

# LaTeX und HTML -- Verwandtschaften

(Günter Partosch, HRZ Gießen)

E-Mail: [Gunter.Partosch@hrz.uni-giessen.de](mailto:Gunter.Partosch@hrz.uni-giessen.de)

## Inhalt des Vortrags

- [HTML -- eine Dokumentenbeschreibungssprache](#)
  - [HTML-Versionen](#)
  - [Informationen über HTML](#)
- [HTML-Dokumente -- eine neue Art von Dokumenten](#)
- [Vergleichende Darstellung von LaTeX und HTML](#)
  - [Allgemeines](#)
    - [The Name of the Game](#)
    - [Eingabedatei](#)
  - [Setzen bzw. Ausgabe von Text](#)
    - [Zeilen- und Seitenumbruch](#)
    - [Spezielle Zeichen](#)
    - [Abstände](#)
    - [Kapitel und Überschriften](#)
    - [Fußnoten](#)
    - [hervorgehobene Textteile](#)
    - [Behandlung spezieller Textteile](#)
    - [Querverweise innerhalb des Dokuments](#)
  - [Spezialitäten](#)
    - [Schriftarten und Schriftgrößen](#)
    - [Abstände \(nur in LaTeX\)](#)
    - [Style-Sheets \(nur in HTML\)](#)
  - ["Übriggebliebenes"](#)
    - [Nur in LaTeX: Formeln, kompliziertere Tabellen](#)
    - [Nur in HTML: Hypertext-Strukturen, Formulare](#)

## HTML - eine Dokumentenbeschreibungssprache

- HTML (Hypertext Markup Language) ist eine SGML-Applikation.
- HTML übernimmt daher sehr viele Eigenschaften von SGML (z.B. logisches Markup, Entities).

- HTML wird durch eine DTD (Document Type Description) definiert.
- Die Norm sagt fast nichts über die tatsächliche Darstellung von Textteilen aus.
- Es gibt Validierer, die ein HTML-Dokument auf Korrektheit überprüfen.
- sehr viele firmenspezifische Entwicklungen
- Die schnell voranschreitende Entwicklung geht einher mit einer abnehmenden Akzeptanz der Normierung.

## HTML-Versionen:

- 2.0: unterstützt schon Bilder, interaktive Formulare, minimale typographische Elemente
- 2.1: minimale Erweiterungen
- 2.2: beinhaltet Version 2.1 und Tabellen aus Version 3
- 3.0: zusätzliche Zeichen- und Absatzformatierungen
- 3.1: Tabellen und mathematische Formeln
- 3.2: die meist eingesetzte Version; verwirft Version 3.0/3.1; zusätzlich Tabellen (einfacheres Modell), Applets, Hoch-/Tiefstellung, Textfluß um Bilder, Zeichensatzwahl, zusätzliche Attribute bei BODY und HR, Style-Sheets, ...
- 4.0: Frames, internationale Sprachunterstützung, Skripts; zusätzliche Möglichkeiten bei Style-Sheets, Formularen, Links

## Informationen über HTML:

- "Uniform Resource Locators (URL)"; RFC 1738; Dezember 1994; <http://rfc.fh-koeln.de/rfc/html/rfc1738.html>
- "Hypertext Markup Language - 2.0"; RFC 1866; November 1995; <http://rfc.fh-koeln.de/rfc/html/rfc1866.html>
- "HTML Tables"; RFC 1942; Mai 1996; <http://rfc.fh-koeln.de/rfc/html/rfc1942.html>
- "RFC 1866 (HTML 2.0 Strict Level 1) conformance"; 12.12.1996; <http://www.hut.fi/~oahvenla/HTML/rfc.html>
- Raggett, Dave**: "Hypertext Markup Language Specification Version 3.0"; 18.03.1995; <http://www.w3.org/pub/WWW/MarkUp/html3/html3.txt>
- "The W3 Consortium Announces HTML 3.2 - providing bright future for HTML as a non-proprietary standard"; <http://www.w3.org/pub/WWW/MarkUp/Wilbur/pr7may96.html>
- "Introducing HTML 3.2"; <http://www.w3.org/pub/WWW/MarkUp/Wilbur/>
- Raggett, Dave**: "W3C Document Type Definition for the Hypertext Markup Language version 3.2"; <http://www.w3.org/pub/WWW/MarkUp/Wilbur/HTML32.dtd>
- Raggett, Dave**: "HTML 3.2 Reference Specification"; 14.01.1996; <http://www.w3.org/pub/WWW/TR/REC-html32.html>
- Raggett, Dave; Le Hors, Arnaud; Jacobs, Ian**: "HTML 4.0 Specification"; 18.12.1997; <http://www.w3.org/pub/WWW/TR/REC-html40/>
- Hubert Partl**: "HTML-Einführung. Hypertext Markup Language die Sprache des World-Wide Web"; August 1996; <ftp://mail.boku.ac.at/www/htmlinf.ps.gz>

