HTTP)

Protocolo utilizado para transferir documentos de hipertexto e seus recursos de máquinas remotas.





Hypertext Transfer Protocol (HTTP) Características gerais

Protocolo da camada de aplicaçãoutilizado para transferir documentos de hipertexto e seus recursos de máquinas remotas.

Modelo de funcionamento baseado em pedido-resposta;

Cabeçalho das mensagens é texto puro (não binário);

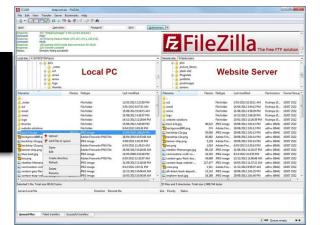
Não orientado a conexões;

Não guarda estado entre conexões distintas, isto é, cada conexão é sempre uma nova conexão para o servidor (isso não acontece, por exemplo, quando estamos a fazer uma

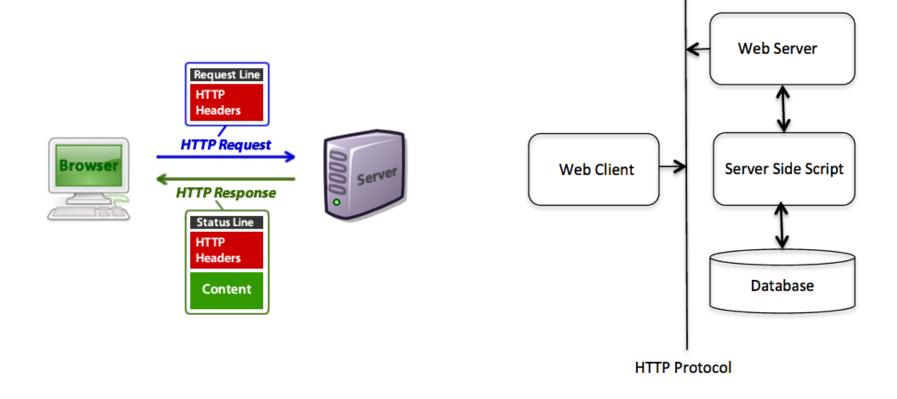
transferência de ficheiros por ftp (file transfer protocol).

* No protocolo ftp, se uma transmissão for quebrada a meio, o protocolo permite retomá-la a partir do ponto em que foi interrompida;

* No protocol http, se uma transmissão for quebrada a meio, é necessário recomeçar a transmissão desde o início.



Hypertext Transfer Protocol (HTTP) Arquitetura



Hypertext Transfer Protocol (HTTP) Sintaxe geral de uma URL

colo>://<servidor>:<porta>/<caminho>/<recurso>?dados

- Protocolo: http / https / ...
- A porta é opcional para serviços em portas default
 - para o http a porta default é a 80;
 - Para o https a porta default é a 443.
- Caminho e recurso podem ser omitidos (URLs parciais)
- As URLs podem conter dados depois do nome do recurso

Hypertext Transfer Protocol (HTTP) Exemplos de URL's

http://www.ua.pt

Neste exemplo de url parcial não se incluiu a porta default (80, para o http) nem o caminho/recurso.

O recurso por default é variável e dependente do tipo de servidor e do tipo de linguagem de programação utilizada no desenvolvimento do servidor.

Nomes típicos: index.htm / index.html / default.htm / default.html / default.asp / default.aspx / default.php / ...

http://www.ua.pt/deti/PageCourse.aspx?id=383

Neste exemplo a URL possui dados depois do recurso, neste caso, id=383. Pergunta: Que representa esta URL?

