



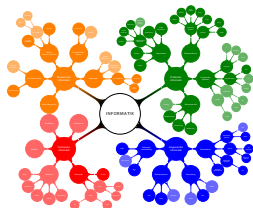
UNIVERSITÄT ZU LÜBECK
INSTITUT FÜR
THEORETISCHE INFORMATIK

Kapitel 7

Hypertext Markup Language

Die Sprache des WWW

Vorlesung Einführung in die Informatik 1 vom 14. November 2013 von Till Tantau



Lernziele von Kapitel 7

1. Die Grundsyntax von HTML kennen
2. Die Grundsyntax von XML kennen
3. Eigene HTML-Seiten erstellen können

Gliederung von Kapitel 7

7.1 HTML

7.1.1 Einführung

7.1.2 Aufbau von HTML-Seiten

7.1.3 Aufbau von Tags

7.2 XML

7.2.1 Einführung

7.2.2 Vergleich mit HTML

7.2.3 Wohlgeformte Texte

Wiederholung: Wohin Seitenbeschreibungssprachen tendieren.

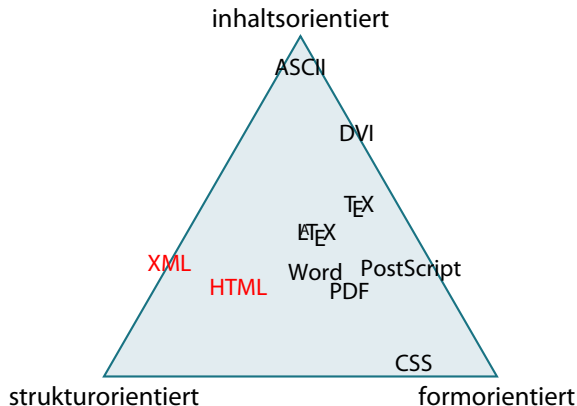
7.1 HTML

Einführung
Aufbau von HTML-Seiten
Aufbau von Tags

7.2 XML

Einführung
Vergleich mit HTML
Wohlgeformte Texte

7-4



7.1 HTML

- Einführung
- Aufbau von HTML-Seiten
- Aufbau von Tags

7.2 XML

- Einführung
- Vergleich mit HTML
- Wohlgeformte Texte

- Die Hypertext-Markup-Language (HTML) ist eine *Seitenbeschreibungssprache*, in der die meisten Webseiten verfasst sind.
- Sie ist *keine Programmiersprache*, HTML beschreibt primär *Inhalt und Struktur* beliebiger Textdokumente, die nicht unbedingt im Netz stehen müssen.
- Die *Form* kann kaum festgelegt werden.
- Wie der Name schon sagt, erlaubt es HTML, *Verweise* auf Bilder oder *Hyperlinks* einfach zu notieren.

7.1 HTML

► Einführung

Aufbau von HTML-Seiten

Aufbau von Tags

7.2 XML

Einführung

Vergleich mit HTML

Wohlgeformte Texte

Originalseite in HTML

```
<p>  
Dies ist  
ein <i>Text</i>.  
</p>
```

Anzeige des Browser

Dies ist ein *Text*.

7.1 HTML

Einführung

- Aufbau von HTML-Seiten
- Aufbau von Tags

7.2 XML

Einführung

Vergleich mit HTML
Wohlgeformte Texte

- HTML-Seiten bestehen aus *Tags* und *normalem Text*.
- Tags sind Ausdrücke der Form `<tag>` und `</tag>`. Die erste Form von Tags heißt *öffnendes Tag*, die zweite Form heißt *schließendes Tag*; sie funktionieren wie öffnende und schließende Klammern.
- Alles, was von einem öffnenden und dem zugehörigen schließenden Tag umschlossen wird, nennt man ein *Element*.
- Je nach Art des Elements (also je nach Art des Tags) dürfen innerhalb des Elements wieder Tags und normaler Text stehen.

7.1 HTML

Einführung

- Aufbau von HTML-Seiten
- Aufbau von Tags

7.2 XML

Einführung

Vergleich mit HTML
Wohlgeformte Texte

Zur Übung

Wie viele Tags und wie viele Elemente enthält die folgende Seite?

```
<p>  
  Dies ist  
  ein <i>Text</i>.<br></br>  
  <ul>  
    <li> Foo </li>  
    <li> Bar </li>  
  </ul>  
</p>
```

7-8

7.1 HTML

Einführung

- Aufbau von HTML-Seiten
- Aufbau von Tags

7.2 XML

Einführung

- Vergleich mit HTML
- Wohlgeformte Texte

- Das äußerste Tag einer HTML-Seite ist immer das Paar `<html>` und `</html>`.
- Innerhalb dieses äußersten Tags kommt erst das Paar `<head>` und `</head>`, dann das Paar `<body>` und `</body>`.
- Innerhalb von `<head>` kann man Titelinformation angeben mittels weiterer Tags wie `<title>`.
- Innerhalb von `<body>` kommt der eigentliche Text der Seite, der später angezeigt werden soll.

