

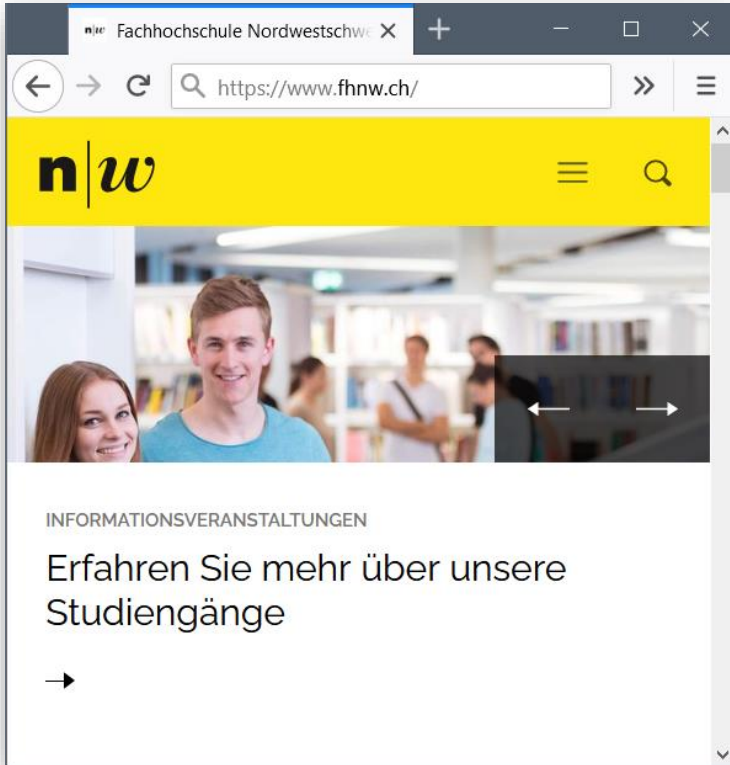
Web Engineering

HTML

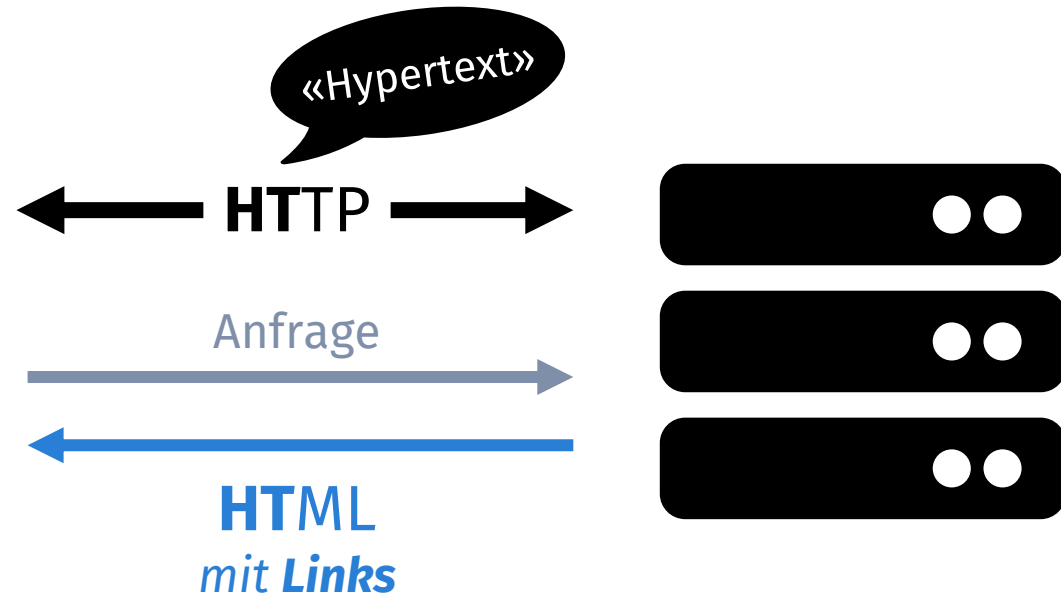
Adrian Herzog

(basierend auf der Arbeit von Michael Faes, Michael Heinrichs & Prof. Dierk König)