Hypertext Transfer Protocol (HTTP) Exemplos de URL's

http://www.ua.pt/deti/PageCourse.aspx?id=383&p=4&a=9

Neste exemplo há 3 parâmetros / dados depois no nome do recurso

No primeiro separador é usado o ponto de interrogação;

Nos demais separadores á utilizado o "i comercial / ampersand"

<u>Pergunta</u>: como enviar o símbolo "&" nos dados, como no exemplo seguinte, sem baralhar o servidor?:

Empresa = Dias & Dias A. Boavida

http://localhost:6067/Specie?Shell ID=265

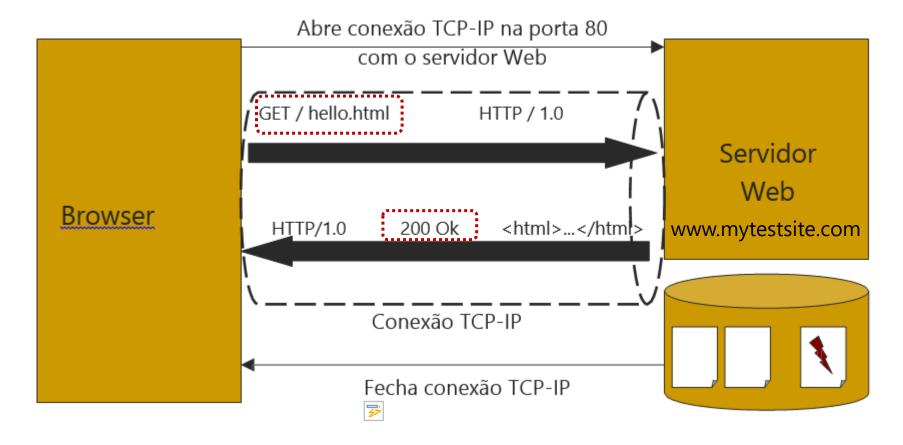
Neste exemplo o nome do servidor é o "localhost", o que significa que o servidor está no computador do próprio utilizador (*local host*).

Neste exemplo o servidor está a responder a pedidos na porta 6067.

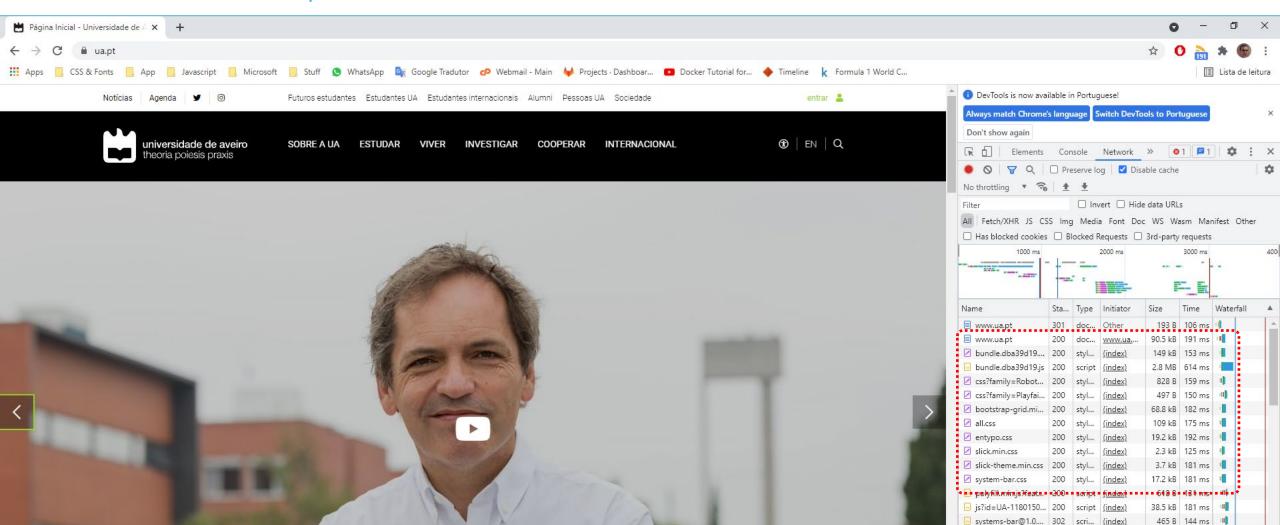


separadores

Hypertext Transfer Protocol (HTTP) Exemplo (http://www.mytestsite.com/hello.htm)



