

<WA/>

2025

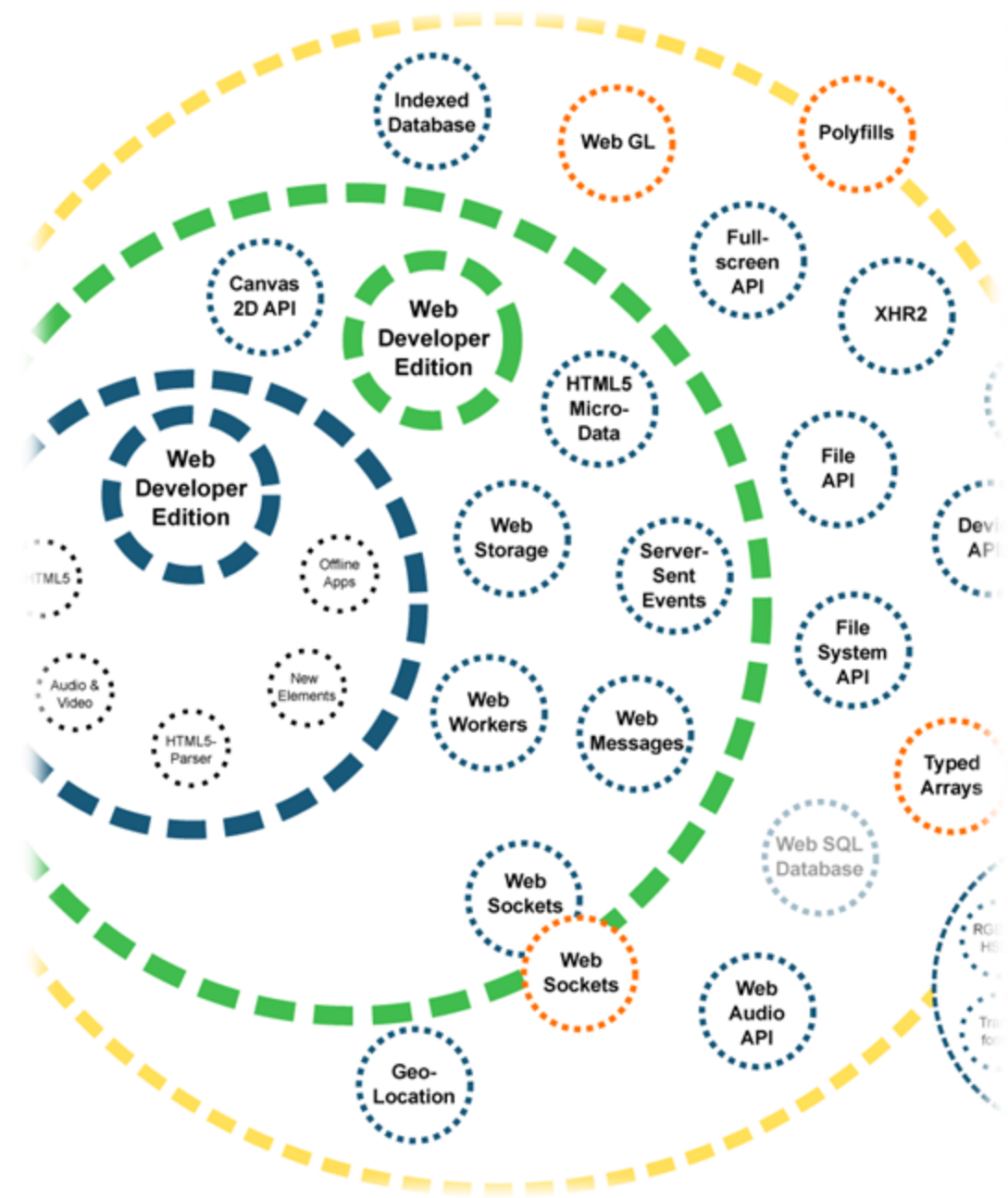
(Modern) HTML

A fast-paced introduction to HTML5

Fulvio Corno

Luigi De Russis

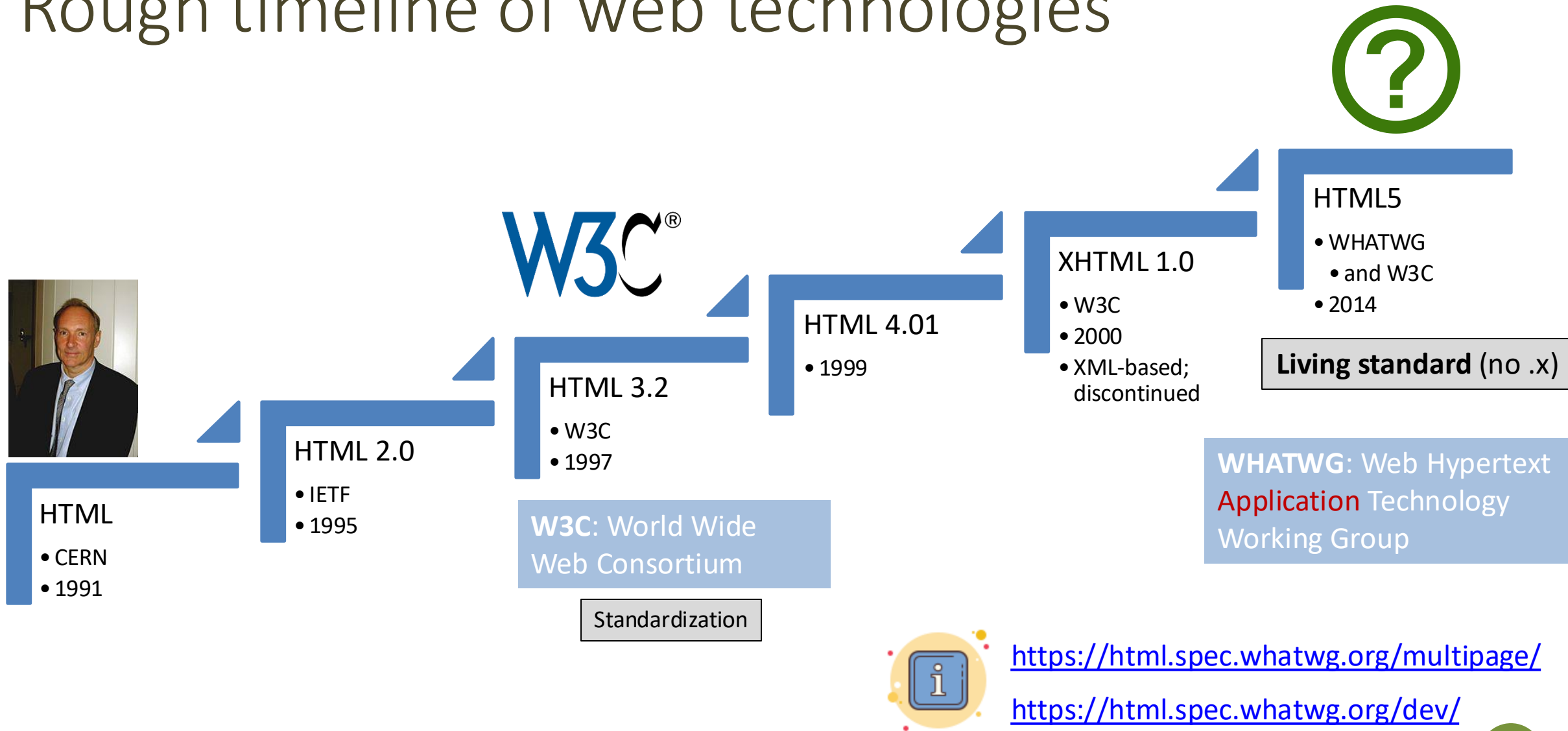
Enrico Masala



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



A different approach



Welcome to the WHATWG community
Maintaining and evolving HTML since 2004

HTML 4.01



web pages



design

