

<WA1/>

2020

# (Modern) HTML

A fast-paced introduction to HTML5

Enrico Masala

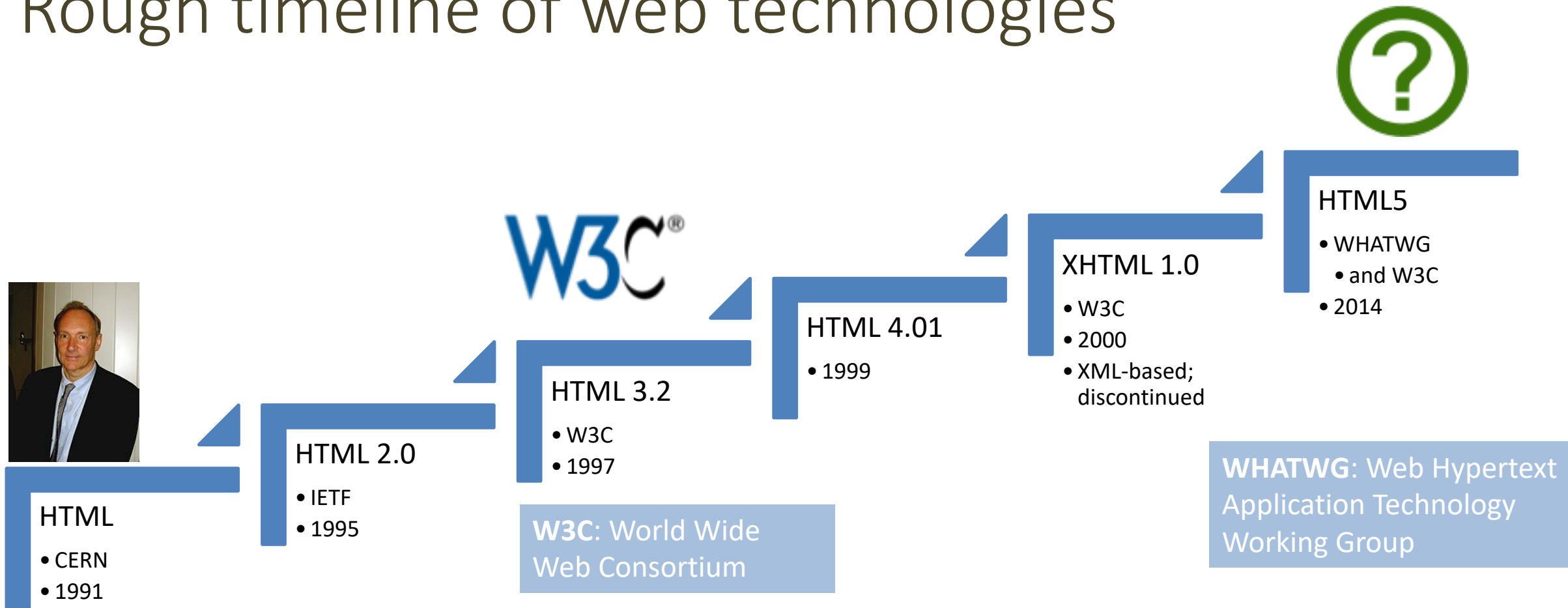
Fulvio Corno



# Goal

- Quick introduction to the essential features of HTML
  - Document structure and syntax
  - Essential elements
  - Semantic elements
  - Page structuring
- Aligned to what HTML 5 looks like in year 2020