Hypertext Transfer Protocol (HTTP) Pedido e resposta a um servidor web (F12 > Network)



Hypertext Transfer Protocol (HTTP) Pedido e resposta a um servidor web

```
GET /hello.htm HTTP/1.1

User-Agent: Mozilla/4.0 (compatible; MSIE5.01; Windows NT)

Host: www.mytestsite.com

Accept-Language: en-us
Accept-Encoding: gzip, deflate

Connection: Keep-Alive
```

Pedido e resposta com sucesso

```
HTTP/1.1 200 OK
Date: Mon, 27 Jul 2009 12:28:53 GMT
Server: Apache/2.2.14 (Win32)
Last-Modified: Wed, 22 Jul 2009 19:15:56 GMT
ETag: "34aa387-d-1568eb00"
Vary: Authorization, Accept
Accept-Ranges: bytes
Content-Length: 88
Content-Type: text/html
Connection: Closed
<html>
<body>
<h1>Hello, World!</h1>
</body>
</html>
```

Hypertext Transfer Protocol (HTTP) Pedido e resposta a um servidor web

```
GET /t.htm HTTP/1.1
User-Agent: Mozilla/4.0 (compatible; MSIE5.01; Windows NT)
Host: www.mytestsite.com
Accept-Language: en-us
Accept-Encoding: gzip, deflate

HTTP/1 1 404
```

Pedido e resposta com erro

```
HTTP/1.1 404 Not Found
Date: Sun, 18 Oct 2012 10:36:20 GMT
Server: Apache/2.2.14 (Win32)
Content-Length: 230
Connection: Closed
Content-Type: text/html; charset=iso-8859-1
<!DOCTYPE HTML PUBLIC "-//IETF//DTD HTML 2.0//EN">
<html>
<head>
   <title>404 Not Found</title>
</head>
<body>
   <h1>Not Found</h1>
   The requested URL /t.html was not found on this server.
</body>
</html>
```

Connection: Keep-Alive

HTTPS

Hypertext Transfer Protocol Secure

HTTPS

Hyper Text Transfer Protocol Secure (1)

O HTTPS (Hyper Text Transfer Protocol Secure - protocolo de transferência de hipertexto seguro) é uma implementação do protocolo HTTP sobre uma camada adicional de segurança que utiliza o protocolo SSL/TLS.

Essa camada adicional permite que os dados sejam transmitidos através de uma conexão criptografada que garante a autenticidade do servidor e do cliente por meio de certificados digitais.

A porta TCP usada por norma para o protocolo HTTPS é a 443, no protocolo HTTP a porta utilizada é a 80, tal como referido previamente.

HTTPS

Hyper Text Transfer Protocol Secure (2)

A ideia principal do HTTPS é criar um canal seguro sobre uma rede insegura. Isso garante uma proteção razoável de pessoas que realizam escutas ilegais (eavesdropping attacks) e de ataques de homem-no-meio (man-in-the-middle), dado que a cifragem foi adequadamente utilizada e que o certificado do servidor é verificável e confiável.

A confiança fornecida pelo HTTPS é baseada em autoridades de certificação que vêm pré-instaladas no navegador

Isto é equivalente a dizer "Eu confio na autoridade de certificação VeriSign / Microsoft / etc. para me dizer em quem devo confiar".

Eavesdropping attacks

Antigamente, este tipo de ataque era realizado sobre as linhas telefónicas, dado origem ao que ainda hoje conhecemos como uma escuta telefónica.

Com o desenvolvimento da tecnologia, os hackers criaram novos métodos e canais para espiar as conversas digitais.

Os analisadores de protocolos podem ser usados para gravar comunicações baseadas em IP e podem converter os dados em arquivos de áudio facilmente acessíveis.