HTML 5

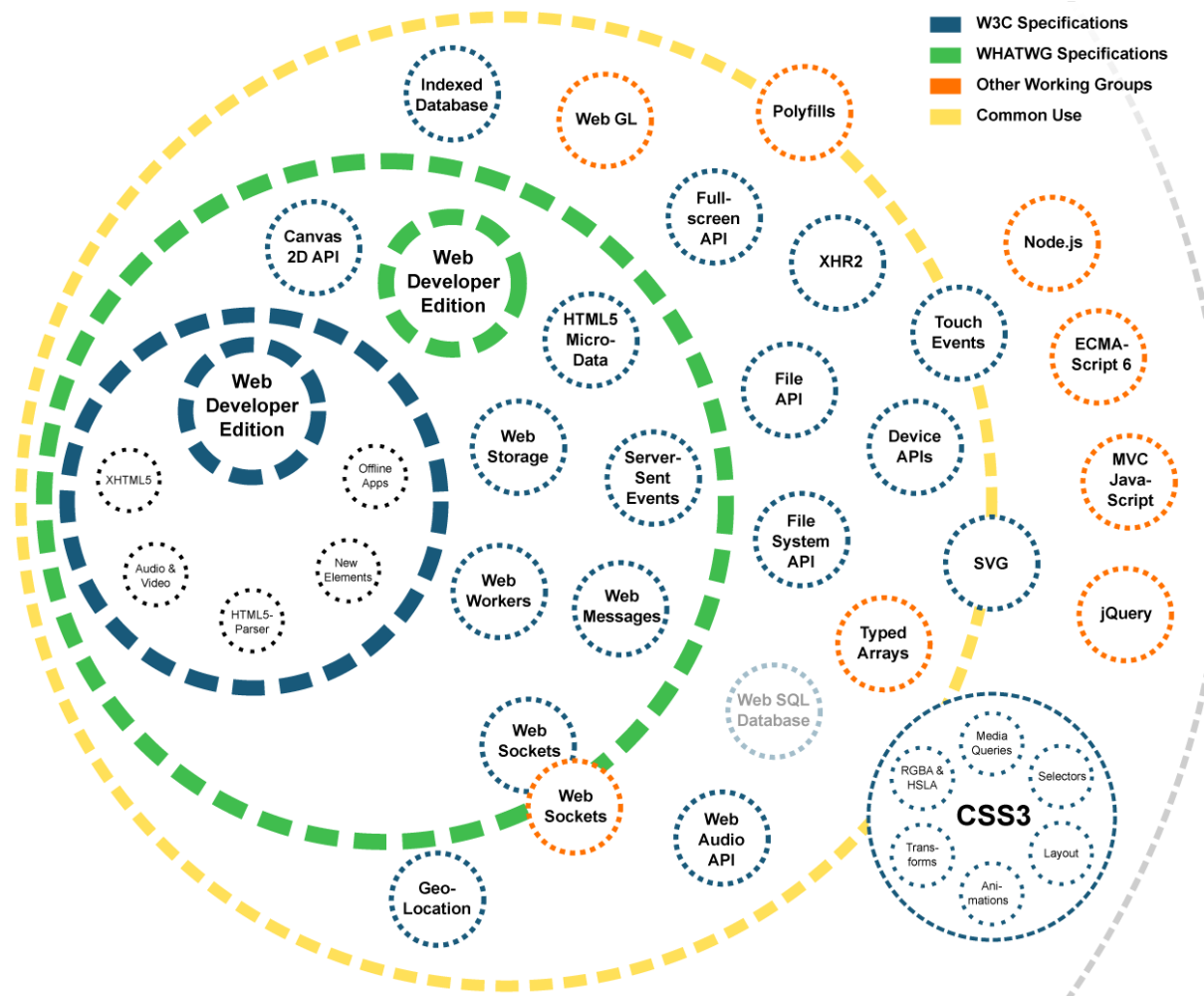


web applications



user interaction

HTML 5 is a family of standards



Live standard: Browser Compatibility?

- <https://caniuse.com/?search=drag%20drop>



- <https://www.browsers.fyi/>



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>
 - **Attribute(s):** href="demo.html"
 - Name=value pairs
 - End tag:
 - Inner text: simple