# Rough timeline of web technologies



<https://html.spec.whatwg.org/multipage/>

<https://html.spec.whatwg.org/dev/>

# A different approach

HTML 4.01



web pages



design

HTML 5

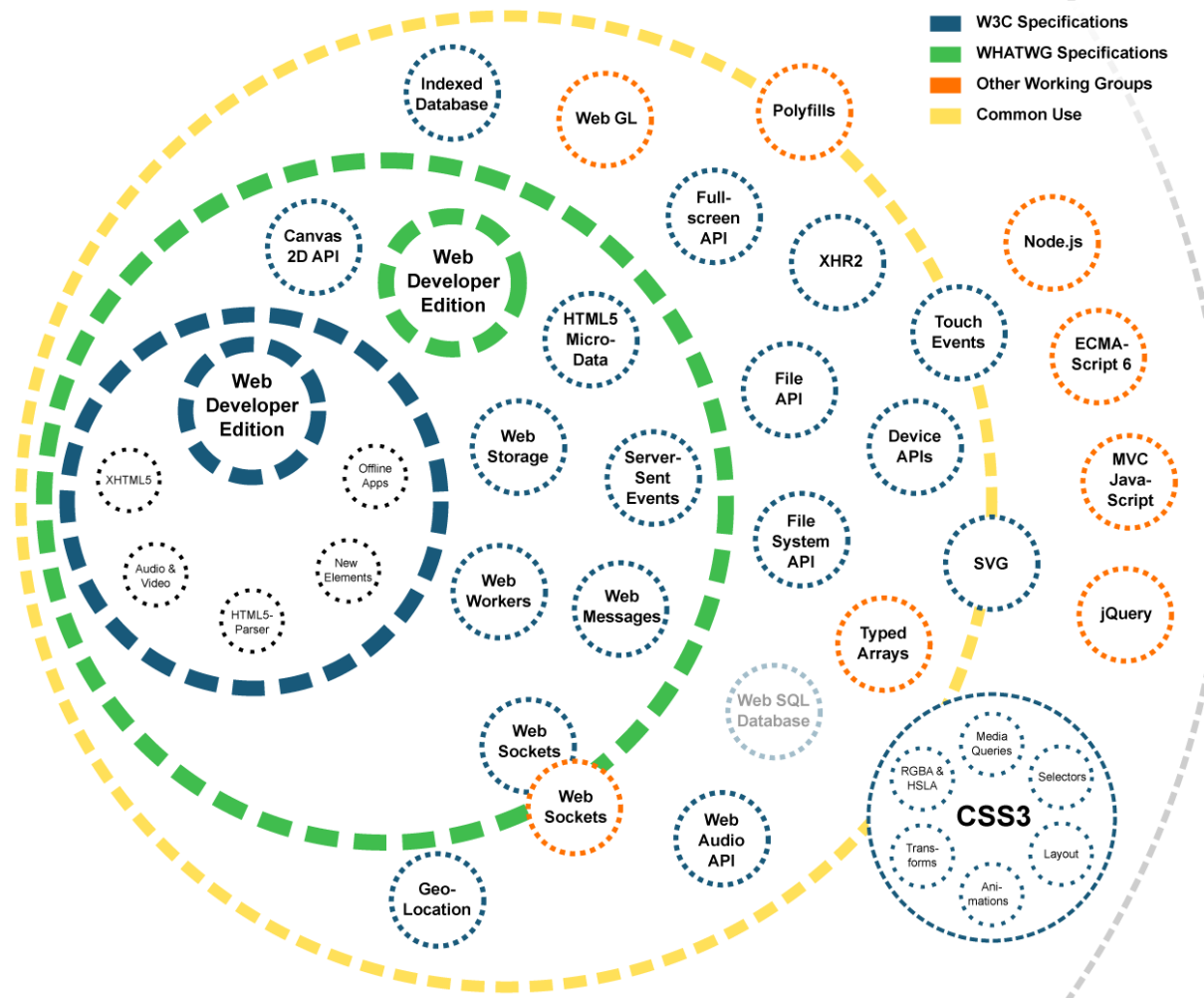


web applications



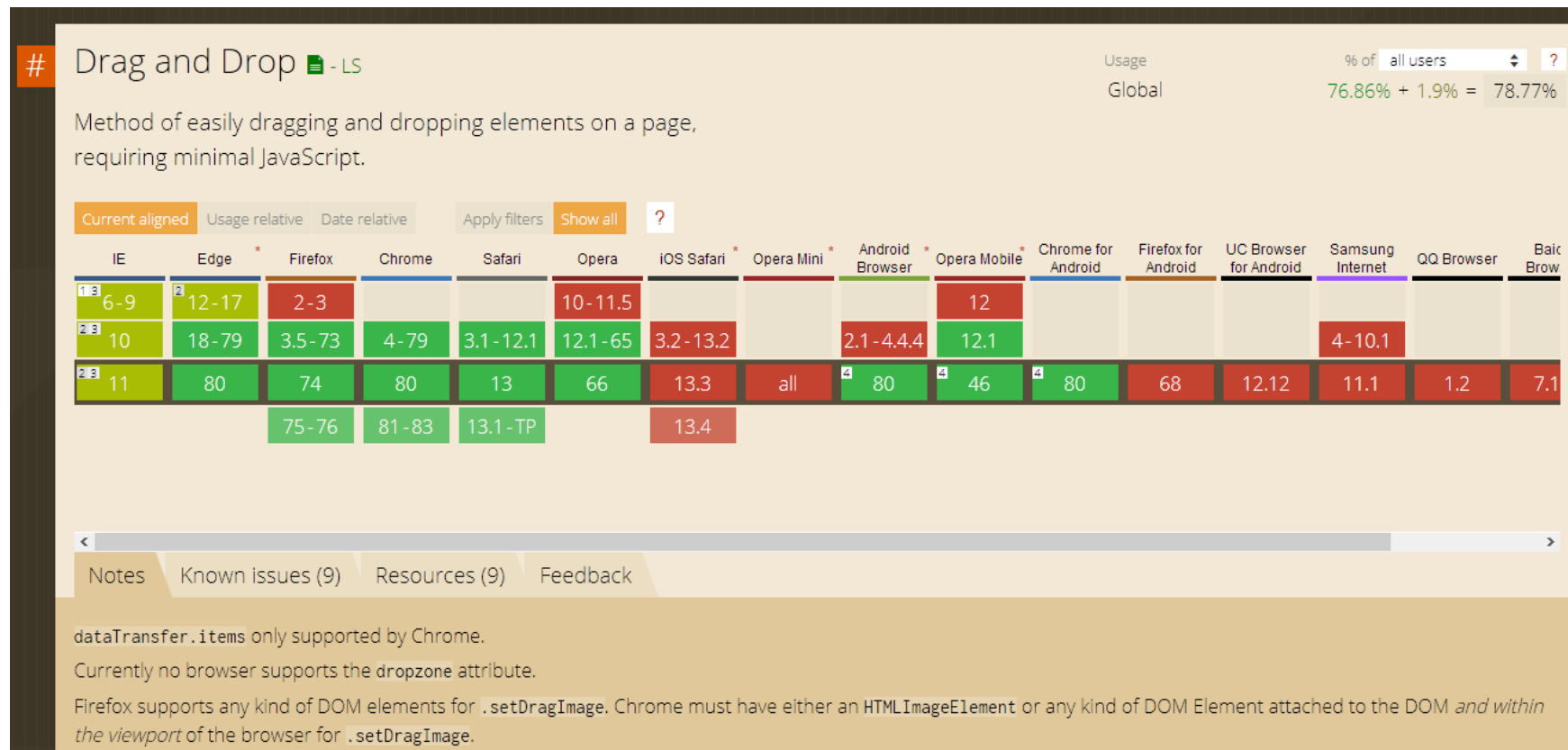
user interaction

# HTML 5 is a family of standards



# Live standard: Browser Compatibility?

- <https://caniuse.com/>



# A basic HTML document

```
<!DOCTYPE html>
<html lang="en">
  <head>
    <title>Sample page</title>
  </head>
  <body>
    <h1>Sample page</h1>
    <p>This is a <a href="demo.html">simple</a> sample.</p>
    <!-- this is a comment -->
  </body>
</html>
```

- Unicode Text File
- DOCTYPE declaration
- Nested tree of Elements
  - Strict nesting
- Element:
  - Start tag: <a>
  - Attributes: href="demo.html"
    - Name=value pairs
  - End tag: </a>
  - Inner text: simple

# Mandatory structure

- First line: `<!DOCTYPE html>`
- `<html>` root element, contains the whole file