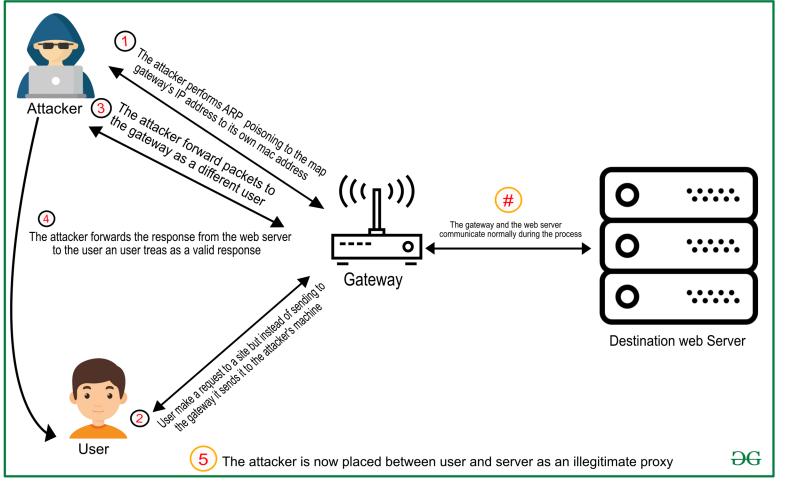
A deteção de dados é outro tipo comum de escuta eletrónica. É usado principalmente nas redes locais empresariais.

Como todas as comunicações dentro da rede são enviadas para os routers/gateways, os hackers podem aceder ao sistema e intercetar esses dados. Da mesma forma, os dados das redes sem fios podem ser facilmente manipulados se informações não seguras chegarem às portas da rede.

A forma mais comum de espionagem é conduzida usando a função de alto-falante ou microfone de um dispositivo.

Esse tipo de ataque de espionagem ocorre mais comumente ao laptop do utilizador, já que o nigorial proposition pode ser ativado discreta e remotamente por um terceiro.

Man In The Middle Attack



Princípio de funcionamento

O SSL (Secure Sockets Layer) usa um sistema de criptografia que utiliza duas chaves para criptografar os dados, uma chave pública conhecida por todos e uma chave privada conhecida apenas pelo destinatário.

O SSL é a única e eficaz maneira de obter segurança de dados em comércio eletrónico.

Quando um SSL – Certificado Digital está instalado no website, <u>um ícone de um cadeado</u> <u>aparece no navegador</u> e o endereço começa com https:// ao invés de http:// informando que os dados serão criptografados.

HTTP: Não Criptografado (sem SSL)

Falsificador de mensagens

Ladrão de dados

Espião

u visitante
(cliente)

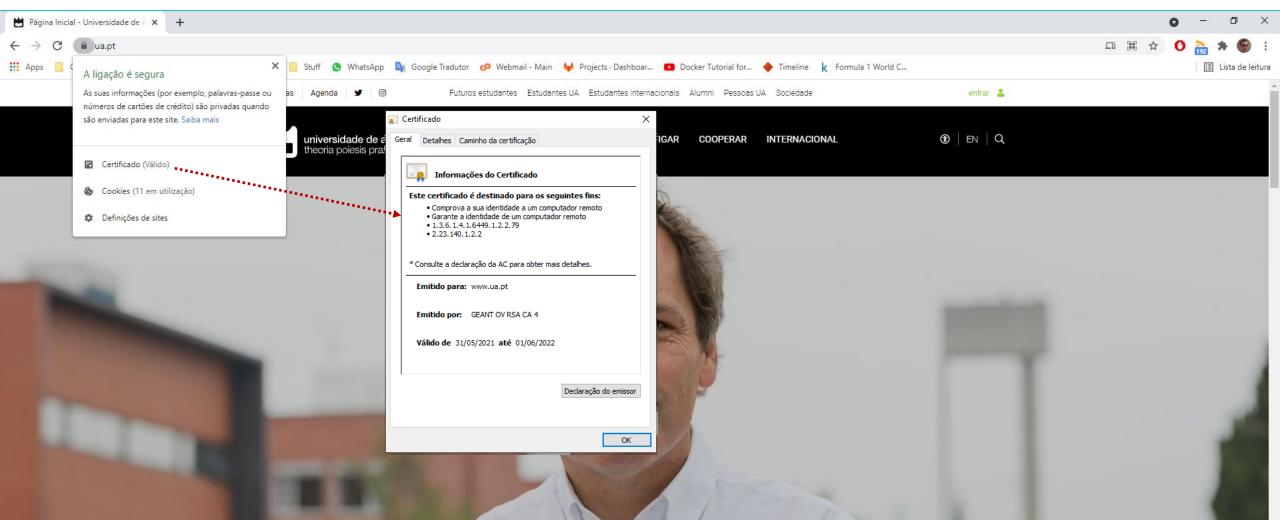
Seu site
(servidor)

