HTML
the language for the web

Our goal is to teach you the basics so that you can pick up the rest on your own.

What is HTML?

Hypertext Markup Language

a simple data format used to create hypertext documents that are portable from one platform to another.

Hypertext

Hypertext is text displayed on a computer display or other electronic devices with references (**hyperlinks**) to other text that the reader can immediately access.

Hypertext documents are interconnected by hyperlinks, which are **typically activated by user actions,** such as a mouse click, keypress set or by touching the screen.

Introduzione alla programmazione web – M
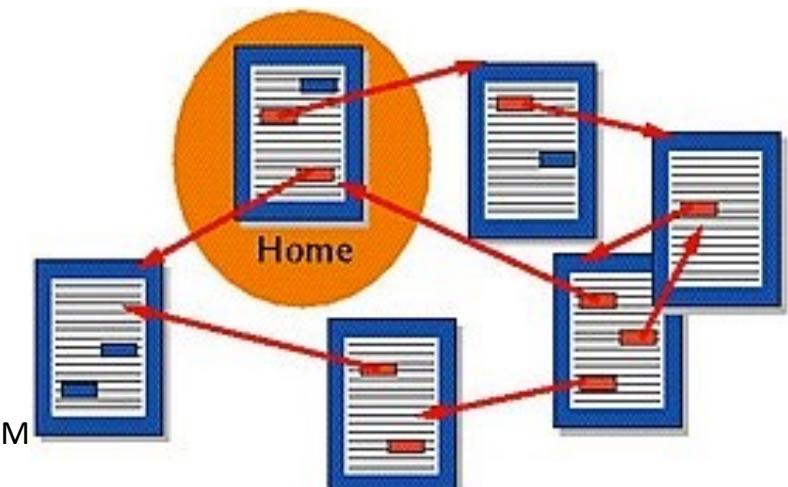
What is HTML?

- Defined in RFC 1866
- Logical markup language
- maximally portable
- **Graceful degradation of presentation**

"A philosophical rule was that HTML should convey the structure of a hypertext document, but not the details of its presentation."

Since then, HTML has **evolved into a presentation language**.

HTML standards are developed under the authority of the World Wide Web Consortium (W3C), headed by Tim Lee

http://www.w3c.org

Media - audio   (HTML 5)

\<audio src=" teenage_dream.mp3"\>\</audio\>

\<audio controls\>
\<source src="teenage_dream.mp3"/\>
\<source src="teenage_dream.ogg"/\>
\</audio\>

Provides a download link for non-supporting browsers:

\<audio src="teenage_dream.ogg" autoplay controls loop\>

\<a href=" teenage_dream.ogg"\>download\</a\>

\</audio\>   *"Graceful degradation of presentation"*

Introduzione alla programmazione web – Marco Ronchetti 2020 – Università di Trento

Versions

version	Date	Notes
HTML	1991	
HTML 2.0	1995	First official version
HTML 3.2	1997	
HTML 4.0 (4.01)	1999 (2001)	
XHTML 1.0	2000 (2001)	XML version of HTML 4
DHTML		HTML 4 +JS + CSS + DOM
XHTML 2.0		Abandoned 2009
HTML 5 (5.2)	2014 (2017)	(2008 first draft, 2021 last version)
XHTML 5.1		XML version of HTML 5

Initially maintained by W3C

Since 2012 also defined by WHATWG (Web Hypertext Application Technology Working Group)

Since 2019 maintained only by WHATWG

# Most desktop browsers still miss about 10% of features!

Maximum points: 555

	Chrome	Opera	Firefox	Edge	Safari
Upcoming	67 ▶ 528		60 ▶ 497	18 ▶ 496	11.2 ▶ 477
Current	66 ▶ 528	45 ▶ 518	59 ▶ 491	17 ▶ 492	11.1 ▶ 471
Older	65 ▶ 528	37 ▶ 489	58 ▶ 486	16 ▶ 476	11 ▶ 452
	64 ▶ 528	30 ▶ 479	57 ▶ 486	15 ▶ 473	10.1 ▶ 406
	63 ▶ 528	12.10 ▶ 309	56 ▶ 478	14 ▶ 460	10.0 ▶ 383
	62 ▶ 528		55 ▶ 478	13 ▶ 433	9.1 ▶ 370
	61 ▶ 526		54 ▶ 474	12 ▶ 377	9.0 ▶ 360
	60 ▶ 523		53 ▶ 474	Internet Explorer	8.0 ▶ 354
				11 ▶ 312	

Mobile and Tablets are even worse...

Introduzione alla programmazione web – Marco Ronchetti 2020 – Università di Trento