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: 
```


Mandatory structure

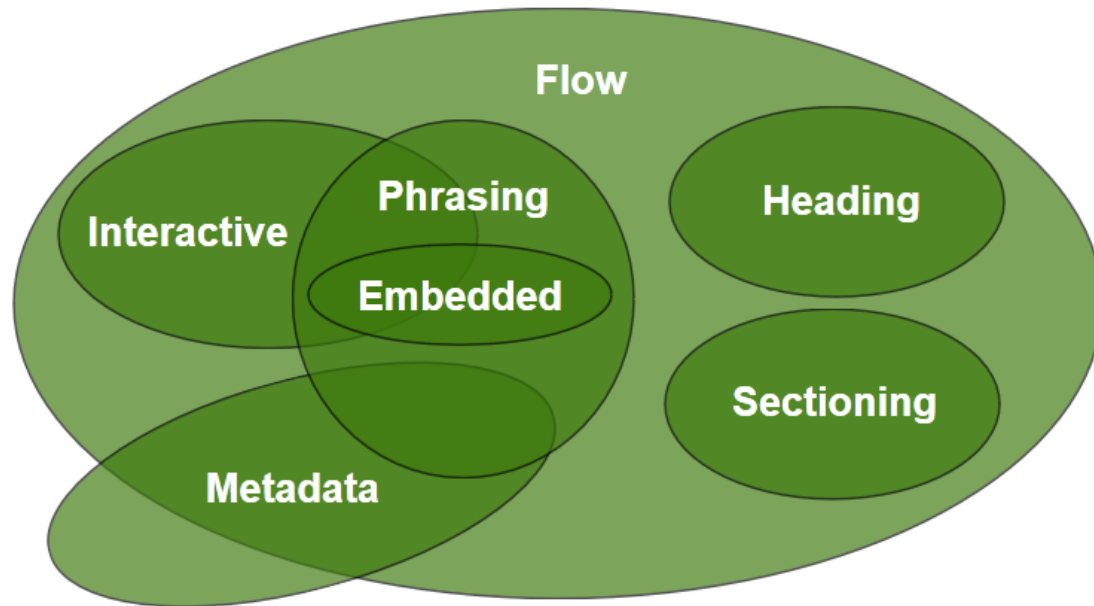
- First line: `<!DOCTYPE html>`
- `<html>` root element, contains the whole file
 - `<head>` (title, other metadata, load style sheets, load JavaScript code)
 - `<body>` (actual page content)

```
<!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 elements

- Semantic HTML tags are used to define the *meaning* of a portion of a document
 - Semantic markup
- The meaning will be rendered graphically, according to the (CSS) style sheets
 - No HTML code should deal with presentation issues
- Example of semantic HTML tags
 - `<title>` `<h1>` `<body>` ...
- Example of non-semantic HTML
 - `<div>` ``
 - No indication as to what type of content they contain or what role that content plays on the page

Categories of HTML elements



- *Details in next slides*
 - *Sectioning and heading* content defines headings and footers
 - *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`