## HTML-Dokumente - eine neue Art von Dokumenten

- HTML-Dokumente nicht vorrangig für das Ausdrucken konzipiert, sondern für die Darstellung in Bildschirmfenstern
- kein Blocksatz; aber: Dokument kann sehr leicht verschiedenen Textbreiten angepaßt werden
- keine Silbentrennung
- Absätze müssen definitiv voneinander abgegrenzt werden

- kein Seitenumbruch, keine Seiteneinteilung, keine Seitennummern
  - kein Inhaltsverzeichnis (Schlagwortverzeichnis, Tabellenverzeichnis, Abbildungsverzeichnis) mit Seitennummern
  - keine Fußnoten, Randnotizen, Kopf- oder Fußzeilen
  - keine automatische Numerierung von Überschriften, Tabellen, Abbildungen
  - keine Seitenverweise
  - keine automatischen Ligaturen
- 
- anklickbare Verweise zu lokalen bzw. externen Datenquellen
  - interaktive Formulare möglich
- 
- logische Auszeichnung von Textteilen vom Autor gefordert; tatsächliche Darstellung Sache des HTML-Browsers

# Vergleichende Darstellung von LaTeX und HTML

## Allgemeines

### The Name of the Game

TeX, LaTeX209, LaTeX2e	HTML 2.0, HTML 2.1, HTML 2.2, HTML 3.0, HTML 3.1, HTML 3.2, HTML 4.0
------------------------	--

### Eingabedatei

#### Leerzeichen:

HTML und LaTeX: Leerzeichen und Zeilenenden einheitlich als Leerzeichen behandelt; mehrere Leerzeichen wie ein Leerzeichen

#### Leerzeilen:

Leerzeile bedeutet Ende eines Absatzes; mehrere Leerzeilen wie eine Leerzeile	keine besondere Bedeutung
--	---------------------------

#### spezielle Zeichen für die jeweilige Beschreibungssprache:

\$ % # _ { } ~ ^ " \   < >	< > &
um sie auszudrucken, z.B.: \ \$ \& \% \# \_ \{ \}	um sie am Bildschirm darzustellen: das jeweilige Entity verwenden, z.B. &lt; ; &gt; ; &amp ;

Befehle:

eingeleitet durch \; danach entweder ein Sonderzeichen oder alphabetische Zeichen (bis zum ersten nicht-aphabetischen Zeichen)	<ul style="list-style-type: none"><li>• Tag: eingeleitet durch &lt;, danach das eigentliche HTML-Element, abgeschlossen durch &gt;</li><li>• Es gibt Befehle, die aus einem Anfang-Tag und Ende-Tag bestehen (Container), z.B. &lt;B&gt;<i>text</i>&lt;/B&gt;, &lt;P ALIGN="right"&gt;<i>text</i>&lt;/P&gt;.</li><li>• In vielen Fällen kann das Ende-Tag weggelassen werden (wenn das im Kontext eindeutig ist).</li><li>• Es gibt ein allgemeines Ende-Tag &lt;/&gt; (wenn das im Kontext eindeutig ist).</li></ul>
LaTeX-Befehle können optionale (in eckigen Klammern) bzw. zwingende (in geschweiften Klammern) Parameter besitzen.	zwingende oder optionale Attribute (Schlüsselwort-Parameter innerhalb der spitzen Klammern), z.B. <P ALIGN="center">
Groß-/Kleinschreibung wird unterschieden.	Groß-/Kleinschreibung wird nicht unterschieden
mit Hilfe von % auf Folgezeilen fortsetzbar	nicht an Zeilenstruktur gebunden