Das World Wide Web



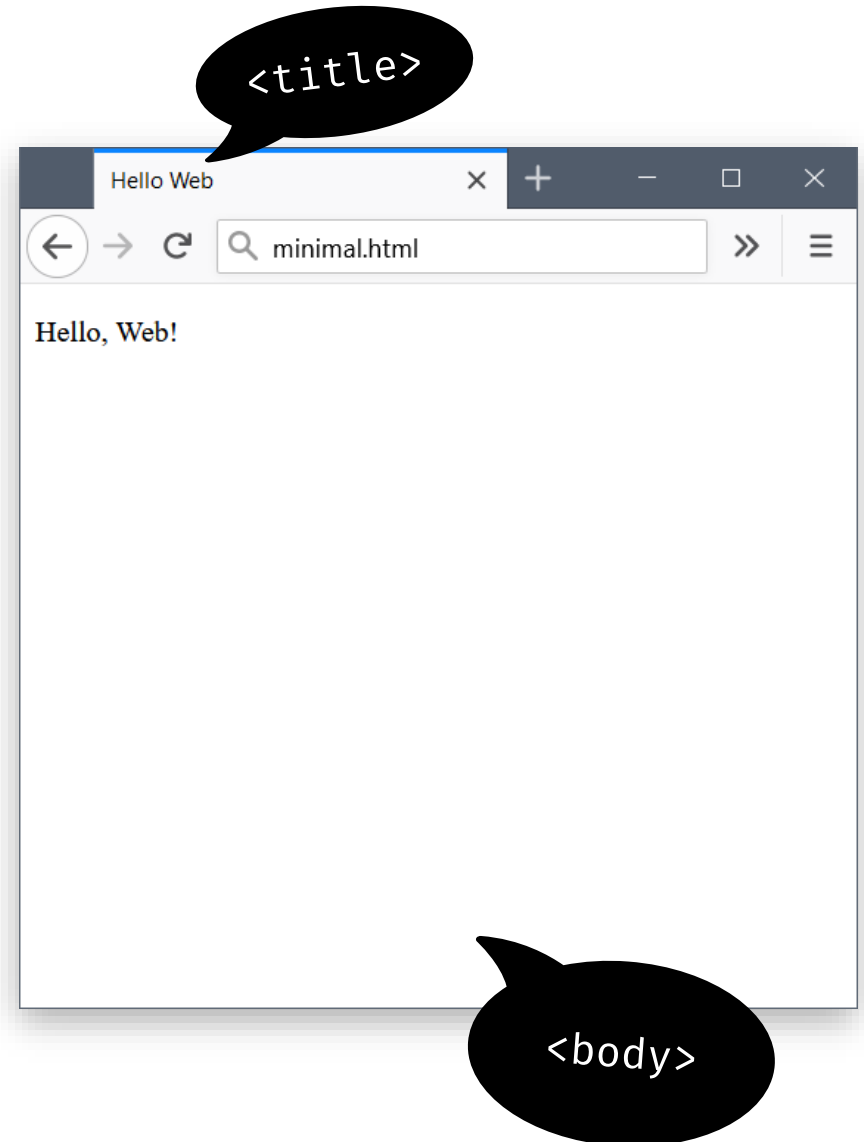
Web-Client
(z. B. Browser)



Web-Server

HTML Grundgerüst

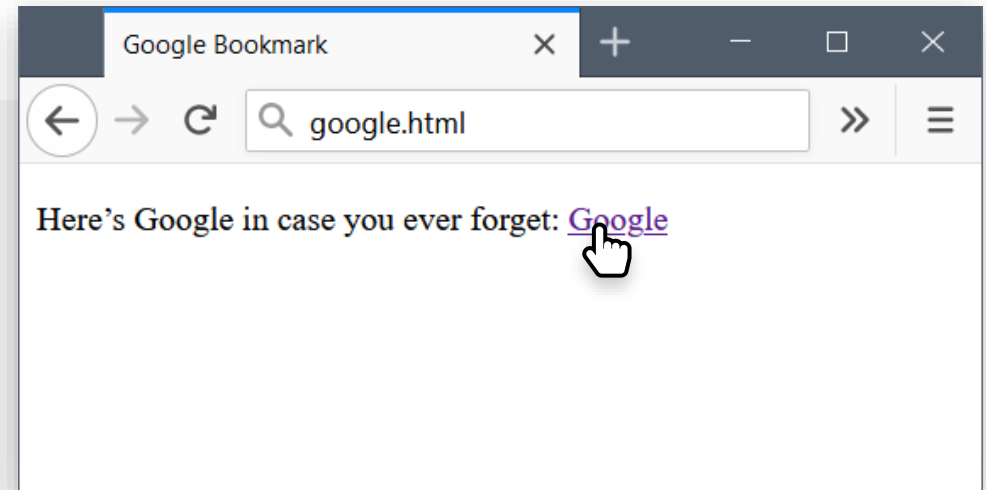
```
<!DOCTYPE html>
<html lang="en">
  <head>
    <title>Hello Web</title>
  </head>
  <body>
    <p>Hello, Web!</p>
  </body>
</html>
```



HT*ML*: eine «*Auszeichnungssprache*»

Verknüpft Text mittels `<tags>` mit «Eigenschaften, Zugehörigkeiten und Darstellungsformen»

```
<!DOCTYPE html>
<html lang="en">
  <head>
    <title>Google Bookmark</title>
    <meta charset="utf-8">
  </head>
  <body>
    <p>
      Here's Google in case you ever forget:
      <a href="https://www.google.com">Google</a>
    </p>
  </body>
</html>
```



Primär Bedeutung,
nicht Präsentation!