HTTPS: Conexão Segura e Barata com SSL



30

HTTPS o site https://www.ua.pt



HTML

HyperText Markup Language

A linguagem html

HTML (abreviatura da expressão inglesa "HyperText Markup Language", que significa "Linguagem de Marcação de Hipertexto")

É a linguagem de marcação utilizada para produzir páginas na Web definida por Tim Berners Lee, conforme já foi visto.

Versões da linguagem html

November 24, 1995 - HTML 2.0 - IETF RFC 1866.

November 25, 1995: RFC 1867 (form-based file upload)

May 1996: RFC 1942 (tables)

August 1996: RFC 1980 (client-side image maps)

January 1997: RFC 2070 (internationalization)

January 1997 – HTML 3.2 published as a W3C Recommendation.

It was the first version developed and standardized exclusively by the W3C, as the IETF

December 1997 – HTML 4.0 published as a W3C Recommendation

Versões da linguagem html

April 1998 – HTML 4.0, reissued with minor edits without incrementing the version number.

December 1999 – HTML 4.01, published as a W3C Recommendation.

Development of the parallel, XML-based language XHTML occupied the W3C's HTML Working Group through the early and mid-2000s.

As of mid-2008 - HTML 4.01, ISO/IEC 15445:2000.

December 2012 – HTML5, candidate recommendation of the World Wide Web Consortium.

October 2014 – HTML5 was published as a W3C Recommendation

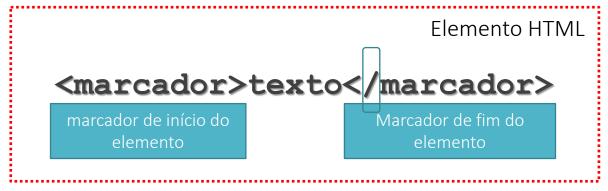
December 2017 – HTML 5.2 as a W3C Recommendation,



Características de um elemento html

Um marcador / etiqueta (*tag* em inglês) num documento html fica sempre colocada entre os símbolos "<" e ">". As etiquetas são responsáveis pela formatação da linguagem.

Nas últimas versões da linguagem, <u>um elemento é constituídos por um par</u> <u>de marcadores: um no início do elemento e outro no fim do elemento</u>.



Características de um elemento HTML

Um elemento HTML é formado por um marcador com o nome de etiqueta (tag), atributos, valores e filhos (que podem ser outros elementos ou texto).

Os atributos modificam os valores padrão dos elementos e os valores caracterizam essa mudança.