Kommentare:

Alles, was hinter dem Zeichen % folgt (bis einschließlich des Zeilenendes), ist Kommentar und wird ignoriert.	Kommentare werden durch <!-- und --> eingeschlossen; können mehrere Zeilen umfassen; z.B. <!-- Das ist ein Kommentar -->
---	--

Aufbau eines Dokuments:

<pre>\documentclass[optionen]{dokumentklasse} ...</pre>	<pre>&lt;!DOCTYPE ...&gt; &lt;HTML&gt; &lt;HEAD&gt; ... &lt;TITLE&gt; <i>Titel des HTML-Dokuments</i>&lt;/TITLE&gt;</pre>
---	---

<i>Präambel mit Vereinbarungen...</i>	...
<code>\begin{document}</code>	<code>&lt;/HEAD&gt;</code>
...	<code>&lt;BODY&gt;</code>
<i>eigentliches LaTeX-Dokument mit dem auszugebenden Text...</i>	...
<code>\end{document}</code>	<i>eigentliches HTML-Dokument mit dem auszugebenden Text</i>
	...
	<code>&lt;/BODY&gt;</code>
	<code>&lt;/HTML&gt;</code>

### Erstellen der Eingabedatei:

in beiden Fällen mit einem normalen Editor; bei HTML auch mit speziellen Editoren (z.B. HoTMetaL Pro) oder durch Konverter in Textverarbeitungsprogrammen (z.B. in WinWord 97 oder Word Internet Assistant für WinWord 6.0/7.0)

### Verarbeiten der Eingabedatei:

Aufruf von LaTeX, gefolgt von einem Aufruf eines DVI-Treibers oder Previewers	Laden der HTML-Datei in einem HTML-Klienten (HTML-Browser)
---	--

## Setzen bzw. Ausgabe von Text

### Zeilen- und Seitenumbruch

#### Blocksatz:

voreingestellt; kann durch verschiedene Parameter (wie z.B. <code>\parindent</code> , <code>\parskip</code> ) gesteuert werden	nicht verfügbar
--	-----------------

#### Silbentrennung und Zeilenumbruch:

automatischer Zeilenumbruch mit Silbentrennung voreingestellt	automatischer Zeilenumbruch (ohne Silbentrennung) voreingestellt
kann durch verschiedene Parameter gesteuert werden (z.B. <code>\</code> , <code>\newline</code> , <code>\-</code> , <code>\hyphenation{muster}</code> , <code>\mbox{text}</code> , <code>~</code> )	nur wenige Steuerungsmöglichkeiten ( <code>&lt;BR&gt;</code> , <code>&amp;nbsp;</code> ; )

#### Seitenumbruch:

automatischer Seitenumbruch voreingestellt; Steuerungsmöglichkeiten durch <code>\newpage</code> , <code>\pagebreak</code> , <code>\nopagebreak</code> , <code>\enlargethispage{<i>maß</i>}</code> , usw.	nicht verfügbar; wird manchmal vom Autor durch eine Querlinie angedeutet ( <code>&lt;HR&gt;</code> )
---	---

**Spezielle Zeichen (Anführungszeichen, Binde- und Gedankenstriche, Auslassungspunkte, Ligaturen, Akzente und spezielle Buchstaben):**