Ein paar Tags

Überschriften

```
<h1>Part I – Silent Spring</h1>
<h2>1. The Madness Years</h2>
```

Part I – Silent Spring

1. The Madness Years

Listen

```
<ol>
  <li>Make it work</li>
  <li>Make it right</li>
  <li>Make it fast</li>
</ol>
```

1. Make it work
2. Make it right
3. Make it fast

Haben alle eine Bedeutung!

Tabellen

```
<table>
  <tr>
    <td>Merkur</td>
    <td>0.4 AU</td>
  </tr>
  <tr>
    <td>Venus</td>
    <td>0.7 AU</td>
  </tr>
</table>
```

Merkur	0.4 AU
Venus	0.7 AU

Bilder

```

```



Was kann man mit HTML alles machen?

What Does HTML5 Do?

Key features of the next Web programming standard.



Aufbau von HTML

*Dokumenttyp-
Deklaration*

```
<!DOCTYPE html>
```

```
<html lang="en" >
```

Attribut

Element

```
<head>
```

```
<title>
```

```
Google Bookmark
```

```
</title>
```

Element

Start-Tag

Text

End-Tag

```
</head>
```

```
<body>...</body>
```

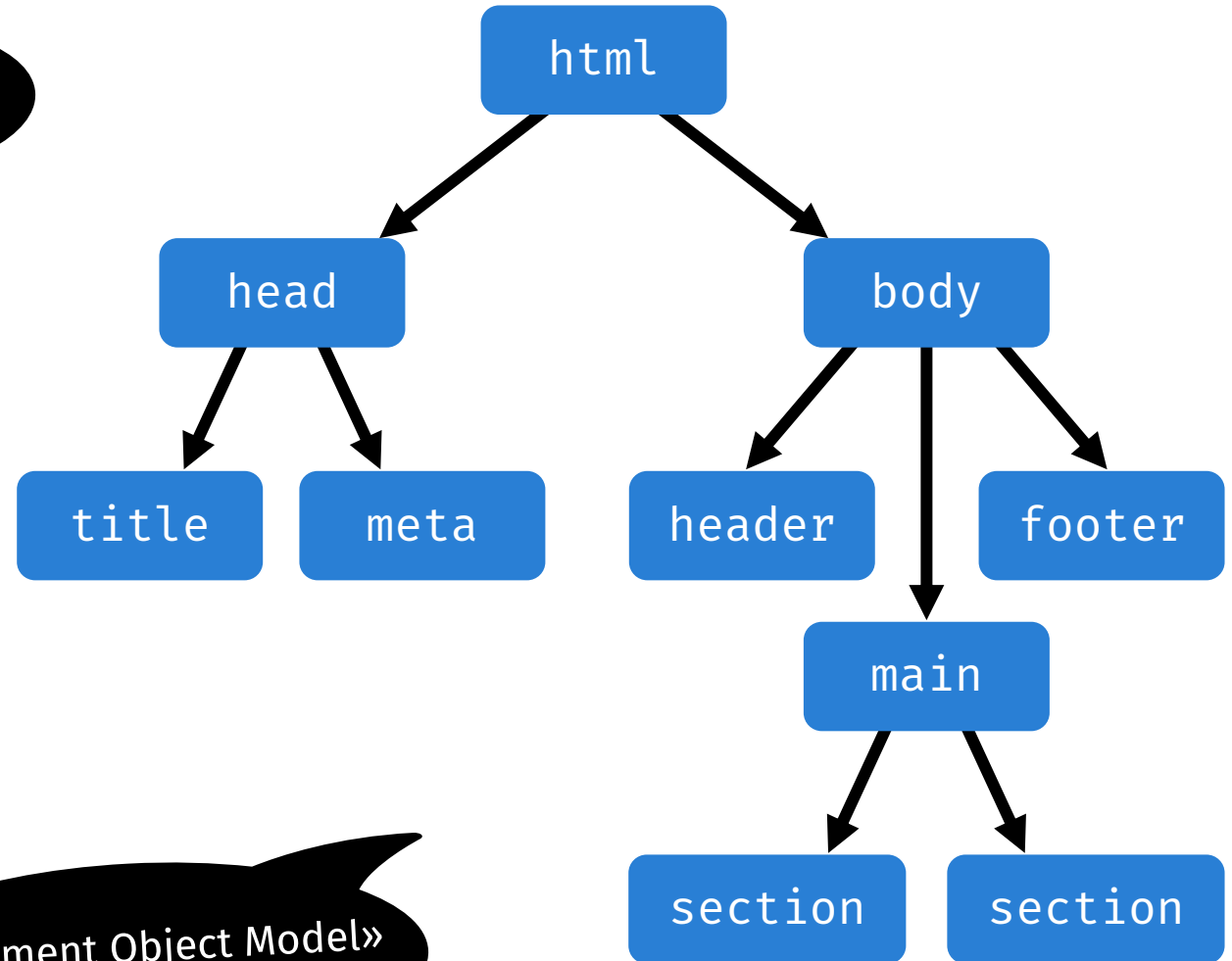
```
</html>
```

HTML-Dokument ist ein Baum

```
<!DOCTYPE html>  
<html lang="en">  
  <head>  
    <title>...</title>  
    <meta>...</meta>  
  </head>  
  <body>  
    <header>...</header>  
    <main>  
      <section>...</section>  
      <section>...</section>  
    </main>  
    <footer>...</footer>  
  </body>  
</html>
```


(nicht Teil des Dokuments)