- `header`
- `footer`

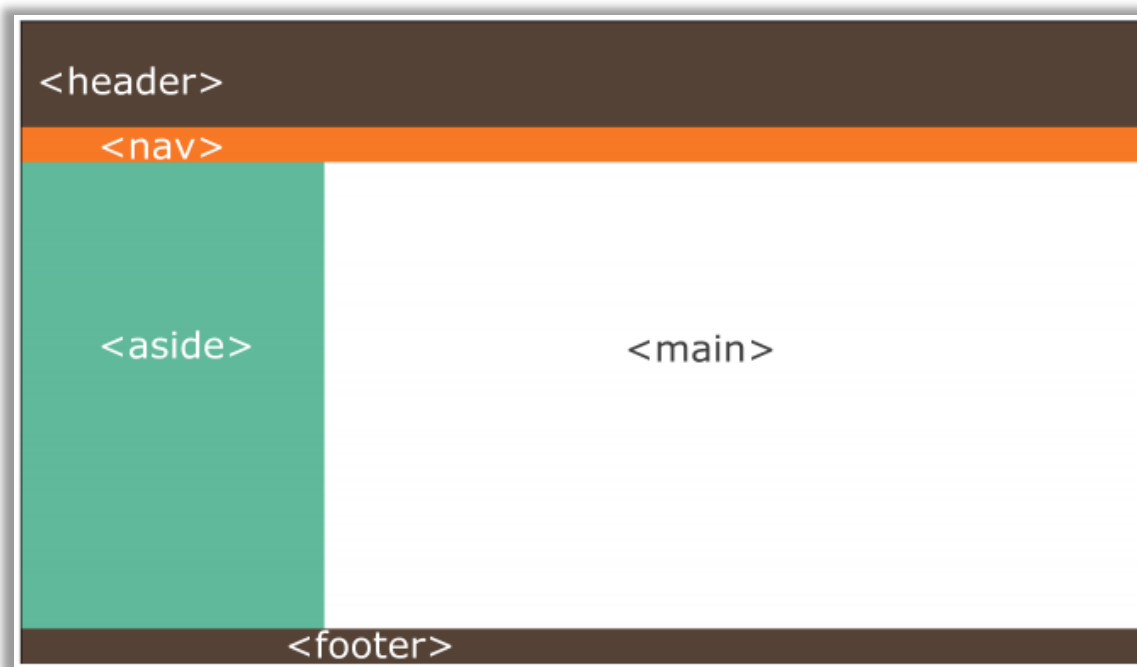
Heading content

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

Usage guide

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

Typical conceptual 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>
```

NB: Actual layout depends on browser default + CSS rules

Grouping content (block)

Element	Usage
<code><p></code>	a paragraph of text
<code><hr></code>	horizontal rule (represents a paragraph-level thematic break)
<code><pre></code>	block of preformatted text
<code><blockquote></code>	a section that is quoted from another source
<code></code>	a list of items, where the items have been intentionally ordered. The items of the list are <code></code> elements
<code></code>	a list of items, where the order of the items is not important. The items of the list are <code></code> elements
<code><menu></code>	an unordered list <code></code> , whose elements are commands that may be executed
<code><dl></code>	a list of definitions. Each definition has a name (<code><dt></code> , definition term) and a value (<code><dd></code> , definition description)
<code><figure></code>	a sectioning element that can be used to annotate illustrations, diagrams, photos, code listings, etc. May include <code></code> or other content. May include <code><figcaption></code>