  - `<head>` (title, other metadata, load style sheets, load javascript code)
  - `<body>` (actual page content)

```
<!DOCTYPE html>
<html lang="en">
  <head>
    <title>Sample page</title>
  </head>
  <body>
    <h1>Sample page</h1>
    <p>This is a <a href="demo.html">simple</a> sample.</p>
    <!-- this is a comment -->
  </body>
</html>
```



# HTML parsing into a DOM tree

```
<!DOCTYPE html>
<html lang="en">
  <head>
    <title>Sample page</title>
  </head>
  <body>
    <h1>Sample page</h1>
    <p>This is a <a href="demo.html">simple</a> sample.</p>
    <!-- this is a comment -->
  </body>
</html>
```

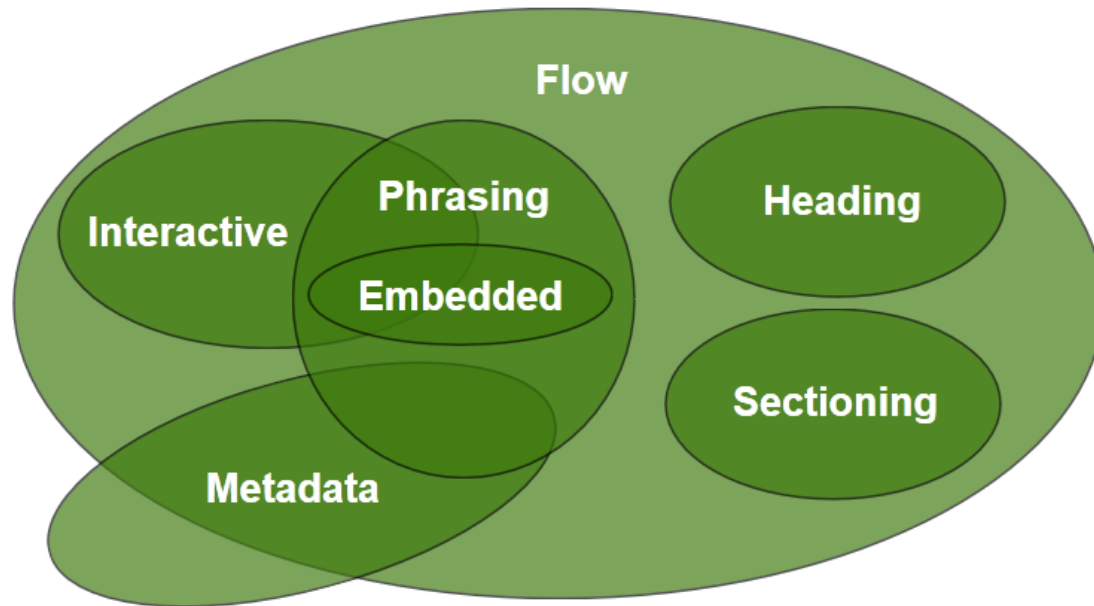
Parsing

```
DOCTYPE: html
html lang="en"
  head
    #text: 
    title
      #text: Sample page
    #text: 
    #text: 
  body
    #text: 
    h1
      #text: Sample page
    #text: 
    p
      #text: This is a
      a href="demo.html"
        #text: simple
      #text: sample.
    #text: 
    #comment: this is a comment
    #text: 
```