«Document Object Model»
(DOM)



HTML-Syntax: Details

Diverse Zeichen haben in HTML spezielle Bedeutung und müssen escapet werden:



Zeichen	Bedeutung	Escape	Wo
<	Tag-Beginn	<	in Text
>	Tag-Ende	>	in Text
"	Attribut-Beginn und Ende	"	in Attributen mit Trennzeichen "
'	Attribute-Beginn und Ende	'	in Attributen mit Trennzeichen '
&	Beginn von Entities	&	es ist kompliziert...

Empfehlung: Mindestens < > und & überall durch Entities ersetzen.

Syntax-Fehler: Beispiele

Welche der folgenden HTML-Ausschnitte sind gültig?

```
<p>Warum ist das <em class="red">rot</em>?</p>
```



```
<p>Er sagte, er sei "krank".</p>
```



```
<p>Es gilt: a < 0.</p>
```



```
<span title='Name des "Buchs"'>Mein Leben als Dino</span>
```



```
Übersetzung: <em title="oder auch "Hi"">Hallo</em>
```



```
<p>Die Hose ist von H&M; das Shirt war ein Geschenk.</p>
```



Gültigkeit

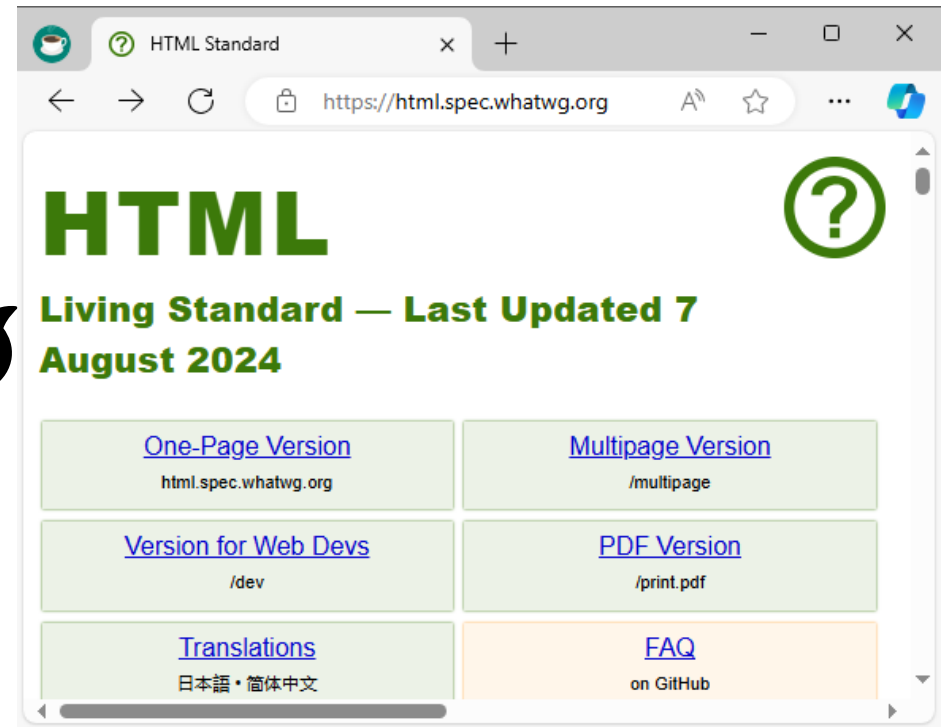
HTML-Sprache ist durch «Standard» definiert

Weitere Regeln für Syntax:

- Elemente müssen geöffnet und geschlossen werden (ausser `
`, `<hr>`, ``, etc.)
- Nicht überlappen: `This is illegal`

Regeln für *Verwendung* von Elementen/Attributen, z. B.:

- Element `` nur innerhalb von ``, `` oder `<menu>`
- Jedes `` muss `alt`-Attribut enthalten



HTML ist komisch...