<code><main></code>	represents the dominant contents of the document
<code><div></code>	the <code><div></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 (“inline”)**, 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
 - `anchor`

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 ²⁴ kg)	Diameter (km)	Density (kg/m ³)	Gravity (m/s ²)	Length of day (hours)	Distance from Sun (10 ⁶ 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 remains controversial .

HTML attributes

- HTML elements may have attributes

`<element attribute="value" ... > ... </element>`

- Attributes can change the element behavior/appearance
- Example
 - `<input type="text" ...>`
 - `<input type="number" ...>`

Attributes applicable to 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 finding/matching portions 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

Attributes applicable to 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

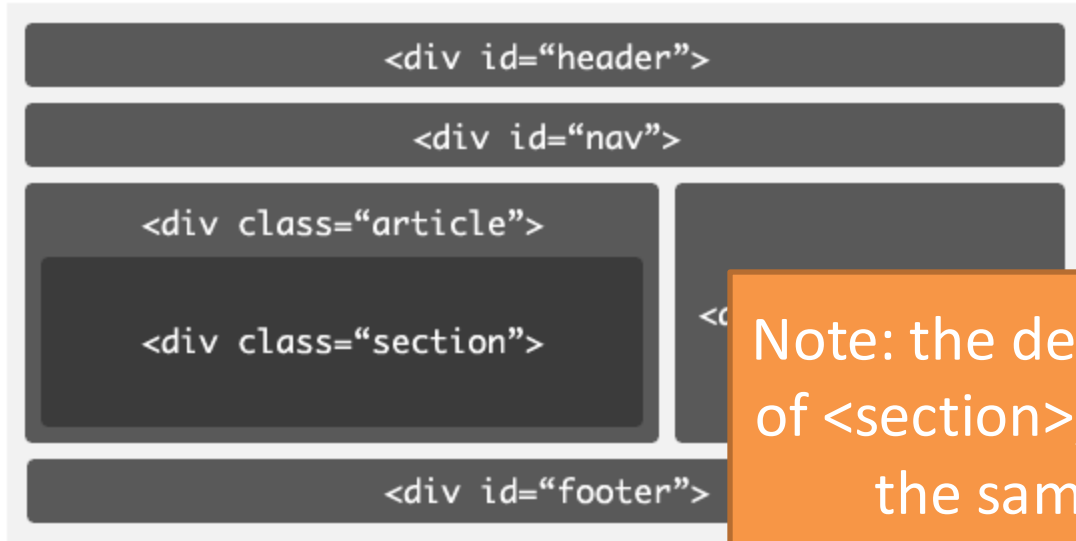
`<div> ... </div>`

- **Block-level** container element
- No predefined semantics
- Invisible layout
- Used to mark-up blocks or groups of blocks, with specific classes
- Shown as `display:block` by default, i.e.
 - Full-width, starts on a new line
 - top-to-bottom layout

` ... `

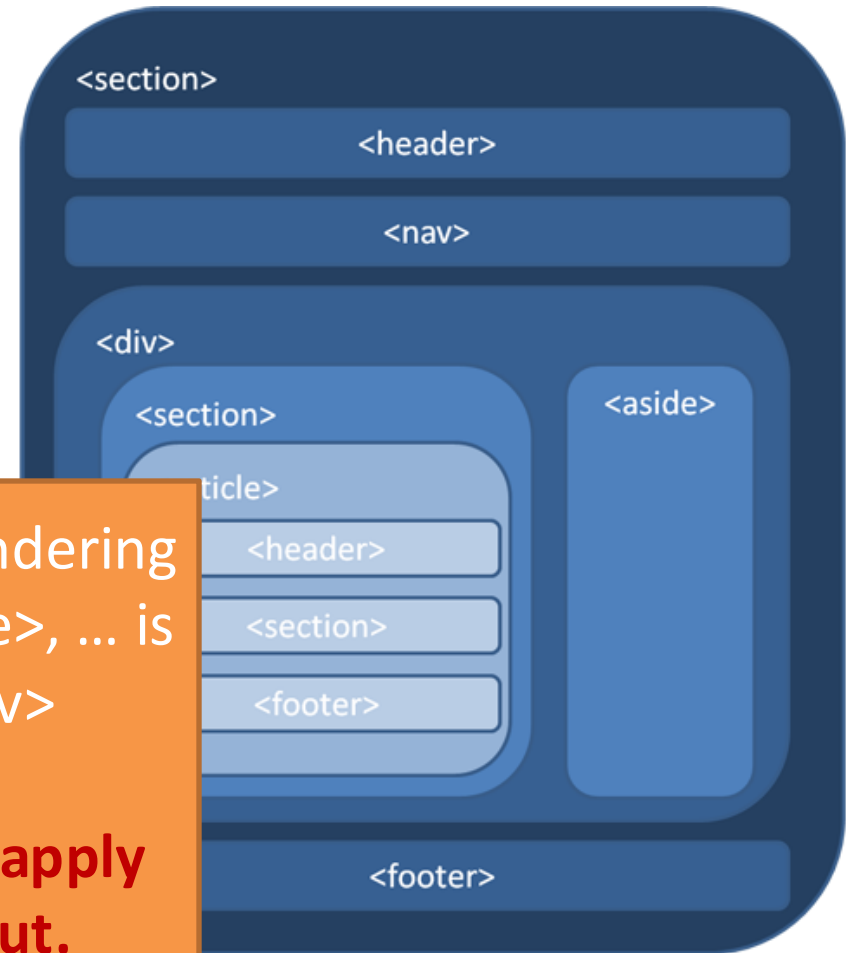
- **Inline-level** container element
- No predefined semantics
- Invisible layout