# HTML elements

- Elements are used to define the *meaning* of a portion of a document
  - Semantic markup
- The meaning will be rendered graphically, according to the style sheets
  - No HTML code should deal with presentation issues
- Each element has a default *display* value
  - display:block
    - Full-width
    - Starts on a new line
    - Top-to-bottom layout
  - display:inline
    - Does not start a new line
    - Occupies just the necessary space
    - Left-to-right layout

# Categories of HTML elements



- *Sectioning* content defines headings and footers
- *Heading content* defines the header of a section
- *Phrasing* content is the *text* of the document, as well as elements that mark up that text *at the intra-paragraph level*.
  - Runs of phrasing content form paragraphs.
- *Flow* content includes most types of elements

<https://html.spec.whatwg.org/dev/dom.html#kinds-of-content>

# Sections and Headings (block)

## Sectioning content

- article
- aside
- nav
- section

## Heading content

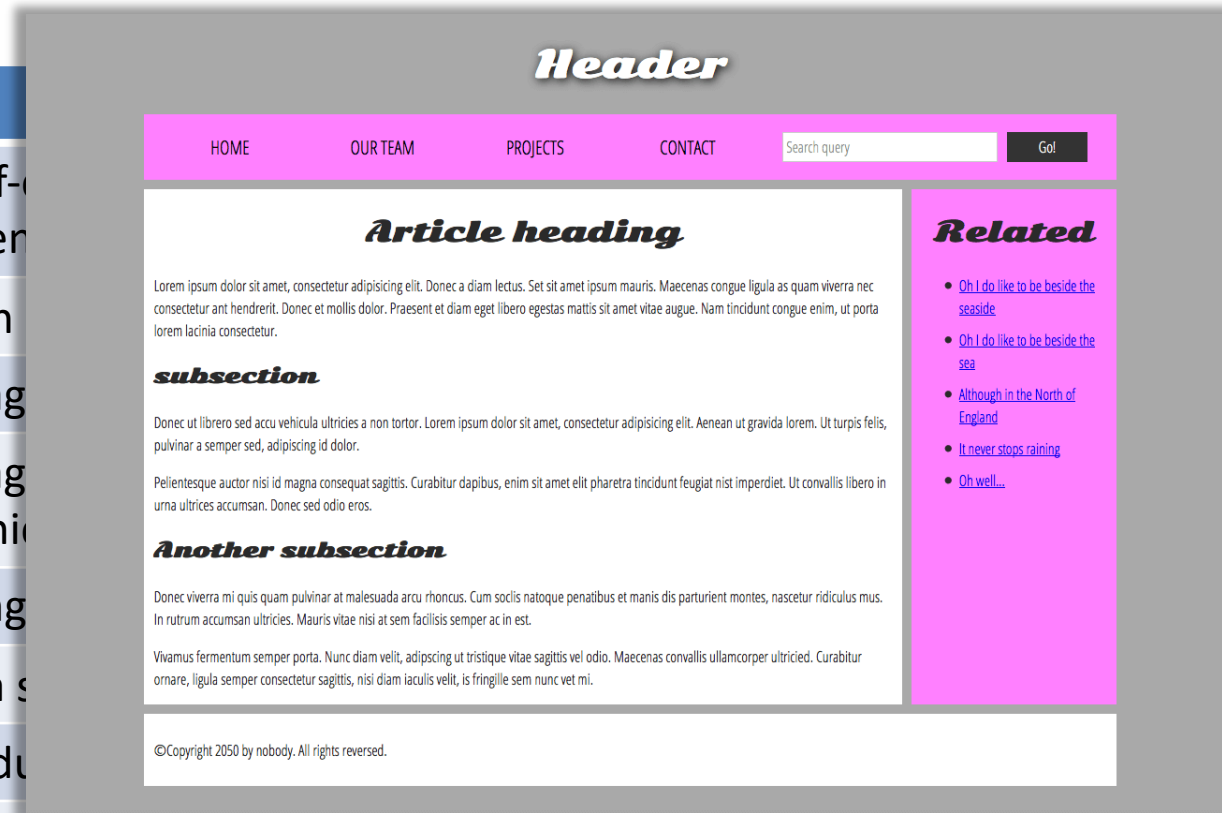
- h1
- h2
- h3
- h4
- h5
- h6
- hgroup

# Usage guide

Element	Usage
<code>&lt;article&gt;</code>	complete, or self-contained, composition in a document, page, application, or site and that is, in principle, independently distributable or reusable
<code>&lt;section&gt;</code>	A generic section of a document or application. A thematic grouping of content, typically with a heading
<code>&lt;nav&gt;</code>	A section of a page that links to other pages or to parts within the page: a section with navigation links
<code>&lt;aside&gt;</code>	A section of a page that consists of content that is tangentially related to the content around the <a href="#">aside</a> element, and which could be considered separate from that content (such as sidebars)
<code>&lt;h1&gt;</code> - <code>&lt;h6&gt;</code>	A section heading
<code>&lt;hgroup&gt;</code>	The heading of a section, that groups multiple h1-h6 elements, e.g., in case of subheadings
<code>&lt;header&gt;</code>	A group of introductory or navigational aids
<code>&lt;footer&gt;</code>	typically contains information about its section such as who wrote it, links to related documents, copyright data, and the like