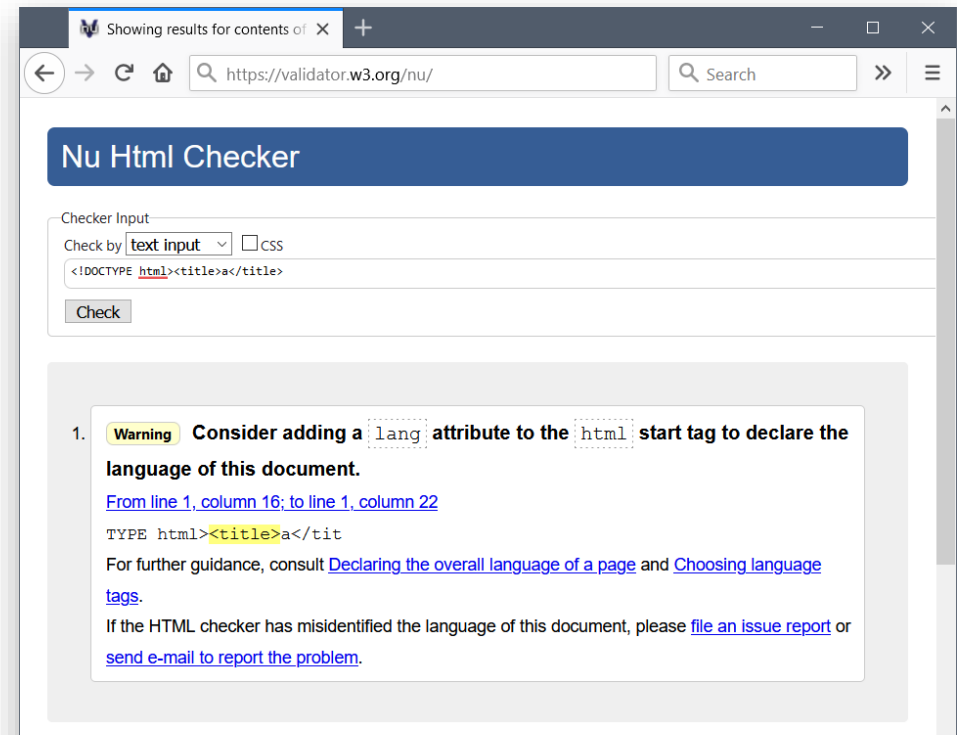
Viele hässliche Dinge sind trotzdem erlaubt:

Elemente müssen geöffnet und geschlossen werden

- Viele End-Tags sind optional: `<p>Paragraph 1` `<p>Paragraph 2`
- Manche Start-Tags sind auch optional! `<html>`, `<head>`, `<body>`

Minimales, valides
HTML-Dokument:

```
<!DOCTYPE html><title>a</title>
```



HTML ist komisch...

(Elemente dürfen nicht überlappen)

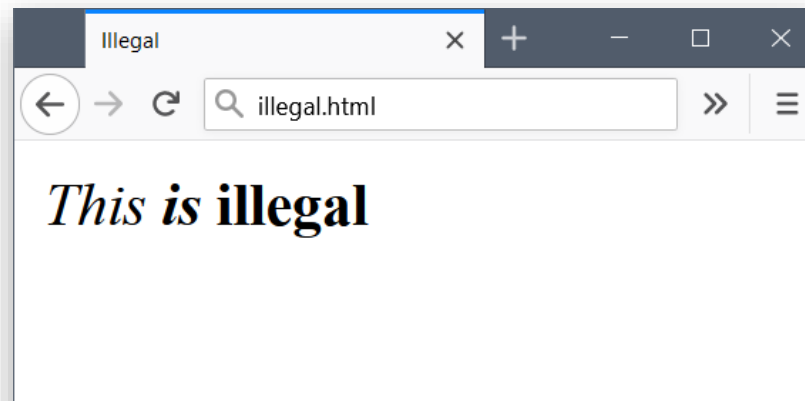
- Ja, aber Spezifikation definiert auch, wie Syntax-Fehler am besten behandelt werden können:

```
<em>This <strong>is</em> illegal</strong>
```



```
<em>This <strong>is</strong></em> <strong>illegal</strong>
```

