

# HTML (I)

Linguagens de Anotação de Documentos

# *HyperText Markup Language*

- Linguagem de anotação para a criação de *documentos* na *web* (*páginas*)
- Tal como no LaTeX, as anotações são estruturais
- O “*esquema*” da linguagem é diferente, mas os conceitos são semelhantes
- Neste caso, as *anotações* são utilizadas pelos *browsers* para apresentar documentos

# WWW

- **Internet:** sistema de computadores ligados em rede que comunicam através do mesmo protocolo (TCP/IP)
- **World Wide Web (WWW):** um modelo de transmissão de dados através da *internet*, disponibilizando recursos ligados entre si
  - um sistema para *identificar* unicamente recursos (URL),
  - um *protocolo* de comunicação para aceder aos recursos (HTTP)
  - uma *linguagem* de anotação para descrever os recursos (HTML)

# WWW

- Proposta em 1989 por Tim Berners-Lee no CERN
- O objetivo era explorar a *internet* para partilhar facilmente documentos com *hiper-ligações (hiper-texto)*
- Desde cedo se decidiu que seria livre e não proprietária, o que contribuiu para a adoção universal

## World Wide Web

The WorldWideWeb (W3) is a wide-area [hypermedia](#) information retrieval initiative aiming to g  
Everything there is online about W3 is linked directly or indirectly to this document, including a [Policy](#) , November's [W3 news](#) , [Frequently Asked Questions](#) .

[What's out there?](#)

Pointers to the world's online information, [subjects](#) , [W3 servers](#) , etc.

[Help](#)

on the browser you are using

[Software Products](#)

A list of W3 project components and their current state. (e.g. [Line Mode](#) ,X11 [Viola](#) , [NeX](#)

[Technical](#)

Details of protocols, formats, program internals etc

[Bibliography](#)

Paper documentation on W3 and references.

[People](#)

A list of some people involved in the project.

[History](#)

A summary of the history of the project.

[How can I help ?](#)

If you would like to support the web..

[Getting code](#)

Getting the code by [anonymous FTP](#) , etc.

# HTML

- Linguagem de anotações que dá instruções aos *browsers* de como apresentar texto, imagens, ...
- Ao contrário do LaTeX, as páginas HTML são difundidas em texto puro, sendo interpretadas localmente pelos *browsers*
- A linguagem tem que estar bem definida e padronizada
- Atualmente gerido pelo *World Wide Web Consortium* (W3C)
- A versão mais recente (HTML5) foi publicada em 2014

# HTML

- Apesar de ser a peça fundamental, é atualmente apoiada por muitas outras tecnologias
- **CSS:** folhas de estilo para definir o aspecto de forma independente da estrutura/conteúdo
- **JavaScript:** linguagem de programação para escrever páginas com conteúdo dinâmico
- Há também ferramentas que automatizam certas partes do desenvolvimento *web*, ocultando o código HTML

# Elementos HTML

- Cada anotação define um *elemento* da página
- As anotações em HTML têm a forma

`<nome_marca> conteúdo </nome_marca>`

- Não são sensíveis à capitalização (mas é recomendado que se usem minúsculas)
- Marcas sem conteúdo podem não ser fechadas (mas é recomendado que sejam escritas como `<nome_marca />`)

# Elementos HTML

- Elementos podem ter *argumentos* adicionais, que configuram o comportamento das anotações
- Passados na marca de abertura:

`<nome_marca nome_arg="valor"> conteúdo </nome_marca>`

- Não são conteúdo, não são apresentados



# Elementos Básicos

- A primeira anotação (não HTML) define o esquema que vai ser usado no documento (`html` é o HTML 5)

```
<!DOCTYPE html>
```

- De seguida dá-se início ao documento com `<html>`, a sua raiz

```
<html>
```

```
outros elementos aninhados
```

```
</html>
```

# Elementos Básicos

- No *cabeçalho* `<head>` define-se meta-informação e estilos
- Por exemplo, o “*título*” da página (apresentado, e.g., no *browser* e em ferramentas de pesquisa)

```
<title> título </title>
```

# Elementos Básicos

- O *corpo* do documento é definido com `<body>`
- Representa o conteúdo que é efetivamente apresentado na página

```
<body>  
  todo o conteúdo  
</body>
```

# Elementos Básicos

- Primeiro documento apresentável:

```
<!DOCTYPE html>
<html>
<head>
<title>Primeira página</title>
</head>
<body>
Olá mundo!
</body>
</html>
```

# Elementos Básicos

- Títulos de secções (*headings*)
- Vários níveis de importância, de `<h1>` a `<h6>`:

`<h1>` Introdução `</h1>`

`<h2>` Objetivos `</h2>`

# Elementos Básicos

- Outras estruturas básicas
- *Parágrafos*: `<p> ... </p>`
- Mudanças de *linha* `<br>` (não têm conteúdo)
- *Separadores* horizontais `<hr>` (não têm conteúdo)

# Elementos Básicos

- *Atributos básicos*
- *Tooltips*

```
<h1 title="O título"> ... </h1>
```

- Dimensões: *width, height*

```
<hr width="50%">
```

- Texto *alternativo* (quando objeto indisponível ou acessibilidade)

```

```

# Elementos Básicos

- *Formatação básica*
- Marcas *lógicas* (e.g., `em`, `strong`) vs. *físicas* (e.g., `i`, `b`)
- Apesar de poderem ter o mesmo aspecto, têm significados distintos
- As lógicas delegam a formatação para as folhas de estilo; as físicas forçam a formatação
- E.g., as partes texto enfatizadas vs. as partes do texto efetivamente em itálico (e.g., palavras inglesas)



# Elementos Básicos

- *Símbolos especiais* reservados para as marcas HTML
  - `&lt` (<), `&gt` (>), `&amp` (&), ...
- Espaços em branco não colapsáveis
  - `&nbsp`

# Elementos Básicos

- *Hiper-ligações* entre elementos

```
<a href="url"> texto </a>
```

- Tanto para páginas locais, como para páginas externas
- Outros atributos, como o `target` que define como a ligação é seguida
  - E.g., `_blank` abre uma nova página

# Elementos Básicos

- *Referências cruzadas*
- O atributo `id` define etiquetas únicas nos elementos

```
<h1 id="intro"> Introdução </h1>
```

- Depois são referenciados por hiper-ligações marcadas por #

```
<a href="#intro"> Ver introdução </a>
```

# Elementos Básicos

- *Comentários* (ignorados pelos interpretadores)
- Elementos definidos como

```
<!-- comentários -->
```

# Questões Técnicas

- Como todas as linguagens de anotação, qualquer *editor de texto* pode ser usado, mas funcionalidades avançadas aumentam a produtividade
  - Extensão *.html* (às vezes *.htm* por motivos históricos)
- Por outro lado, não há processo de compilação: os *browsers interpretam* diretamente os ficheiros HTML

# *Take-home Lesson*

- Apesar das imensas tecnologias atuais, o HTML continua no centro da *web*
- É uma linguagem de anotação: texto puro anotado, que depois é interpretado pelos *browsers*
- Atualmente as funcionalidades básicas de formatação devem ser evitadas para separar as preocupações

# Tutorial

<https://www.w3schools.com/html>