# Usage guide

Element	Usage
<article>	complete, or self-contained principle, independent of the rest of the document
<section>	A generic section of a document
<nav>	A section of a page containing navigation links
<aside>	A section of a page that is separate from the main content, and which can be removed without affecting the overall meaning of the document (e.g. sidebars)
<h1> - <h6>	A section heading
<hgroup>	The heading of a section
<header>	A group of introductory elements
<footer>	typically contains information about its section such as who wrote it, links to related documents, copyright data, and the like



# Grouping content (block)

Element	Usage
<code>&lt;p&gt;</code>	a paragraph of text
<code>&lt;hr&gt;</code>	horizontal rule (represents a paragraph-level thematic break)
<code>&lt;pre&gt;</code>	block of preformatted text
<code>&lt;blockquote&gt;</code>	a section that is quoted from another source
<code>&lt;ol&gt;</code>	a list of items, where the items have been intentionally ordered. The items of the list are <code>&lt;li&gt;</code> elements
<code>&lt;ul&gt;</code>	a list of items, where the order of the items is not important. The items of the list are <code>&lt;li&gt;</code> elements
<code>&lt;menu&gt;</code>	an unordered list <code>&lt;ul&gt;</code> , whose elements are commands that may be executed
<code>&lt;dl&gt;</code>	a list of definitions. Each definition has a name ( <code>&lt;dt&gt;</code> , definition term) and a value ( <code>&lt;dd&gt;</code> , definition description)
<code>&lt;figure&gt;</code>	a sectioning element that can be used to annotate illustrations, diagrams, photos, code listings, etc. May include <code>&lt;img&gt;</code> or other content. May include <code>&lt;figcaption&gt;</code>
<code>&lt;main&gt;</code>	represents the dominant contents of the document
<code>&lt;div&gt;</code>	the <code>&lt;div&gt;</code> element has no special meaning at all. It can be used with the <code>class</code> , and <code>title</code> attributes to mark up semantics common to a group of consecutive elements. Use as a last resort.

# Phrasing content (inline)

- **a**
- abbr
- audio
- **b**
- br
- button
- canvas
- cite
- code
- data
- datalist
- del
- dfn
- **em**
- embed
- **i**
- **img**
- input
- ins
- kbd
- label
- mark
- output
- picture
- progress
- q
- s
- samp
- select
- slot
- small
- **span**
- **strong**
- sub
- sup
- template
- textarea
- time
- var
- video
- wbr
- Insert content or format content inside a single paragraph
- Normally formatted left-to-right, and wrap at the end of line

<https://html.spec.whatwg.org/dev/text-level-semantic.html>



# Interactive content (inline)

- **a**
- audio
- **button**
- details
- embed
- iframe
- **input**
- label
- **select**
- **textarea**
- video
- Elements specifically intended for user interaction
- May be used to compose a form
- May include multimedia interactive content
- In the simplest form, they may just be a link
  - `<a href="URL">anchor</a>`

# Tabular data

- `<table>` container
- `<thead>` groups header rows, `<tbody>` groups data rows, `<tfoot>` groups summary rows
- `<tr>` table row, includes
  - `<td>` cell with table data
  - `<th>` cell with table heading
- May have a `<caption>`
- `<colgroup>` may apply common attributes to a set of `<col>` columns
- `rowspan` and `colspan` attributes for creating complex grids

Data about the planets of our solar system (Planetary facts taken from [Nasa's Planetary Fact Sheet - Metric](#)).

		Name	Mass (10 <sup>24</sup> kg)	Diameter (km)	Density (kg/m <sup>3</sup> )	Gravity (m/s <sup>2</sup> )	Length of day (hours)	Distance from Sun (10 <sup>6</sup> km)	Mean temperature (°C)	Number of moons	Notes
Terrestrial planets		Mercury	0.330	4,879	5427	3.7	4222.6	57.9	167	0	Closest to the Sun
		Venus	4.87	12,104	5243	8.9	2802.0	108.2	464	0	
		Earth	5.97	12,756	5514	9.8	24.0	149.6	15	1	Our world
		Mars	0.642	6,792	3933	3.7	24.7	227.9	-65	2	The red planet
Jovian planets	Gas giants	Jupiter	1898	142,984	1326	23.1	9.9	778.6	-110	67	The largest planet
		Saturn	568	120,536	687	9.0	10.7	1433.5	-140	62	
	Ice giants	Uranus	86.8	51,118	1271	8.7	17.2	2872.5	-195	27	
		Neptune	102	49,528	1638	11.0	16.1	4495.1	-200	14	
Dwarf planets		Pluto	0.0146	2,370	2095	0.7	153.3	5906.4	-225	5	Declassified as a planet in 2006, but this <a href="#">remains controversial</a> .