- Die meisten Browser haben keine Probleme damit:



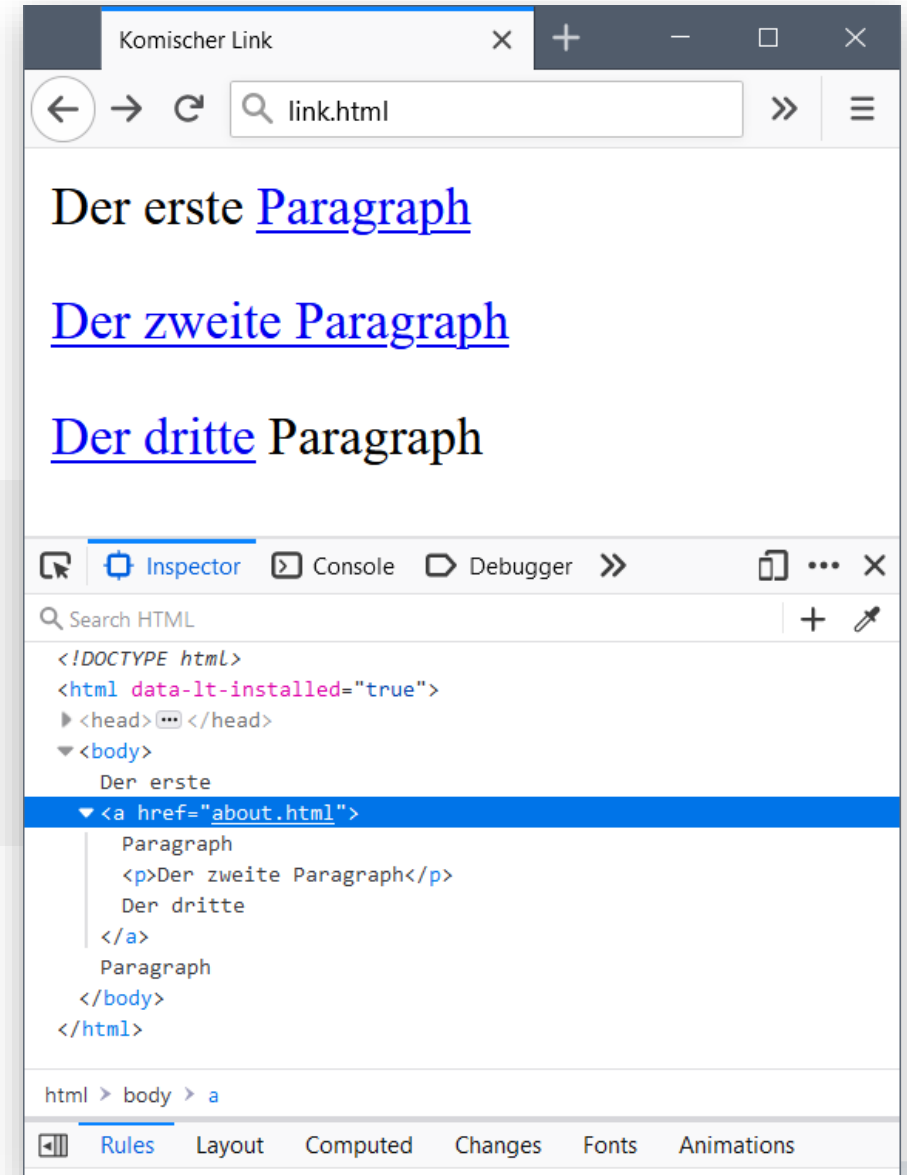
HTML ist komisch...

(Elemente dürfen nicht überlappen)

- Ja, aber Elemente in Dokument stimmen nicht 100% mit logischen Einheiten (Paragraphen) überein:

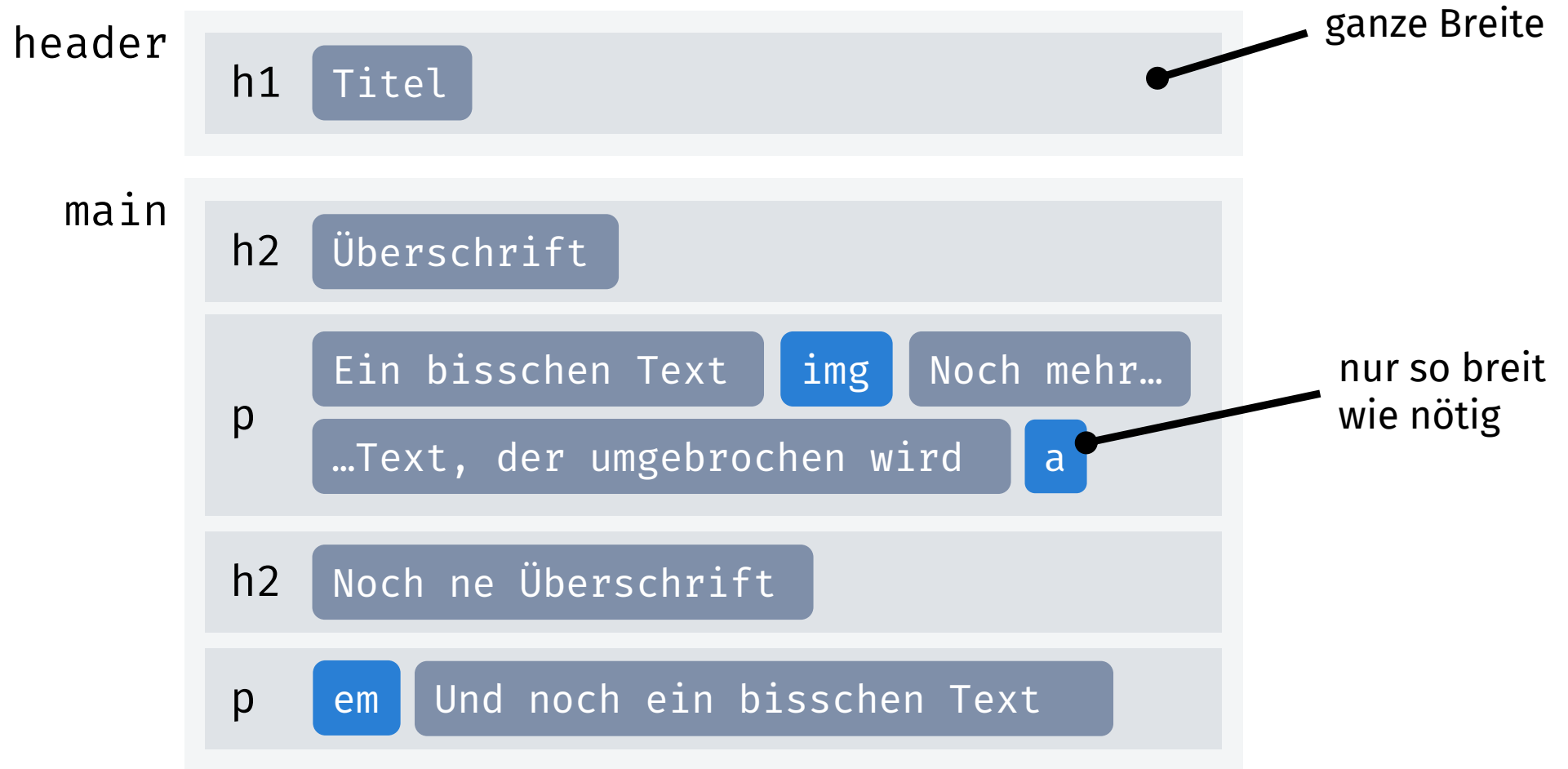
```
<!DOCTYPE html>
<title>Komischer Link</title>
Der erste <a href="about.html">Paragraph
<p>Der zweite Paragraph</p>
Der dritte</a> Paragraph
```