Exemplo de um elemento HTML simples (não possui filhos):

```
<hr/>
```

Exemplo de um elemento HTML com atributos:

```
<a href="http://www.ua.pt/">Universidade de Aveiro</a>
```

Exemplo de um elemento HTML com filhos e atributos:

```
<a href="http://www.ua.pt/">Universidade de Aveiro</a> é a minha Universidade.
```

Estrutura base de um documento html

```
<!doctype html>
<html lang="en">
 <head>
  <meta charset="UTF-8">
                                                                             Cabeçalho do documento
  <meta name="Author" content="">
                                                                             [Este conteúdo não é representado]
  <meta name="Keywords" content="">
  <meta name="Description" content="">
 <title>Document</title>
 </head>
                                                                     Conteúdo do documento
 <body>
                                                                     [Este conteúdo é representado]
 </body>
</html>
```

Elementos base de um documento html

Os elementos básicos de um documento html são:

<html> – define o início de um documento html e indica ao navegador que todo conteúdo posterior a este marcador deve ser tratado como uma série de marcadores html.

Neste marcador pode ter como atributo a língua em que o documentos está escrito. Essa informação é importante para os motores de pesquisa / indexação.

<head> – define o cabeçalho do documento e possui informações sobre o documento que vai ser representado;

<body> – define o conteúdo principal do documento e é a paret que é exibida no navegador. No marcador do corpo podem-se definir atributos comuns a toda a página, como cor de fundo, margens, e outras formatações.

Elementos do cabeçalho de um documento html

Dentro do cabeçalho podemos ter os seguintes elementos:

```
<title> – define o título da página; é exibido na barra de título do browser;

<style type="text/css"> – define a formatação das etiquetas em CSS;

<script type="text/javascript"> – define a programação de certas funções da página escritas na linguagem JavaScript;
```

- define ligações da página com outros arquivos, como por exemplo, feeds, CSS, scripts, tipos de letra, etc.

<meta> – define propriedades da página, como, por exemplo, a codificação dos caracteres (UTF-8, ISO-8859-1, ...), a descrição da página (autor, keywords,...)

Etiquetas do corpo de um documento html

```
<h1>, <h2>, ... <h6> – define os cabeçalhos e títulos no documento em diversos tamanhos;
```

... – define um novo parágrafo;

 - impõe uma quebra de linha num texto;

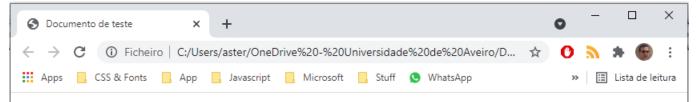
..., <i>...</i>, <u>...</u> e <s>...</s> – determina diversas formatações de letra: negrito, itálico, sublinhado e riscado, respectivamente.

<a>... – cria uma hiperligação para um outro local, seja uma página, um e-mail ou outro serviço.

Exemplo Código fonte

```
<!doctype html>
<html lang="en">
<head>
  <meta charset="UTF-8">
  <meta name="Author" content="Joaquim Sousa Pinto">
  <meta name="Keywords" content="test page; html">
  <meta name="Description" content="">
  <title>Documento de teste</title>
</head>
 <body>
    <h1>Vidit Homerus probari fabulam non posse</h1>
    <h2>I cantiunculis tantus irretitus vir teneretur;</h2>
    Lorem ipsum dolor sit amet, consectetur adipiscing elit. Quantum Aristoxeni ingenium consumptum
videmus in musicis? Quid ait Aristoteles reliquique Platonis alumni? Hic nihil fuit, quod quaereremus.
    <b>Eademne, quae restincta siti?</b> Quae diligentissime contra Aristonem dicuntur a Chryippo.
Quippe: habes enim a rhetoribus; Negat enim summo bono afferre incrementum diem.
    Sed quid sentiat, non videtis. Minime vero, inquit ille, consentit. Beatus sibi videtur esse
moriens. Duo Reges: constructio interrete. <i>Nihil illinc huc pervenit.</i><br/>
    Sed haec nihil sane ad rem;
</body>
</html>
©2014-21, JOAQUIM SOUSA PINTO
```

Exemplo Representação no navegador



Vidit Homerus probari fabulam non posse

I cantiunculis tantus irretitus vir teneretur;

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Quantum Aristoxeni ingenium consumptum videmus in musicis? Quid ait Aristoteles reliquique Platonis alumni? Hic nihil fuit, quod quaereremus.