# Key attributes (on all elements)

## `class="aa bb cc"`

- Space-separated list of class identifiers
- Represents all the classes that this element belongs to
- Heavily used in CSS and JS for matching portion of contents

## `id="unique"`

- Represents an element's unique identifier
- Must be unique within this specific HTML document
- Heavily used in CSS and JS for finding/matching this specific element

# Key attributes (on all elements)

**style="css fragment"**

- Apply a set of CSS declarations to this specific HTML element
- Use sparingly, prefer CSS rules that map to element `id` and `class`

# Wildcard elements

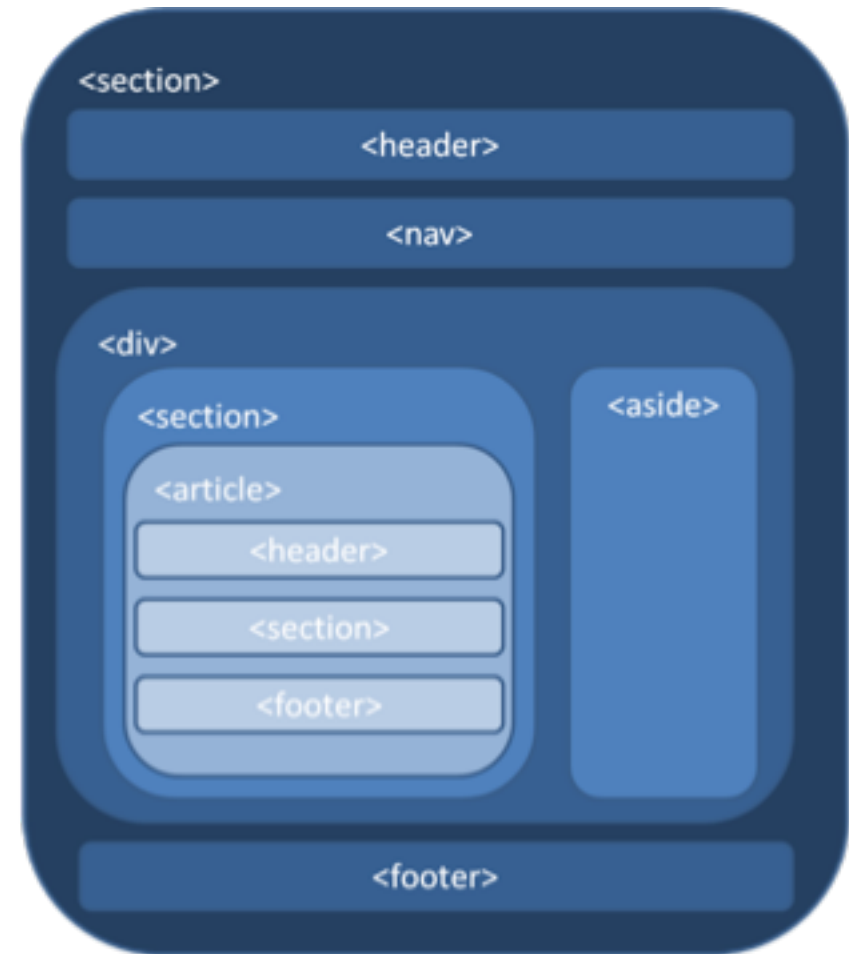
## `<div>...</div>`

- Block-level container element
- No predefined semantics
- Invisible layout
- Used to mark-up blocks or groups of blocks, with specific classes