HTML5 over the years

HTML5test.com score over the years

Introduzione alla programmazione web – Marco Ronchetti 2020 – Università di Trento
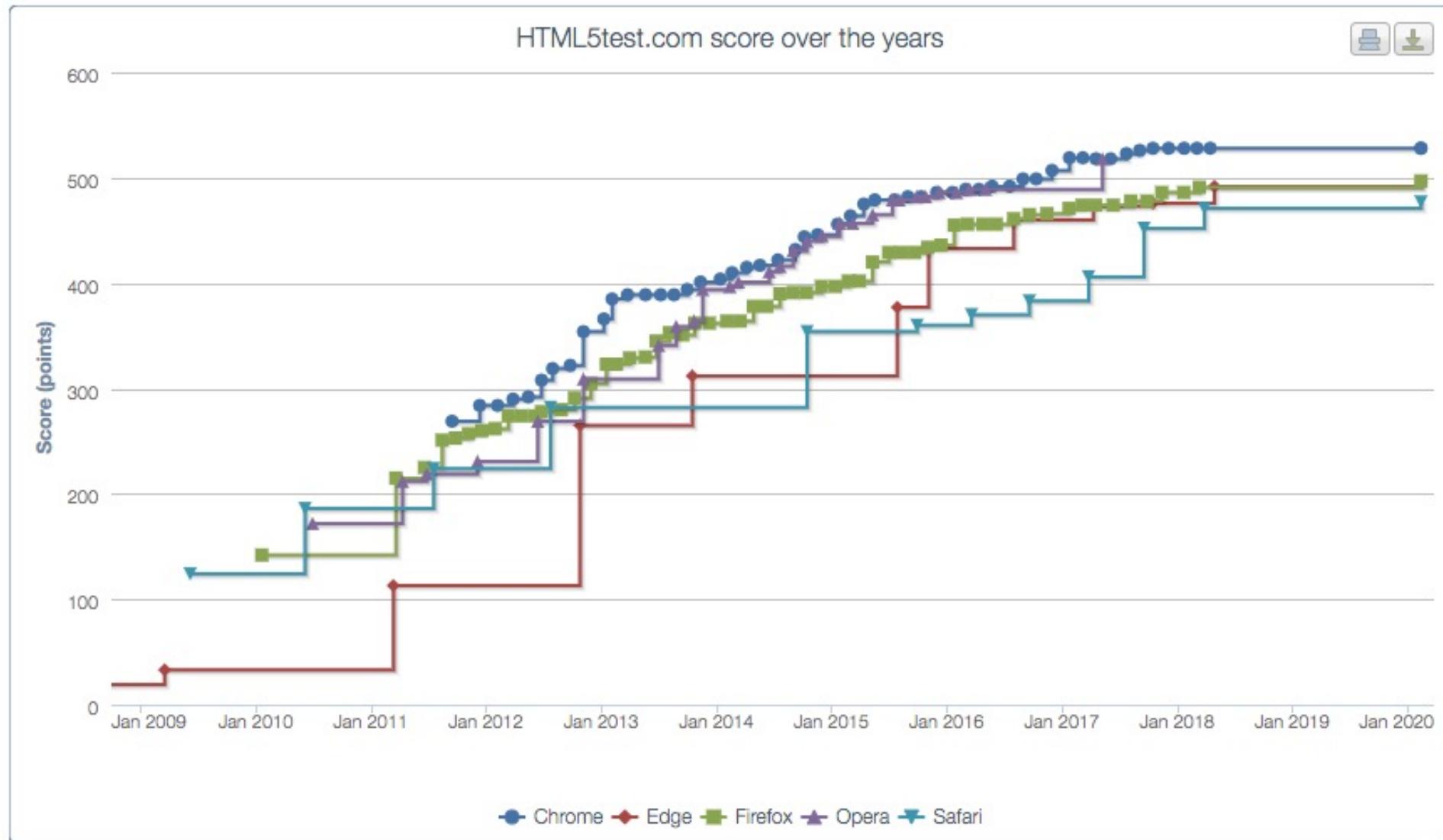
Content/Aspect separation
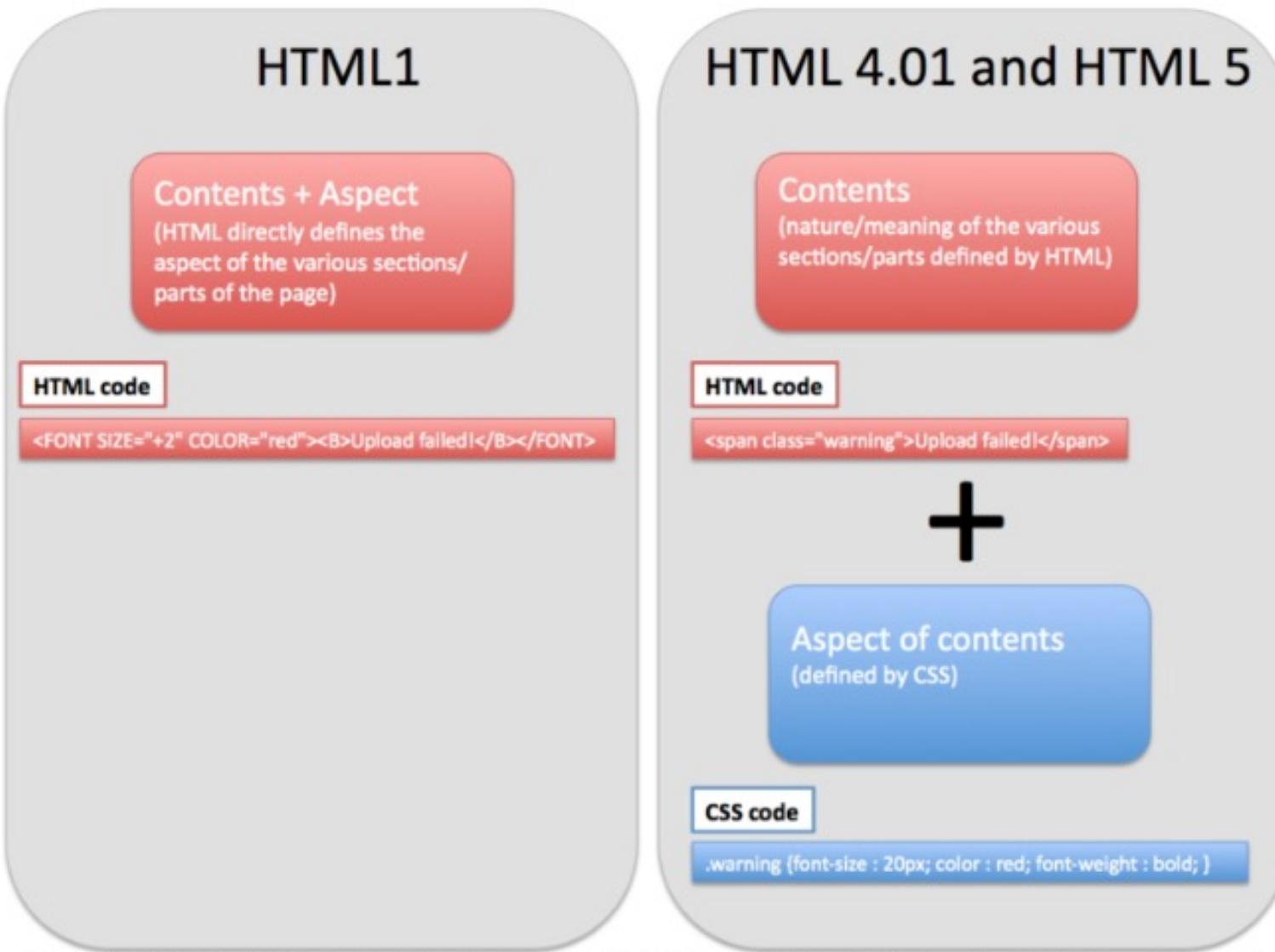
HTML1