Anführungszeichen: <code>`<i>text</i>`</code> , <code>"<i>text</i>"</code> , usw.	Anführungszeichen: <code>"<i>text</i>"</code>
Binde- und Gedankenstriche: <code>-</code> , <code>--</code> , <code>---</code>	Binde- und Gedankenstriche: nur <code>-</code> , <code>--</code>
Auslassungspunkte: <code>\ldots</code>	Auslassungspunkte: <code>...</code>
automatische Ligaturen: bei <code>ff</code> , <code>fi</code> , <code>fl</code> , usw.	automatische Ligaturen: nicht verfügbar
Akzente: <code>\^{<i>zeichen</i>}</code> , <code>\`{<i>zeichen</i>}</code> usw.	Akzente: nicht verfügbar; Zeichen direkt auswählen
spezielle Buchstaben: <code>\OE</code> , <code>\ae</code> , <code>\ss</code> , usw.	spezielle Buchstaben: <code>&amp;Oelig;</code> <code>&amp;aelig;</code> <code>&amp;szlig;</code>

**Abstände zwischen Wörtern:**

<code>\frenchspacing</code> , <code>\nofrenchspacing</code> , <code>\ , \@</code> , <code>\ , ~</code>	nur <code>&amp;nbsp;</code> ;
---	-------------------------------

**Kapitel und Überschriften:**

<code>\part</code> , <code>\chapter{<i>text</i>}</code> , <code>\section{<i>text</i>}</code> , <code>\subsection{<i>text</i>}</code> , ...	<code>&lt;H1&gt;<i>text</i>&lt;/H1&gt;</code> , <code>&lt;H2&gt;<i>text</i>&lt;/H2&gt;</code> , ... <code>&lt;H6&gt;<i>text</i>&lt;/H6&gt;</code>
---	--

**Fußnoten:**

<code>\footnote{<i>text</i>}</code>	nicht verfügbar
-------------------------------------	-----------------

**hervorgehobene Textteile:**

<code>\emph{<i>text</i>}</code> oder <code>{\em <i>text</i>}</code> (ggf. mit <code>\/</code> )	leichte Hervorhebung: <code>&lt;EM&gt;<i>text</i>&lt;/EM&gt;</code>
	starke Hervorhebung:

	<code>&lt;STRONG&gt;text&lt;/STRONG&gt;</code>
--	--

Andere logische Textauszeichnungen in HTML:

- Zitat: `<CITE>text</CITE>`
- Code-Beispiel: `<CODE>text</CODE>`
- Benutzereingabe: `<KBD>text</KBD>`
- Folge von literalen Zeichen: `<SAMP>text</SAMP>`
- Variablennamen: `<VAR>text</VAR>`

**Behandlung spezieller Textteile (Zitate, Listen, Textausrichtung, direkte Ausgabe, Abbildungen, Tabellen)**

Zitate:

quote-, quotation- und verse-Umgebung	<code>&lt;BLOCKQUOTE&gt;text &lt;/BLOCKQUOTE&gt;</code>
---------------------------------------	---

Listen:

einfache Listen: <code>itemize</code> -Umgebung	einfache Listen: <code>&lt;UL&gt;liste&lt;/UL&gt;</code>
numerierte Aufzählungen: <code>enumerate</code> -Umgebung	numerierte Aufzählungen: <code>&lt;OL&gt;liste&lt;/OL&gt;</code>
Beschreibungen: <code>description</code> -Umgebung	Beschreibungen: <code>&lt;DL&gt;liste&lt;/DL&gt;</code>
Listeneinträge jeweils durch <code>\item</code> bzw. <code>\item[term]</code> eingeleitet	Listeneinträge jeweils durch <code>&lt;LI&gt;</code> bzw. <code>&lt;DT&gt;term&lt;DD&gt;</code> eingeleitet

Textausrichtung:

Blocksatz voreingestellt	Blocksatz nicht verfügbar
andere Ausrichtungen durch <code>flushleft</code> -, <code>flushleft</code> - oder <code>center</code> -Umgebung	andere Ausrichtungen durch <code>&lt;CENTER&gt;text&lt;/CENTER&gt;</code> oder das Attribut <code>ALIGN</code> bei den Elementen <code>CAPTION</code> , <code>DIV</code> , <code>INPUT</code> , <code>P</code> , <code>TABLE</code> , <code>TD</code> , <code>TH</code> , <code>TR</code> (jeweils als Container formuliert); mögliche Attribut-Werte: <code>center</code> , <code>left</code> , <code>right</code> ; z.B. <code>&lt;P ALIGN=center&gt;text&lt;/P&gt;</code>

direkte Ausgabe:

<code>\verb text </code>	<code>&lt;SAMP&gt;text&lt;/SAMP&gt;</code> oder <code>&lt;CODE&gt;text&lt;/CODE&gt;</code>
verbatim-Umgebung	beste Entsprechung durch <code>&lt;XMP&gt;text&lt;/XMP&gt;</code> ( <code>&lt;LISTING&gt;text&lt;/LISTING&gt;</code> )  <code>&lt;PRE&gt;text&lt;/PRE&gt;</code> läßt Zeichenformatierungen und Konstruktionen der Art <code>&lt;A attribute&gt;text&lt;/A&gt;</code> zu

Abbildungen:

PostScript-Bilder können mit Hilfe zusätzlicher Pakete eingebaut werden.	Inline-Bilder: <code>&lt;IMG SRC="URL" attribute&gt;</code>  anklickbare Verweise auf externe Bilder durch <code>&lt;A HREF="URL" attribute&gt; anker&lt;/A&gt;</code>
Bildunterschrift durch die Anweisung <code>\caption{text}</code> in einer figure-Umgebung	keine Bildunterschrift in diesem Sinne verfügbar

Tabellen:

Tabellierungen: tabbing-Umgebung	Tabellierungen: nicht verfügbar
eigentliche Tabellen: tabular-Umgebung	eigentliche Tabellen: <code>&lt;TABLE&gt;tabelle&lt;/TABLE&gt;</code>
Tabellenüberschrift durch die Anweisung <code>\caption{text}</code> in einer table-Umgebung	Tabellentitel durch <code>&lt;CAPTION&gt;text&lt;/CAPTION&gt;</code>
Beispiel für eigentliche Tabelle: <code>\begin{table}</code> <code>\caption{Kleine Tabelle}</code> <code>\begin{tabular}{ l c r}</code> <code>\hline</code> zelle 1 & zelle 2 & zelle 3 \\ zelle 4 & zelle 5 & zelle 6 \\ <code>\hline</code> <code>\end{tabular}</code> <code>\end{table}</code>	Beispiele für (einfache) Tabelle: <code>&lt;TABLE BORDER&gt;</code> <code>&lt;CAPTION&gt;</code> Kleine Tabelle <code>&lt;/CAPTION&gt;</code> <code>&lt;TR&gt;&lt;TD&gt;zelle 1</code> <code>&lt;TD&gt;zelle 2</code> <code>&lt;TD&gt;zelle 3</code> <code>&lt;TR&gt;&lt;TD&gt;zelle 4</code> <code>&lt;TD&gt;zelle 5</code> <code>&lt;TD&gt;zelle 6</code> <code>&lt;/TABLE&gt;</code>



Querverweise innerhalb des Dokuments:

Verweisziel: <code>\label{label}</code>	Verweisziel: <code>&lt;A NAME="name" optionen&gt;text&lt;/A&gt;</code>
Verweis: <code>\ref{label}</code> bzw. <code>\pageref{label}</code>	Verweis: <code>&lt;A HREF="#name" optionen&gt;anker&lt;/A&gt;</code>

Spezialitäten

Schriftarten und Schriftgrößen

normal: <code>\textrm{text}</code> bzw. <code>{\rm text}</code>	normal:
Sans-Serif: <code>\textsf{text}</code> bzw. <code>{\sf text}</code>	Sans-Serif: nicht direkt verfügbar (ggf. über FONT)
Typewriter: <code>\texttt{text}</code> bzw. <code>{\tt text}</code>	Typewriter: <code>&lt;TT&gt;text&lt;/TT&gt;</code>
fett: <code>\textbf{text}</code> bzw. <code>{\bf text}</code>	fett: <code>&lt;B&gt;text&lt;/B&gt;</code>
kursiv: <code>\textit{text}</code> bzw. <code>{\it text}</code>	kursiv: <code>&lt;I&gt;text&lt;/I&gt;</code>
in LaTeX nicht direkt verfügbar	durchgestrichen: <code>&lt;STRIKE&gt;text&lt;/STRIKE&gt;</code>
schräg: <code>\textsl{text}</code> bzw. <code>{\sl text}</code>	schräg: nicht direkt verfügbar (ggf. über FONT)
Kapitälchen: <code>\textsc{text}</code> bzw. <code>{\sc text}</code>	Kapitälchen: nicht direkt verfügbar
Schriftgrößen: <code>\tiny, ..., \small, \normalsize, \large, ..., \Huge</code>	Schriftgrößen: <code>&lt;FONT SIZE=größe&gt;text&lt;/FONT&gt;</code> , wobei <i>größe</i> eine konkrete Zahl (mit oder ohne Vorzeichen) ist