7.1 HTML

Einführung

- Aufbau von HTML-Seiten
- Aufbau von Tags

7.2 XML

Einführung

- Vergleich mit HTML
- Wohlgeformte Texte

- Zur Gliederung des Haupttextes stehen weitere Tags zur Verfügung.
- Überschriften kann man mit den Tags `<h1>` bis `<h6>` einfügen.
- Absätze umschließt man mit den Tags `<p>` und `</p>`.
- Listen umschließt man mit `` für unsortierte Listen oder `` für sortierte Listen.
- Einzelne Listenelemente umschließt man mit `` (list item).
- Tabellen umschließt man mit `<table>`, Tabellenzeilen mit `<tr>` (table row) und Zellen mit `<td>` (table data).
- `<div>` erzeugt »generische« Unterteilungen, sie werden nur in Verbindung mit so genannten Style-Sheets gebraucht.

- ▶ Die bisherigen Gliederungselemente konnten nicht mitten in einem Absatz vorkommen. Man nennt sie *Blockelemente*.
- ▶ Es gibt aber auch Tags und Elemente, die mitten in einem Absatz vorkommen dürfen wie
 - ▶ `<i>` (italics) setzt Text kursiv,
 - ▶ `` (bold) setzt Text fett,
 - ▶ `` hebt Text besonders hervor,
 - ▶ `` macht selber nichts, wird aber für Style-Sheets gebraucht.
- ▶ Solche Elemente nennt man *Inlineelemente*.

Merkregel

Inlineelemente können mitten in einer Zeile beginnen und auf einer anderen Zeile enden. Blockelemente hingegen nicht.

7.1 HTML

Einführung

- ▶ Aufbau von HTML-Seiten
- Aufbau von Tags

7.2 XML

Einführung

Vergleich mit HTML
Wohlgeformte Texte

```
<html>
  <head>
    <title>
      Dies ist eine Beispielseite.
    </title>
  </head>
  <body>
    <h1> Haupt&uuml;berschritt </h1>
    <p>
      Hier ist <strong>normaler</strong> Text. Man
      beachte, dass f&uuml;r den Umlaut &uuml; eine
      merkw&uuml;rdige Zeichenfolge benutzt wurde.
    </p>
    <h2> Unterabschnitt </h1>
    <p>
      Blah, blah.
    </body>
  </html>
```

7.1 HTML

Einführung

- Aufbau von HTML-Seiten
- Aufbau von Tags

7.2 XML

Einführung

Vergleich mit HTML
Wohlgeformte Texte

7-12

Zur Übung

Finden Sie zwei Fehler.

7.1 HTML

Einführung

Aufbau von HTML-Seiten

► Aufbau von Tags

7.2 XML

Einführung

Vergleich mit HTML

Wohlgeformte Texte

- Tags können *Parameter* bekommen.
- Diese werden *innerhalb des öffnenden Tags* in der Form
`parameter="wert"` angegeben,
- mehrere Parameter werden durch Leerzeichen getrennt.

Beispiel (Ein Verweis auf eine Webseite)

Auf `dieser Seite` ist
ein Wiki zu finden.

7.1 HTML

Einführung

Aufbau von HTML-Seiten

► Aufbau von Tags

7.2 XML

Einführung

Vergleich mit HTML

Wohlgeformte Texte

Zur Übung

Geben Sie den Originalquelltext einer HTML-Seite an, die in einer Liste Links auf die deutsche, die englische und die französische Wikipedia enthält.

7.1 HTML

Einführung

Aufbau von HTML-Seiten

► Aufbau von Tags

7.2 XML

Einführung

Vergleich mit HTML

Wohlgeformte Texte

- Oft gibt es Situationen, in denen ein Tag keinen »Inhalt« hat.
- So erzeugt das Tag-Paar `
` und `</br>` einen Zeilenumbruch (`br` steht für break).
- Innerhalb dieses Paares kann man aber nicht sinnvoll etwas schreiben.
- Statt `
</br>` darf man kürzer schreiben `
`.
- Allgemein bedeutet `<tag/>` dasselbe wie `<tag></tag>`.