Contents + Aspect
(HTML directly defines the aspect of the various sections/ parts of the page)

HTML code

`<FONT SIZE="+2" COLOR="red"><B>Upload failed!</B></FONT>`

HTML 4.01 and HTML 5

Contents
(nature/meaning of the various sections/parts defined by HTML)

HTML code

`<span class="warning">Upload failed!</span>`

+

Aspect of contents
(defined by CSS)

CSS code

`.warning {font-size : 20px; color : red; font-weight : bold; }`

Da http://www.cellbiol.com/bioinformatics_web_development

Introduzione alla programmazione web – Marco Ronchetti 2020 – Università di Trento

Structure of an HTML doc

```
<!DOCTYPE html>
<html>
  <head>
    <meta charset="utf-8">
    <title>My test page</title>
  </head>
  <body>
    <p>This is my page</p>
  </body>
</html>
```


Introduzione a
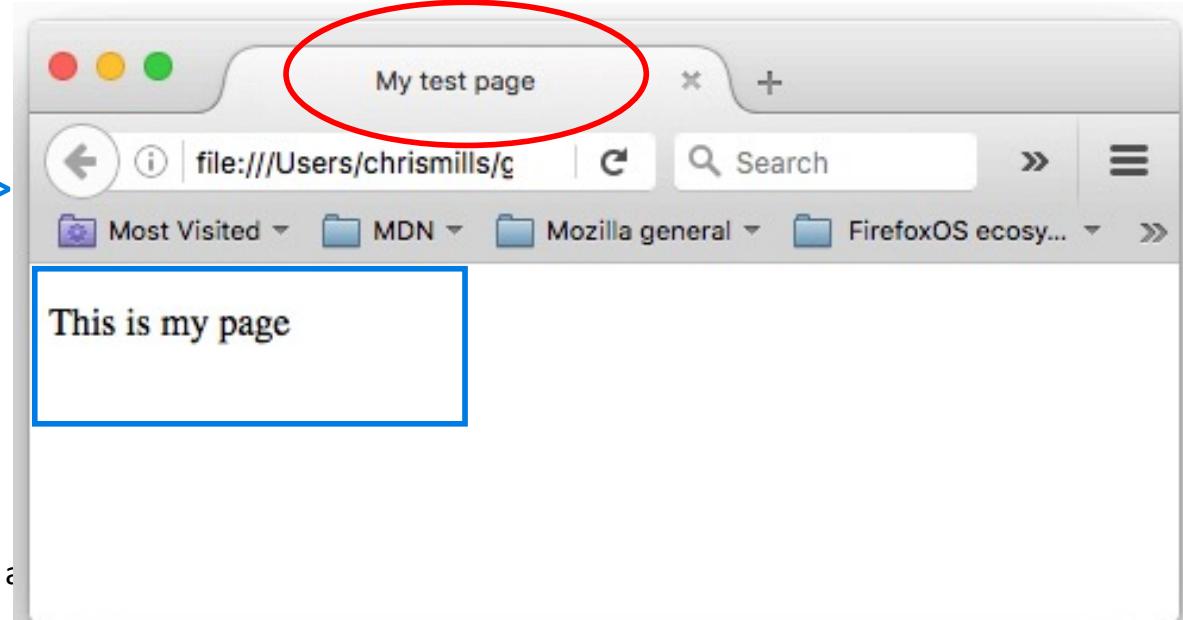
Structure of an HTML doc

- HEAD: contains information about the document (metadata):

    - Title of the page (which appears at the top of the browser window)

    - Meta tags: used to describe the content (used by Search engines)

    - Statements or references to JavaScript and StyleSheets code

- BODY: Contains the actual content of the document

    - This is the part that will be displayed in the browser window

DocType declaration

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

Today (simplified in HTML 5.0):

<!DOCTYPE html>

The HTML DTDs!

https://www.w3.org/TR/REC-html40/sgml/dtd.html

https://www.w3.org/TR/html401/sgml/dtd.html

https://www.w3.org/TR/xhtml1/dtds.html#a_dtd_XHTML-1.0-Strict

There is no HTML5 DTD!

Structure of an HTML doc

```
<!DOCTYPE html>
<html>
  <head>
    <meta charset="utf-8">
    <title>My test page</title>
  </head>
  <body>
<p>This is


        my page</p>

  </body>
</html>
```

Indentation is irrelevant!
Multiple spaces, line breaks, tabs
Are mapped into a single space.

Introduzione a
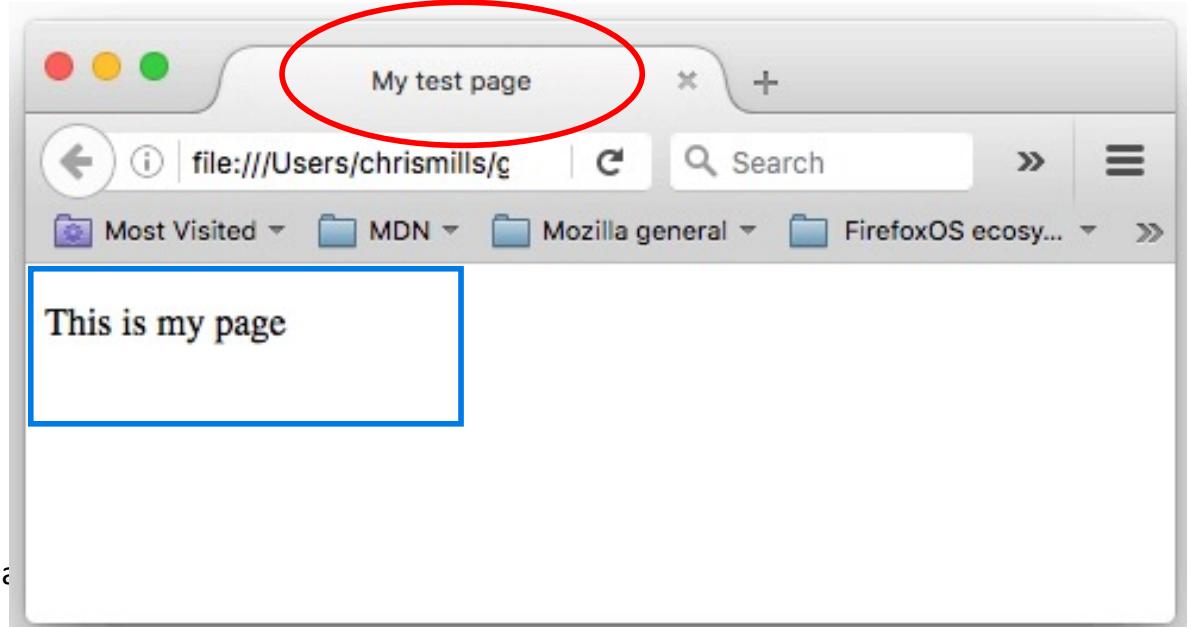
Structure of an HTML doc

```
<!DOCTYPE html>
<html>
  <HEAD>
    <meta charset="utf-8">
    <title>My test page</title>
  </head>
  <BOdy>
<p>This is

      my page</p>

    </body>
</html>
```