Eademne, quae restincta siti? Quae diligentissime contra Aristonem dicuntur a Chryippo. Quippe: habes enim a rhetoribus; Negat enim summo bono afferre incrementum diem.

Sed quid sentiat, non videtis. Minime vero, inquit ille, consentit. Beatus sibi videtur esse moriens. Duo Reges: constructio interrete. Nihil illinc huc pervenit.

Sed haec nihil sane ad rem;

Exemplos de hiperligações

Ligação para <u>outro site</u> internet:

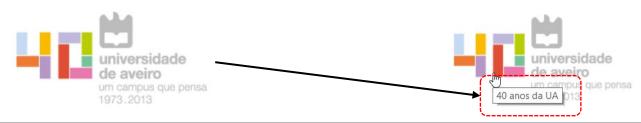
Ligação à Wikipedia!

Ligação para outra página do mesmo site:

Estando <u>dentro do site</u> <u>www.ua.pt</u> e pretendendo aceder à página do curso da LEI ...

Licenciatura em Engenharia Informática

Exemplos da utilização de hiperligações com imagens



Parâmetros da hiperligação:

href = destino

Parâmetros da imagem:

- src= localização da imagem
- alt = texto apresentado nos browsers de texto ou quando as imagens não são apesentadas
- title = texto que é mostrado quando o rato passa sobre a imagem (pode ser também utilizado o atributo aria-label)
- width = largura da imagem
- height = altura da imagem

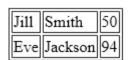
Etiquetas do corpo de um documento html

```
... - cria uma tabela
As linhas são criadas com o marcador ...
As células são criadas com o marcador ...

Os cabeçalhos das colunas são criados com o marcador ...

A legenda da tabela é inserda com o marcador <caption>...</caption>
```

Exemplo de uma tabela () com duas linhas () e 3 colunas ()



Adaptado de : http://www.w3schools.com/html/html tables.asp

Exemplo de tabela com 2 <u>linhas</u> () sendo a primeira com <u>cabeçalhos</u> () e a segunda com <u>dados</u> ()

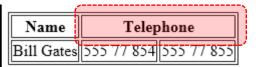
```
Firstname Lastname Points
Eve Jackson 94
```

Adaptado de: http://www.w3schools.com/html/html tables.asp

Exemplo de uma tabela com uma <u>célula colapsada na horizontal</u> (colspan).

Note que a 1^a linha possui apenas 2 colunas enquanto a 2^a linha possui 3 colunas. Na primeira linha, as 2^a e 3^a colunas foram unidas numa única (atributo colspan="2").

```
        \lambda name
        \lambda name
```



Fonte: http://www.w3schools.com/html/html tables.asp

Exemplo de uma tabela com uma <u>célula colapsada na vertical</u> (rowspan).

Note que a 3ª linha possui somente uma coluna enquanto as duas anteriores possuem duas.

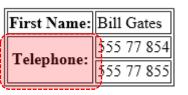
Na 1^a coluna da 2^a linha o atributo rowspan="2" une esta célula com a seguinte.

```
        First Name:
        Bill Gates

                rowspan="2">Telephone:
                555 77 854

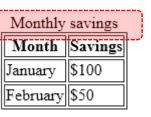
                      555 77 855

                      555 77 855
```



Adaptado de: http://www.w3schools.com/html/html tables.asp

O atributo <caption> introduz uma legenda na tabela (caption)



Adaptado de: http://www.w3schools.com/html/html tables.asp

Etiquetas do corpo de um documento html Imagens

 - Insere uma imagem.

```
<img src="https://www.w3schools.com/images/lamp.jpg" alt="Lamp" width="32" height="32" />
```

```
<a><img/></a>
```