Es lohnt sich trotzdem,
möglichst valides HTML zu
schreiben

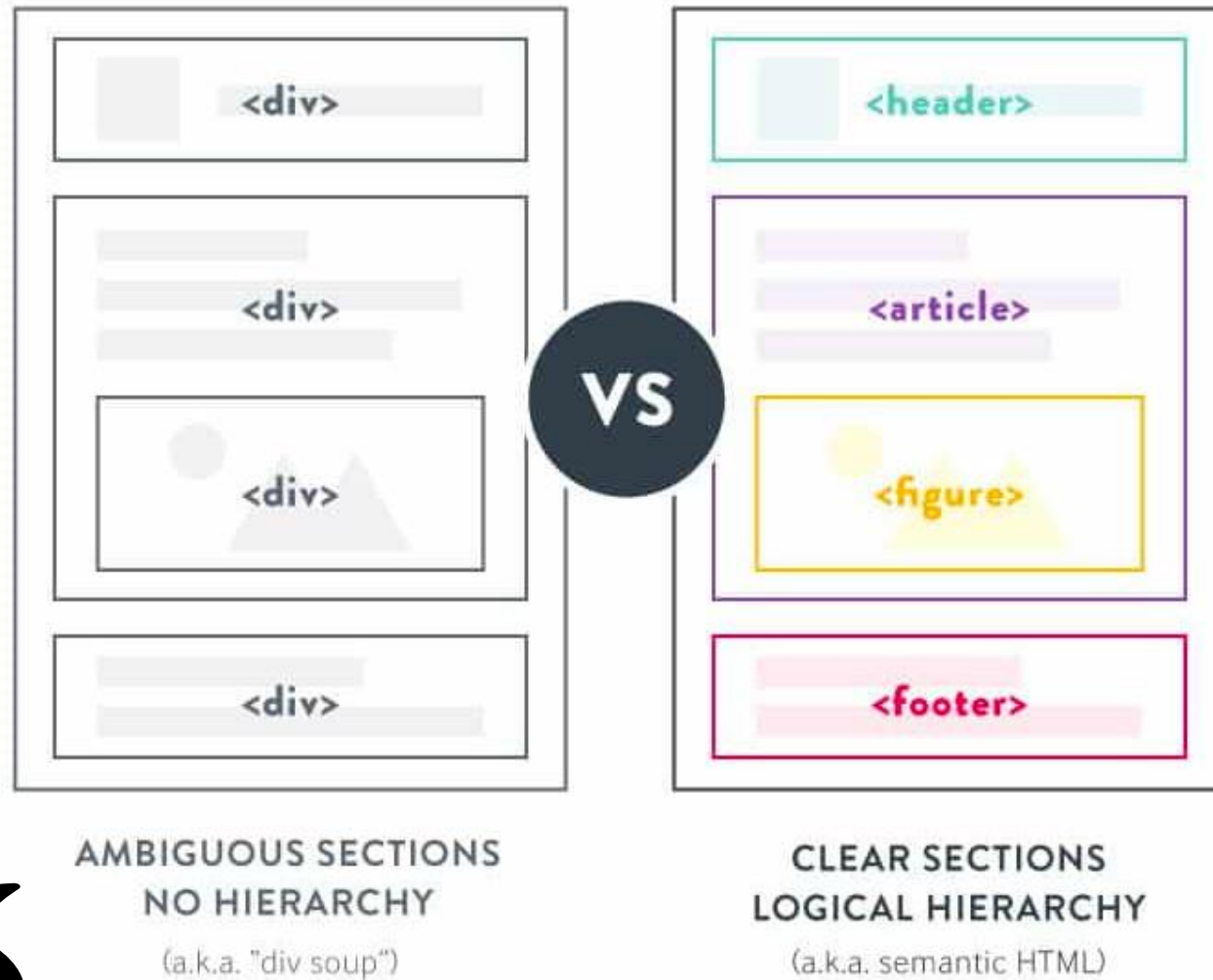


Block- und Inline-Elemente

Zwei grobe Hauptkategorien von Elementen innerhalb von `<body>`



«Semantische Elemente»



Bedeutung,
nicht Präsentation!

Semantische Elemente: Beispiele

`<article>`

`<aside>`

`<details>`

`<figcaption>`

`<figure>`

`<footer>`

`<header>`

`<main>`

`<mark>`

`<nav>`

`<section>`

`<summary>`

`<time>`

Aber auch:

`<h1>`, `<h2>`, ...

``

``

`<table>`

``

``

Warum nicht einfach alles mit `<div>`?

- Browser sind nicht die einzigen Programme, die HTML verarbeiten!
- Suchmaschinen, Crawler, **Screenreader**, ...

<https://developer.mozilla.org/en-US/docs/Web/HTML/Element>

Die wichtigsten Elemente

- Doctype: `<!DOCTYPE html>`
- Grundstruktur: `<html>`, `<head>`, `<body>`
- Links: `<a>`
- Überschriften und Abschnitte: `<h1>`, `<h2>`, `<h3>`, ..., `<p>`
- Blöcke / Inline-Elemente / Umbruch: `<div>`, ``, `
`
- Listen: ``, ``, ``
- Tabellen: `<table>`, `<tr>`, `<th>`, `<td>`
- Textformatierung und Hervorhebung: ``, ``, ``, `<i>`, `<u>`
- Bilder: ``, `<svg>`
- Formulare: `<form>`, `<input>`, `<button>`, `<textarea>`, `<select>`, `<option>`, `<label>`

Auch wichtig

- JavaScript: `<script>`, `<noscript>`
- CSS: `<style>`
- Multimedia: `<audio>`, `<video>`
- Leinwand für Grafikausgabe: `<canvas>`
- Seite einbetten: `<iframe>`
- Zusatzinfos im Head: `<meta>`
- Kommentare: `<!-- dies wird nicht dargestellt -->`

strong / em / b

Tag	Default styling	Meaning
	Bold	No special meaning, just used to draw attention to the text.
	Bold	Content has greater importance.
	Italics	Emphasise when reading out loud.

See [MDN](#)

Fragen?