Uppercase and lower case do not matter in tags. (It does in XHTML!) Using lowercase is recommended.

Introduzione a
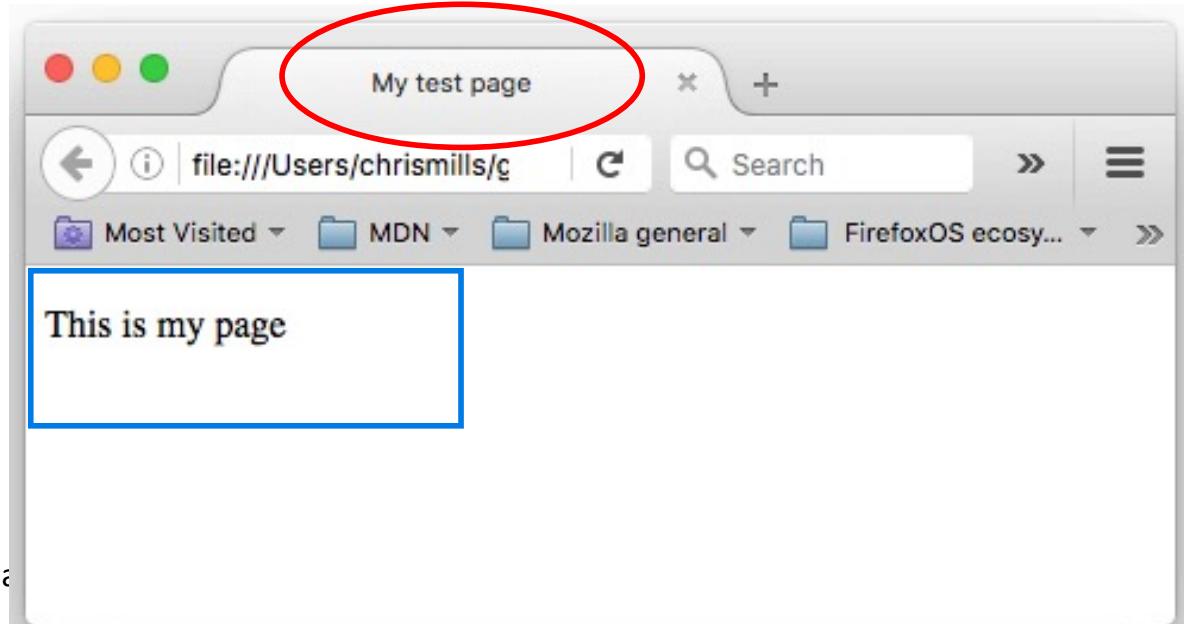
Structure of an HTML doc

```html
<!DOCTYPE html>
<html>
  <head>
    <meta charset="utf-8">
    <title>My test page</title>
  </head>
  <body>
    <p>This is my page</p>
  </body>
</html>
```

HTML also works without a server!