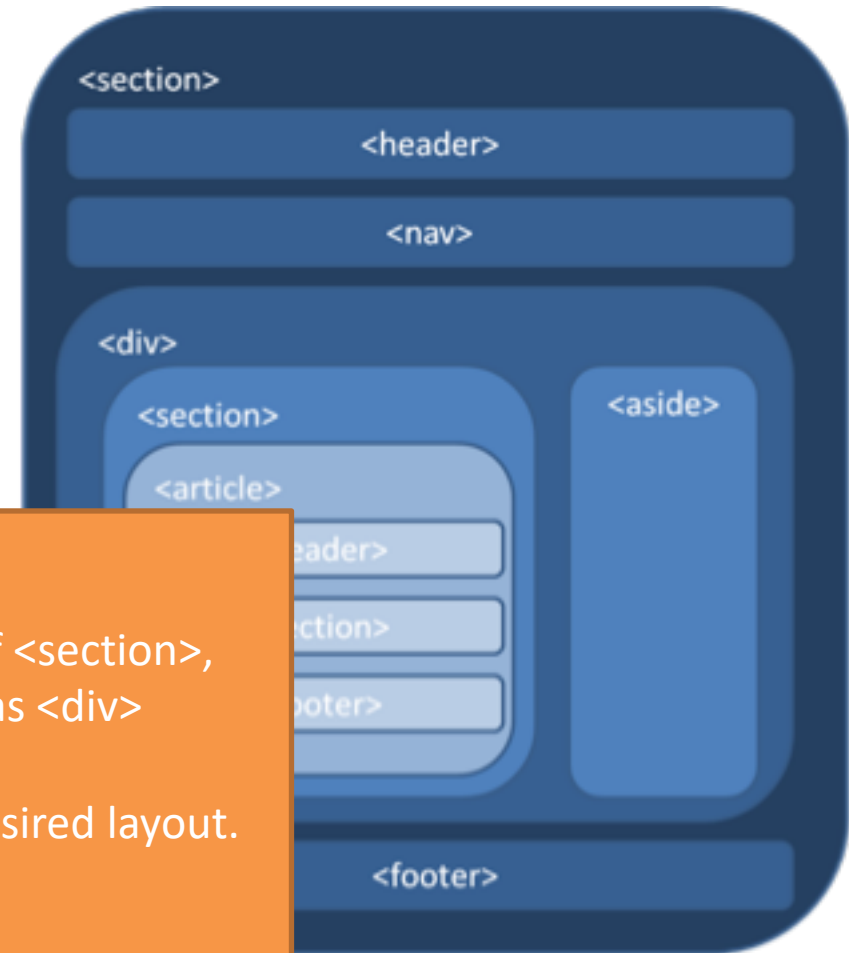
## `<span>...</span>`

- Inline-level container element
- No predefined semantics
- Invisible layout
- Used to mark-up parts of a paragraph, with specific classes

# Opaque vs. Semantic elements



# Opaque vs. Semantic elements



Note: the default rendering of `<section>`, `<article>`, etc is the same as `<div>`

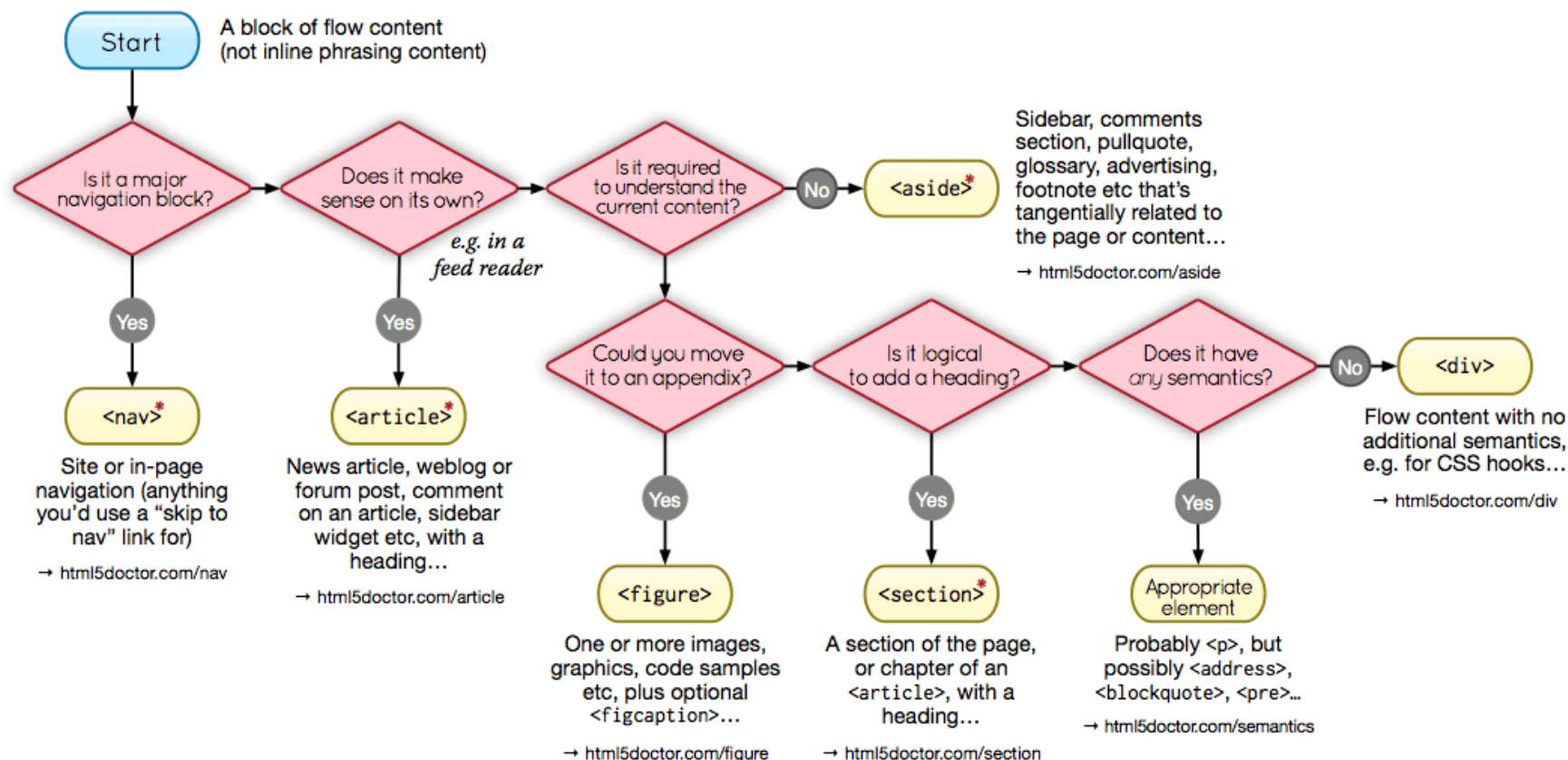
Only CSS rules will apply the desired layout.