- Used to mark-up parts of a paragraph, with specific classes
- Shown as `display:inline` by default, i.e.
 - Does not start a new line, occupy only the necessary space
 - Left-to-right layout, wraps around at the end of the line

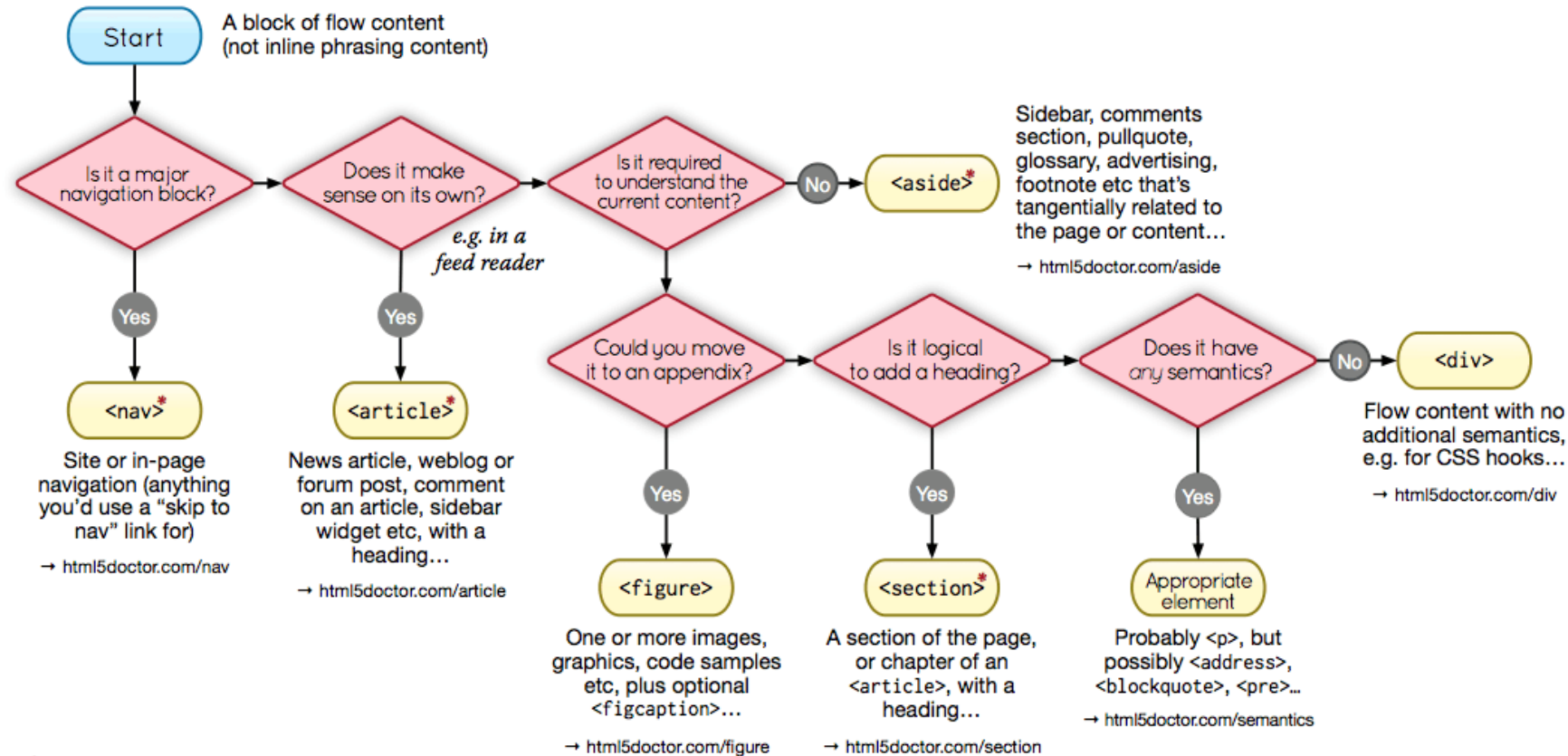
Opaque vs. Semantic elements



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

Only CSS rules will apply the desired layout.





* 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

2011-07-22 v1.5
For more information:
www.html5doctor.com/semantics

HTML5 has the best DOCTYPE ever.

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

Only one single stylesheet is used (media types are declared inside stylesheet), and only served to good browsers. IE 6 is served a universal stylesheet. Below IE 6 gets nothing.

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

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

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

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

Lists are marked up as lists, depending on the needs of the list: unordered, ordered, and the underused definition list.

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

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

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

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