Introduzione a
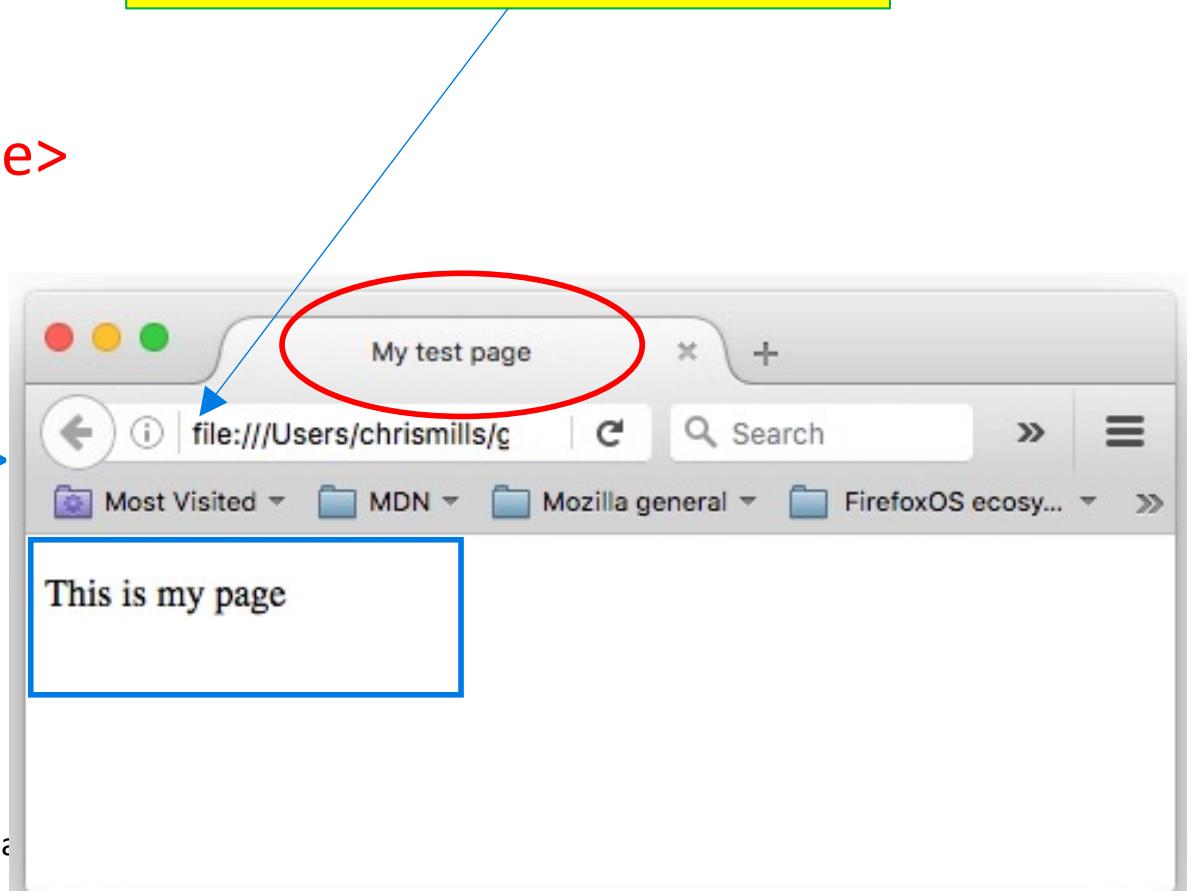
View source

W3school resources

https://www.w3schools.com/html/
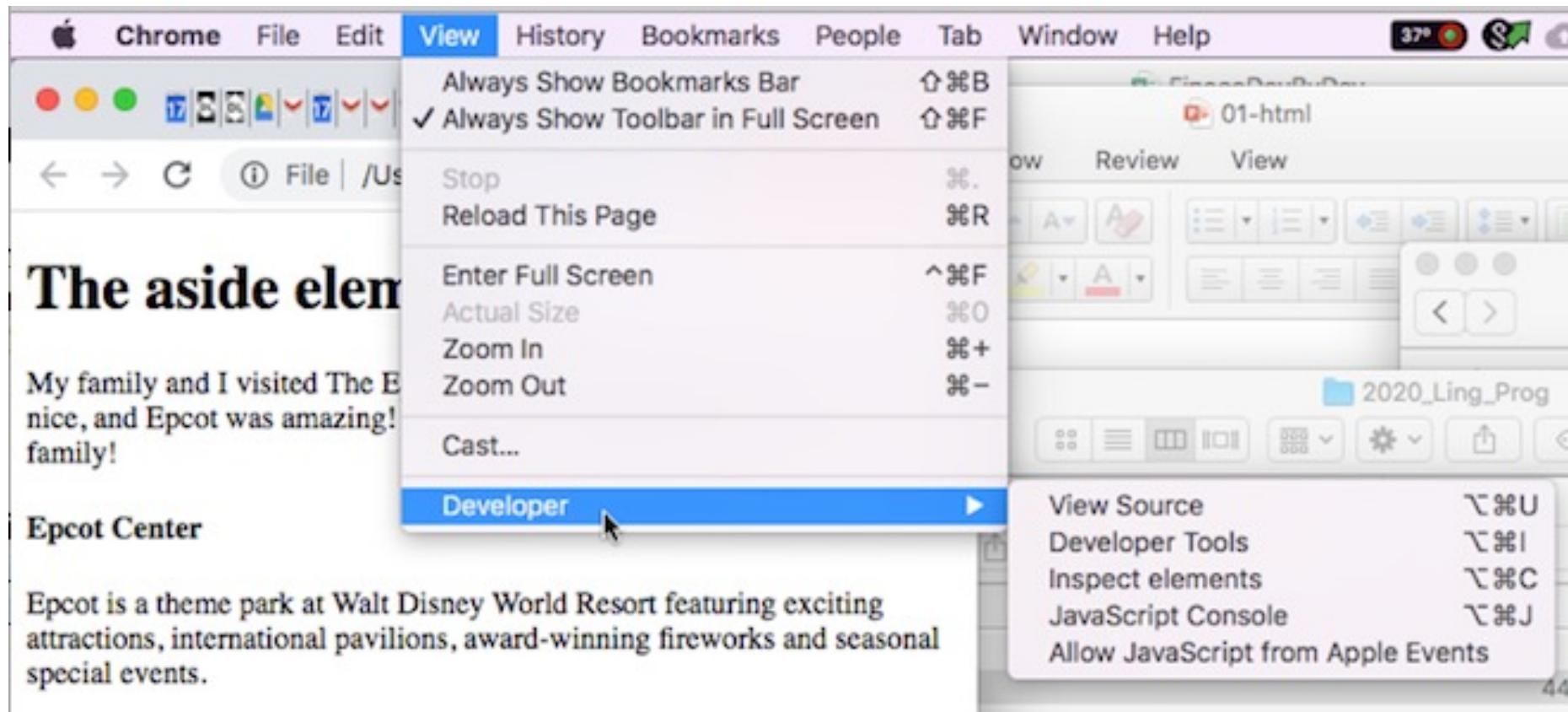
Example

```
<!DOCTYPE html>
<html>
<head>
<title>Page Title</title>
</head>
<body>

<h1>This is a Heading</h1>
<p>This is a paragraph.</p>

</body>
</html>
```