# Typical layout



```
1  <body>
2      <header>
3          <nav>
4
5          </nav>
6      </header>
7      <aside>
8
9      </aside>
10     <main>
11
12     </main>
13     <footer>
14
15     </footer>
16 </body>
```





\* Sectioning content element

These four elements (and their headings) are used by HTML5's outlining algorithm to make the document's outline  
→ [html5doctor.com/outline](http://html5doctor.com/outline)

2011-07-22 v1.5  
For more information:  
[www.html5doctor.com/semantics](http://www.html5doctor.com/semantics)

# WHAT Beautiful HTML LOOKS LIKE

HTML5 and it's new elements make for the most beautiful HTML, yes.

## DOCTYPE

HTML has the best DOCTYPE ever.

## TITLE

Title of the site is simple and clear. Purpose of page is first, a sequence is used, and ends with title of the site.

## CSS

Only one single stylesheet is used (media types are declared inside stylesheet), and only served as paid browsers. If it is served a natural stylesheet. Better if it gets ending.

## BODY

ID applied to body to allow for unique page styling without any additional markup.

## DESCRIPTIVE BLOCKS

header, section, article, aside... all appropriately describe the content they contain better than the div of old.

## HEIRARCHY

Title tags are reserved for real content, and follow a clear hierarchy.

## IMAGE ATTRIBUTES

Images include alternate text, mostly for visually impaired users but also for validation. Height and width applied for rendering efficiency.

## APPROPRIATE TAGS

Links are marked up as lists, depending on the needs of the list: numbered, ordered, and the unordered definitions.

## SEMANTIC CLASSES

Beyond appropriate element names, classes and IDs are semantic: they describe without specifying (e.g. "url" is much better than "url")

## CLASSES

Are used any time similar styling needs to be applied to multiple elements (reusability).

## IDS

Are used any time an element appears only once on the page and cannot be targeted successfully any other way.

## JAVASCRIPT

jQuery (the most beautiful JavaScript library) is served from Google. Only a single JavaScript file is loaded. Both scripts are collected at the bottom of the page.

## INDENTATION

Tabs or spaces are used to indent the code to properly indicate parent-child relationship of tags.

## FILE PATHS

File resources use relative file paths for efficiency. Content file paths are absolute, ensuring content is syndicated.

## INCLUDES

Things common across multiple pages are inserted via server side includes. (Doesn't have to be PHP of course, any type of include is fine)

## CHARACTERS ENCODED

With a special character, it's encoded.

## DYNAMIC

Things that need to be dynamic, are dynamic.

## COMMENTS

Comments are included for things that may not be immediately obvious upon reviewing the code.

## FREE FROM STYLING

Nothing on the page applies styling or even implies what the styling might be. Everything on the page is either composed via content, context, or describing.

## VALID

The markup adheres to W3C validation. Tags are closed, required attributes used, etc.

<https://css-tricks.com/what-beautiful-html-code-looks-like/>

POLITECNICO DI TORINO

Applicazioni Web I - Web Applications I - 2019/2020

# References

- HTML: The Living Standard, Developer's Edition — Last Updated 20 March 2020 - <https://html.spec.whatwg.org/dev/>
- MDN: HTML: Hypertext Markup Language - <https://developer.mozilla.org/en-US/docs/Web/HTML>
- A Modern HTML Introductory Tutorial - <http://fastwebstart.com/modern-html-tutorial/> (and linked resources)
- <https://css-tricks.com/what-beautiful-html-code-looks-like/>

# License

- These slides are distributed under a Creative Commons license “**Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)**”
- **You are free to:**
  - **Share** — copy and redistribute the material in any medium or format
  - **Adapt** — remix, transform, and build upon the material
  - The licensor cannot revoke these freedoms as long as you follow the license terms.
- **Under the following terms:**
  - **Attribution** — You must give [appropriate credit](#), provide a link to the license, and [indicate if changes were made](#). You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
  - **NonCommercial** — You may not use the material for [commercial purposes](#).
  - **ShareAlike** — If you remix, transform, or build upon the material, you must distribute your contributions under the [same license](#) as the original.
  - **No additional restrictions** — You may not apply legal terms or [technological measures](#) that legally restrict others from doing anything the license permits.
- <https://creativecommons.org/licenses/by-nc-sa/4.0/>