```
<!DOCTYPE HTML>
<html>

<head>
<meta http-equiv="Content-Type" content="text/html; charset=UTF-8" />

<title>Portfolio | Chris Coyler</title>

<!--[[[ IE]]]><!-->
<link rel="stylesheet" type="text/css" href="/css/main.css" />
<!--<[[[endif]]><!-->

<!--[[[ gte IE 7]]>
<link rel="stylesheet" type="text/css" href="/css/main.css" media="screen, projection" />
<[[[endif]]><!-->

<!--[[[ IE 6]]>
<link rel="stylesheet" type="text/css" href="http://universal-ie6-css.googlecode.com/files/ie6.0.3.css" media="screen, projection" />
<[[[endif]]><!-->
</head>

<body id="home">

<header>
<a id="logo" href="/">Site Title</a>
<div id="slogan">web craftsman, blogger, author, speaker</div>
<nav>
<?php include("inc/main-menu.php"); ?>
</nav>
</header>

<section class="container">

<article>
<h1>Hipsters</h1>



You can&#8217;t dress up as a hipster for Halloween. Their attire is already so bizarre that there&#8217;s an exaggeration of it that looks like a costume. It would just look like you are another hipster about to read a poem about reading poems.</p>

<h2>Secondary Title</h2>
<p>Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas.</p>
</article>

<article>
<!-- Additional Article -->
</article>
</section>

<aside>
<h3>My Major Projects</h3>
<ul>
<dt><a href="http://aremysitesup.com">Are My Sites Up?</a></dt>
<dd>Monitor your sites</dd>