Try it Yourself »
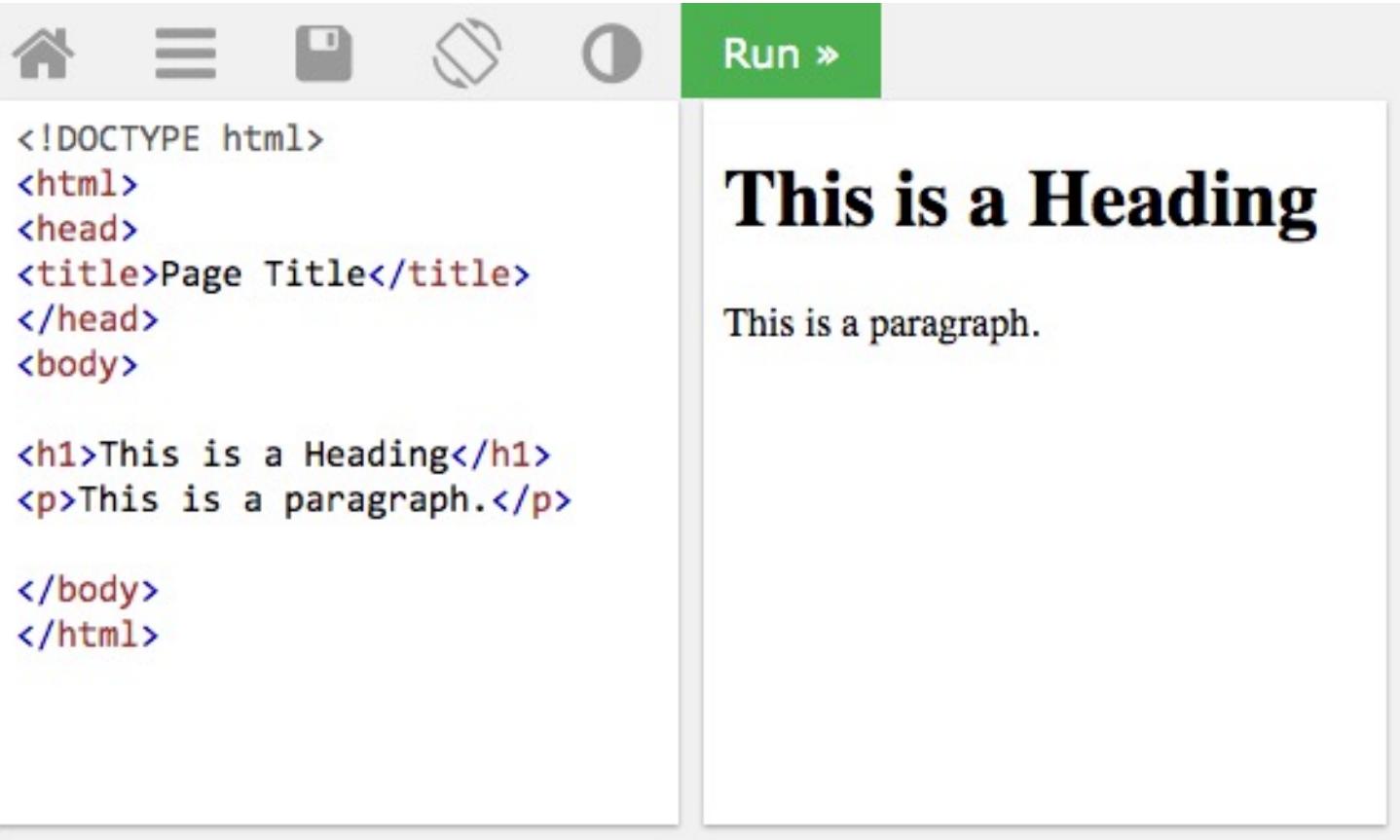
Run »

```
<!DOCTYPE html>
<html>
<head>
<title>Page Title</title>
</head>
<body>

<h1>This is a Heading</h1>
<p>This is a paragraph.</p>

</body>
</html>
```

This is a Heading

This is a paragraph.

HTML Tags

All HTML tags are made up of a tag name, sometimes followed by an optional list of attributes.

Nothing within the brackets will be displayed by the browser (unless the HTML is incorrectly written and the browser interprets the tags as part of the content).

Tags

- A tag is a keyword enclosed by angle brackets (Example: <B>)

- The opening and closing tags use the same command except the closing tag contains and additional forward slash /

- Example: `<B> Warning </B>`
  causes the word 'Warning' to appear in bold face on a Web page

- Learning HTML essentially means learning a vocabulary of tags.

Tags – void elements

- There are opening and closing tags for many but not all tags; The affected text is between the two tags.

- See: https://html.spec.whatwg.org/multipage/syntax.html#void-elements

-
 for instance does not strictly require a closing tag – *in principle* it would be better to use
</BR> or the self-closing tag

Is the Ending Slash Optional?

HTML5: the slash is optional.

HTML4: the slash is technically invalid.
          However, it's accepted by W3C's HTML validator.

XHTML: The slash is REQUIRED.

Nested Tags

- Whenever you have HTML tags within other HTML tags, you *should* close the most inner tag first

- Example:

<I> <B> The winner is: Sofia Goggia! </B>Congratulations! </I>

Bold italic   *italic*

- However, this is legal (NOT in XHTML):

<B> The winner is: <I> Sofia Goggia! </B>Congratulations! </I>

Bold   *Bold italic*   *italic*

Attributes

Name   Value

<body bgcolor="green" text="#000000" link='blue'>

Double quotes   Single quotes

Attributes: Name-value pairs which refine or extend tag's functions

- You can add multiple attributes within a single tag.
- Some attributes only have name (no associated value).
- Values are limited to 1024 characters in length.
- Attributes values are within single or double quotes.

Basic formatting tags

- Headings **<h1>**…**<h6>**
- Emphasis **<em>,<strong>,<b>,<i>**
- Subscript, superscript **<sub><sup>**

```
<!DOCTYPE html>
<html>
<body>
<h1>this is H1</h1>
<h2>this is H2</h2>
<h3>this is H3</h3>
<h4>this is H4</h4>
this is normal text<br>
<b>this is bold text</b><br>
<em>this is emphasized text</em><br>
<strong>this is strong text</strong>
<br>
<cite>this is a citation</cite><br>
</body>
</html>
```

this is H1

this is H2

this is H3

this is H4

this is normal text
this is bold text
this is emphasized text
this is strong text
this is a citation

Basic formatting tags

■ Lists <**ol**>, <**ul**>, <**li**>

```
<ol>
    <li>…</li>
    <li>…</li>
</ol>
```

OL - Ordered **L**ist

1. First item
2. Second item
3. Third item

UL - Unodered **L**ist

• First item
• Second item
• Third item

Basic formatting: blocks

- Line breaks:

- Paragraphs and blocks: <p>, <div>, <span>

- **P** and **DIV** have *semantic* difference
    - **<div>** describes a container of data (*with graphic separation*: it has a line-break before and after it)
    - **<span>** is a logical, in-line grouping *without graphic evidence*. It's mostly used for a small chunk of HTML inside a line

    - **<p>** describes a paragraph of content (with graphic separation). *It cannot contain block-level elements (P, DIV) but can contain SPAN*).

Don't mark up your document based on how it should look - mark it up based on what it is.