Beispiel (Eine Stelle auf die man verweisen kann)

`` Im Anfang war das Wort...

- ▶ Die Liste der Tags hängt von der HTML-Version ab.
- ▶ In der aktuellen Version HTML5 gibt es eine feste Liste, mit der man Dinge beschreiben kann wie:
 - ▶ Überschriften,
 - ▶ Angabe von Textarten wie Zitate, Programmtexte,
 - ▶ Listen,
 - ▶ Tabellen,
 - ▶ Verweise und eingebettete Objekte wie Graphiken.
- ▶ Eine ausführliche Liste aller Tags mit Erklärungen und Tutorien findet sich auf der Webseite <http://de.selfhtml.org/>.
- ▶ Es ist *nicht möglich*, neue Tags hinzuzufügen für Dinge wie
 - ▶ Atome,
 - ▶ Gene,
 - ▶ Moleküle,
 - ▶ Spezies.

7.1 HTML

Einführung

Aufbau von HTML-Seiten

▶ Aufbau von Tags

7.2 XML

Einführung

Vergleich mit HTML

Wohlgeformte Texte

7.1 HTML

Einführung
Aufbau von HTML-Seiten
Aufbau von Tags

7.2 XML

► Einführung
Vergleich mit HTML
Wohlgeformte Texte

- Die Beschreibung der Struktur und des Inhalts von Dokumenten mit HTML ist, wenn sie richtig angewendet wird, *sowohl für Menschen wie für Maschinen gut lesbar*.
- Ein großes Problem ist jedoch, dass *sich HTML nicht erweitern lässt*.
- Die *eXtensible Markup Language (XML)* löst dieses Problem, indem sie es erlaubt, *neue Tags zu definieren*.
- Damit ist aber noch nicht gesagt, wie diese Tags dann *angezeigt werden sollen*, was ein Style-Sheet regelt.
- Man kann XML aber auch einfach nur benutzen, um die Struktur von Dokumenten deutlich zu machen.

Ein XML-Dokument muss immer mit folgender Zeile beginnen:

```
<?xml version="1.0"?>
```

In HTML

Ein Kaiserschmarrn besteht aus:

```
<ol>  
  <li> 150g Mehl </li>  
  <li> 1/8l Milch </li>  
  <li> 3 Eier (getrennt) </li>  
  <li> Puderzucker und eine Prise Salz </li>  
</ol>
```

In XML

Ein Kaiserschmarrn besteht aus:

```
<rezept>  
  <zutat> 150g Mehl </zutat>  
  <zutat> 1/8l Milch </zutat>  
  <zutat> 3 Eier (getrennt) </zutat>  
  <zutat> Puderzucker und eine Prise Salz </zutat>  
</rezept>
```

7.1 HTML

Einführung
Aufbau von HTML-Seiten
Aufbau von Tags

7.2 XML

Einführung
► Vergleich mit HTML
Wohlgeformte Texte

7.1 HTML

Einführung
Aufbau von HTML-Seiten
Aufbau von Tags

7.2 XML

Einführung
► Vergleich mit HTML
Wohlgeformte Texte

Zur Übung

Geben Sie die Beschreibung eines Atoms in XML-Syntax an. (Es gibt nicht die »richtige« Lösung, es kommt darauf an, was Ihnen wichtig erscheint und auf den Kontext.)

7.1 HTML

Einführung
Aufbau von HTML-Seiten
Aufbau von Tags

7.2 XML

Einführung
Vergleich mit HTML
► Wohlgeformte Texte

Definition

Ein XML-Text heißt *wohlgeformt*, wenn alle Tags korrekt geschachtelt sind.

Definition

Eine *Document Type Description (DTD)* gibt eine Menge von erlaubten Tags und erlaubten Parametern an. (Beispiel: Es gibt eine DTD für HTML.)

Definition

Ein XML-Text heißt *gültig in Bezug auf eine DTD*, wenn die Tags wie von der DTD vorgeschrieben benutzt werden.

7.1 HTML

Einführung
Aufbau von HTML-Seiten
Aufbau von Tags

7.2 XML

Einführung
Vergleich mit HTML
► Wohlgeformte Texte

1. HTML-Texte beschreiben Inhalt und Struktur, XML-Texte ebenso.
2. HTML- und XML-Texte bestehen aus *korrekt geschachtelten Tags*.

7-21



Götz Bürkle et al.

<http://de.selfhtml.org/>, Zugriff Mai 2007.