<dt><a href="http://css-tricks.com">CSS-Tricks</a></dt>
<dd>A web design community</dd>

<dt><a href="http://dignop.com">Digging Into WordPress</a></dt>
<dd>Learn about WordPress</dd>
</ul>
</aside>

<footer class="container">
<h4>People I Enjoy</h4>
<ul class="col">
<li><a href="http://fastfoodreviewed.com">Jesse Lynch</a></li>
<li><a href="http://jeffcampana.com">Jeff Campana</a></li>
<li><a href="http://perishablepress.com">Jeff Starr</a></li>
</ul>
<ul class="col">
<li><a href="http://davidwalsh.name">David Walsh</a></li>
<li><a href="http://thestrategicretreat.com">Jeff Pennam</a></li>
<li><a href="http://http://shiftedfrequency.com">Richard Felix Jr.</a></li>
</ul>

<h4>Sandwiches</h4>
<ul class="col container" id="sandwich-list">
<li><a href="http://jimjohns.com">Jimmy Johns</a></li>
<li><a href="http://subway.com">Subway</a></li>
<li><a href="http://potbelly.com">Potbelly</a></li>
</ul>

&copy;2007 -<?php echo date("Y"); ?> Chris Coyler
</footer>

<script type="text/javascript" src="http://ajax.googleapis.com/ajax/libs/jquery/1.3.2/jquery.min.js?ver=1.3.2"></script>
<script type="text/javascript" src="/js/main.js"></script>

<!-- Google Analytics Code -->
<?php include_once("inc/analytics.php"); ?>

</body>

</html>
```

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

Site resources use relative file paths for efficiency.
Content file paths are absolute, assuming content is syndicated.

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.

If it's a special character, it's encoded.

Things that need to be dynamic, are dynamic.

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


Nothing on the page applies styling or even implies what the styling might be. Everything on the page is either a required site resource, content, or describing

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



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

HTML Validator

**Markup Validation Service**
Check the markup (HTML, XHTML, ...) of Web documents

Validate by URI

Validate by File Upload

Validate by Direct Input

Validate by URI

Validate a document online:

Address:

▼ More Options

Character Encoding

(detect automatically)

☐ Only if missing

Document Type

(detect automatically)

☐ Only if missing

☒ List Messages Sequentially

☐ Group Error Messages by Type

☐ Show Source

☐ Clean up Markup with HTML-Tidy

☐ Show Outline

☐ Validate error pages

☐ Verbose Output

Check

This validator checks the [markup validity](#) of Web documents in HTML, XHTML, SMIL, MathML, etc. If you wish to validate specific content such as [RSS/Atom feeds](#) or [CSS stylesheets](#), [MobileOK content](#), or to [find broken links](#), there are [other validators and tools](#) available. As an alternative you can also try our [non-DTD-based validator](#).



Interested in "developing" your developer skills? In W3C's hands-on Professional Certificate Program, learn how to code the right way by creating Web sites and apps that use the latest Web standards. [Find out more!](#)

[Donate](#) and help us build better tools for a better web.

[Home](#) [About...](#) [News](#) [Docs](#) [Help & FAQ](#) [Feedback](#) [Contribute](#)



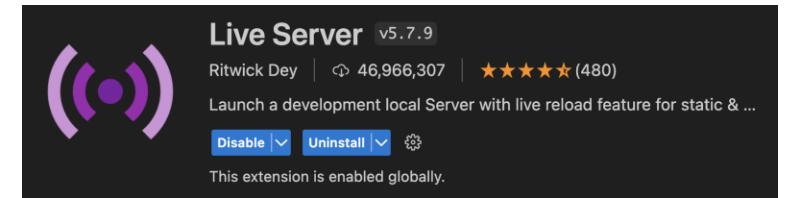
This service runs the W3C Markup Validator, [v1.3+bg](#).
COPYRIGHT © 1994-2013 W3C® (MIT, ERCIM, KEIO, BEIHANG), ALL RIGHTS RESERVED. W3C LIABILITY, TRADEMARK, DOCUMENT USE AND SOFTWARE LICENSING RULES APPLY. YOUR INTERACTIONS WITH THIS SITE ARE IN ACCORDANCE WITH OUR PUBLIC AND MEMBER PRIVACY STATEMENTS.



<https://validator.w3.org/>



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> (includes “Introduction to 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/>