Basic formatting: blocks

- This works (because HTML is tolerant) but is conceptually WRONG.

```
<!DOCTYPE html>
<html>
<body>
<p>
This is a paragraph
<P> this is within an inner P</P>
This is the rest of the paragrah
</p>
</body>
</html>
```

This is a paragraph

this is within an inner P

This is the rest of the paragrah

Comments

Comments <!-- This is a comment -->

Images

`<img src="url" >`

Figure and Figure caption (HTML 5)

Cool Ninja Guy

```
<figure>
  <img src="ninja_guy.jpg" alt="A Standing Ninja Guy.">
  <figcaption>Cool Ninja Guy</figcaption>
</figure>
```

How are images indexed by search engines?

Introduzione alla programmazione web – Marco Ronchetti 2020 – Università di Trento

File paths

<img src="http://machine.domain/images/picture.jpg">	Absolute Path
<img src="picture.jpg">	picture.jpg is located in the same server, same folder as the current page
<img src="images/picture.jpg">	picture.jpg is located in the images folder in the current folder
<img src="/images/picture.jpg">	picture.jpg is located in the images folder at the root of the current web
<img src="../picture.jpg">	picture.jpg is located in the folder one level up from the current folder

Hyperlinks

testo del link

- Also URLs with other protocols:

write to me!

```
<base href="https://www.mydomain.com/" target="_blank">
```

Hyperlinks – inner target

```
<!DOCTYPE html>
<html>
<body>
<h2>The inner target</h2>
<a href="#pippo" >GO TO PIPPO!</a>
<p>one of many lines...</p>

...

<p>one of many lines...</p>
<a name="pippo"/>
<H1>PIPPO IS HERE!</H1>
<p>one of many lines...</p>

...

<p>one of many lines...</p>
</body>
</html>
```

Iframes

`<iframe src="demo_iframe.htm"></iframe>`

`<iframe height="300px" width="100%" src="demo_iframe.htm"` `name="iframe_a"``></iframe>`

`<p><a href="https://www.w3schools.com"` `target="iframe_a"``>W3Schools.com</a></p>`

Warning:
if the site you are trying to access limits embedding
(via iframe, frame, embed, object) to the same origin using the X-Frame-Options header, you can not include it!

Character Encoding

Describes the mapping between chars and integer numbers

- ASCII was the first **character encoding standard** (also called character set). 128 different alphanumeric characters: {numbers (0-9), English letters (A-Z), some special characters e.g. ! $ + - ( ) @ < >, some non printable ones like ESC}.

- ISO-8859-1: default character set for HTML 4. This character set supported 256 different character codes.

- ANSI (Windows-1252): the original Windows character set. Adde 32 extra characters to ISO-8859-1.

- UTF-8 (Unicode) was supported by HTML4, is the default character set for HTML5.

Entities

Result	Description	Entity Name	Entity Number
	non-breaking space		
<	less than	<	<
>	greater than	>	>
&	ampersand	&	&
"	double quotation mark	"	"
'	single quotation mark (apostrophe)	'	'
¢	cent	¢	¢
£	pound	£	£
¥	yen	¥	¥
€	euro	€	€
©	copyright	©	©
®	registered trademark	®	®

URL Encoding

Character	From Windows-1252	From UTF-8
space	%20	%20
!	%21	%21
"	%22	%22
#	%23	%23
$	%24	%24
%	%25	%25
&	%26	%26
'	%27	%27
(	%28	%28
)	%29	%29

- See also https://www.w3schools.com/tags/ref_urlencode.asp

Introduzione alla programmazione web – Marco Ronchetti 2020 – Università di Trento

Metadata (head)

```
<meta charset="UTF-8">
<meta name="description" content="Free Web tutorials">
<meta name="keywords" content="HTML,CSS,XML,JavaScript">
<meta name="author" content="John Doe">


<meta http-equiv="refresh" content="30">
```

<meta charset="utf-8"> Charset declaration

- Link CSS/JS

Metadata (head)

\<meta name="viewport" content="width=device-width, initial-scale=1.0"\>

With viewport, on phone

Without viewport, on phone

Introduzione alla programmazione web – Marco Ronchetti 2020 – Università di Trento

What I expect you to know…

Ignore (for now) references to "styles"

HTML Tutorial

HTML HOME
HTML Introduction
HTML Editors
HTML Basic
HTML Elements
HTML Attributes
HTML Headings
HTML Paragraphs
~~HTML Styles~~
HTML Formatting
HTML Quotations
HTML Comments
HTML Colors

~~HTML CSS~~
HTML Links
HTML Images
HTML Tables
HTML Lists
HTML Blocks
~~HTML Classes~~
~~HTML Id~~
HTML Iframes
~~HTML JavaScript~~
HTML File Paths
HTML Head
~~HTML Layout~~
~~HTML Responsive~~

HTML Computercode
HTML Entities
HTML Symbols
HTML Charset
HTML URL Encode
~~HTML XHTML~~

HTML Forms

HTML Forms
HTML Form Elements
HTML Input Types
HTML Input Attributes

Introduzione alla programmazione web – Marco Ronchetti 2020 – Università di Trento

Tables

<table> defines the tabella

<th> or <tr> define the rows

<td> defines every single cell

<caption> defines the caption (didascalia)

The attributes
- colspan="n"
- rowspan="n"

define the cell span over columns and rows.

Tables

Example

```html
<h2>Cell that spans two columns</h2>

<table style="width:100%">
  <tr>
    <th>Name</th>
    <th colspan="2">Telephone</th>
  </tr>
  <tr>
    <td>Bill Gates</td>
    <td>555 77854</td>
    <td>555 77855</td>
  </tr>
    <tr>
    <td>Donald Duck</td>
    <td>945 2626</td>
    <td>0471 23706</td>
  </tr>
</table>
```

Cell that spans two columns

Name	Telephone	
Bill Gates	555 77854	555 77855
Donald Duck	945 2626	0471 23706

HTML Editors

HTML editors are called "WYSIWYG" (Uizziuaig)

- *What You See Is What You Get!*

Some free, multiplatform HTML Editors:

- Aptana
- Atom
- Bluefish
- Netbeans

(see next slides)

Multiplatform, free HTML editors - 1
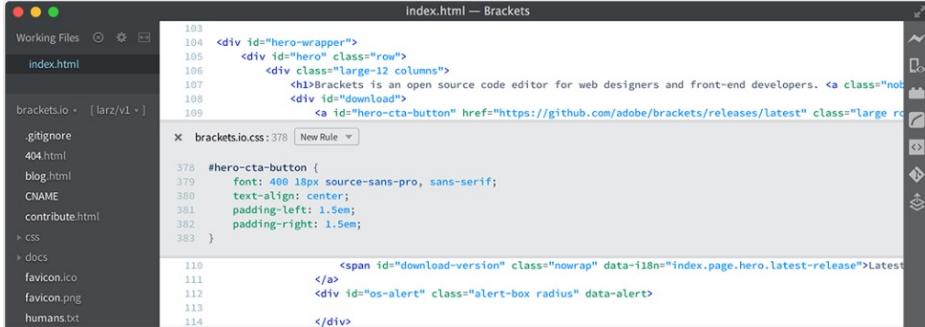
Brackets - open source text editor.
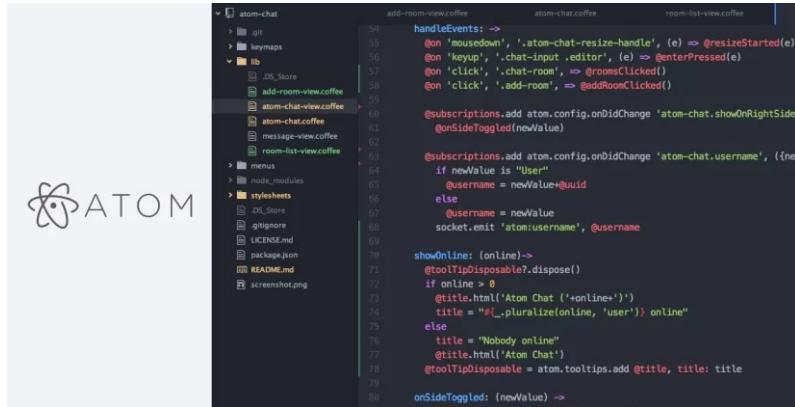
Atom - open source web application IDE.
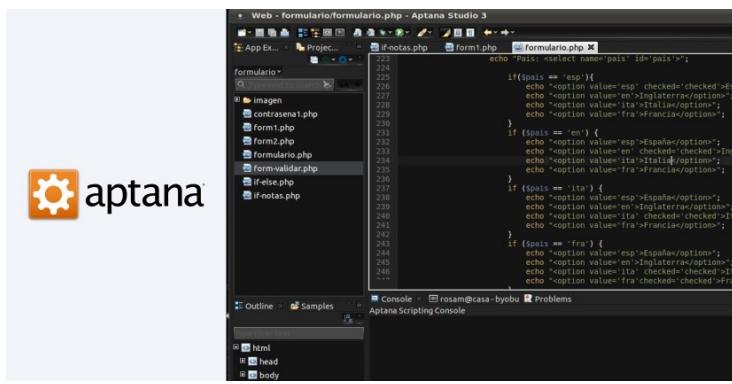
Aptana Studio 3 - open source
web application IDE.
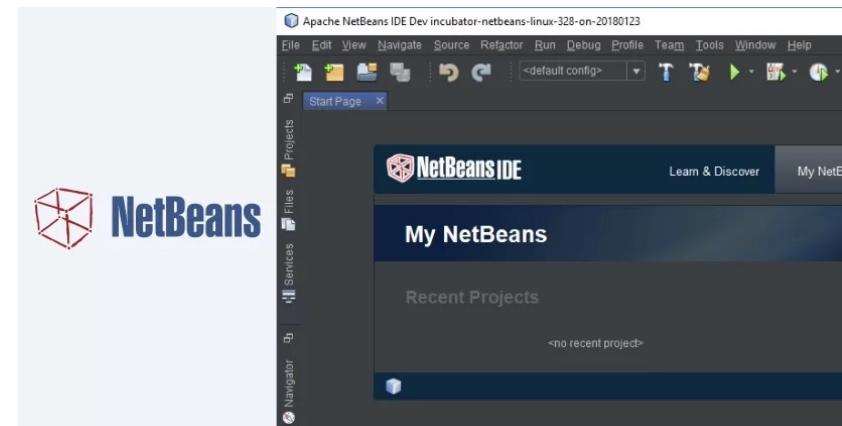
NetBeans contains a free
and open source HTML editor.

HTML Editors pros and cons:

- Pros
  - Good for beginners
  - You do not have to know much HTML to use editors
  - It is easy to create complex tables, image maps, and use advanced functions such as Style Sheets and JavaScript
  - Easy for design due to the WYSIWYG interface

- Cons
  - They do not always generate clean HTML
    - Add redundant tags
    - Add their own tags
  - They do not offer good graphic tools to manipulate images
  - It can be hard to fix HTML tags since every time you open a document, code is added.

Homework -1

Create a webpage that looks like this: ➡

The image can be retrieved from
https://latemar.science.unitn.it/segue_userFiles/2021ProgrammazioneWeb/2.jpg

The index items should point to the corresponding sections below.

The final link should connect to
https://piazza.com/unitn.it/spring2022/introduzione allaprogrammazioneweb2122/home

My First Website

Indice

1. Images
2. Formatting
3. Lists
4. Links

Images

Formatting

This is an example of **bold**, *italic*, ***bold italic***,

Lists

List of tags used in the body of this page

- <h1>, <h2>
- , <ul>,
- <a>
- <b>, <i>
- dragonfruit

Links

The web site of this course on Piazza

Introduzione alla programmazione web – Marco Ronchetti 2

Homework -2

Change the webpage so that it takes the following aspect:

Introduzione alla programmazione web – Marco Ronchetti 2020 – Università di Trento

Homework -3

Create a webpage that looks like this:

Suggestion: bgcolor="red"

Readings

- https://developer.mozilla.org/en-US/docs/Learn/HTML/Introduction_to_HTML

- https://www.w3schools.com/html/

- http://www.cellbiol.com/bioinformatics_web_development/chapter-3-your-first-web-page-learning-html-and-css/