Abstände (nur in LaTeX)

- Zeilenabstand: `\linespread{faktor}`
- spez. horizontale Abstände: `\, \enspace, \quad, \qquad, \hfill`
- spez. vertikale Abstände: `\vspace{länge}, \smallskip, \medskip, \bigskip, \vfill`

Style-Sheets (nur in HTML)

- HTML war ursprünglich als Dokumentenbeschreibungssprache konzipiert.
- Die Einführung vieler Formatierungselemente und -attribute auf niedriger Ebene verwischt diese Zielsetzung.

Style-Sheets ermöglichen dem Autor, Inhalt und Darstellung des Textes wieder zu trennen.

- Ein Style-Sheet kann beispielsweise zwischen `<STYLE>...</STYLE>` im Kopf des Dokuments vereinbart werden.
- Der Autor kann sich auch auf externe Style-Sheets beziehen.
- eine sehr einfache Form eines Style-Sheets:

```
P      {font-size : 12pt}
STRONG {color : #800000}
```

---> "normale" Absätze: in der Schriftgröße 12pt,

---> "starke" Hervorhebungen: dunkelrot

- noch zahlreiche andere Mechanismen verfügbar

## "Übriggebliebenes"

### Nur in LaTeX: Formeln, kompliziertere Tabellen

- Außer `<SUB>text</SUB>` und `<SUP>text</SUP>` sind in HTML 3.2 keine weiteren Konstruktionen für Formeln verfügbar.
- Im Vergleich zu HTML 3.2 können in LaTeX deutlich komplexere Tabellen erstellt werden.

### Nur in HTML: Hypertext-Strukturen, Formulare

#### Hypertext-Strukturen:

Mit Hilfe der Konstruktion `<A HREF="URL" optionen>anker</A>` kann in HTML auf lokale oder externe Datenquellen verwiesen werden.

*URL* (Uniform Resource Locator) hat hierbei den folgenden Aufbau:

*protokoll*: *//rechner/pfad*

mit

*protokoll*

Protokoll oder Zugriffsmechanismus auf *rechner*;

folgende Angaben sind u.a. möglich (http, mailto, ftp, news, telnet)

*rechner*

Internet-Adresse der Datenquelle (ggf. mit Port-Nummer)

*pfad*

führt auf *rechner* zur gewünschten Datei

#### Beispiele:

```
http://www.uni-giessen.de/hrz/texservice.html
ftp://ftp.uni-giessen.de/pub/doc/gopher.faq
```

news:de.comp.text.tex  
mailto:Gunter.Partosch@hrz.uni-giessen.de

Formulare:

```
...
<FORM
ACTION="mailto:dante@dante.de"  METHOD="POST"
ENCTYPE="text/plain">

<H1>Anmeldung zur Tagung DANTE'99</H1>

<P>
Name   :<INPUT NAME="Text1"  VALUE="Partosch">
<P>
Vorname :<INPUT NAME="Text2"  VALUE="Günter">
<P>
Ich werde
<P>
<INPUT TYPE="CHECKBOX" NAME="Kontrollkästchen1" CHECKED>
am Er&ouml;ffnungsvortrag,
<P>
<INPUT TYPE="CHECKBOX" NAME="Kontrollkästchen2" CHECKED>
an den Votr&auml;gen am Vormittag,
<P>
<INPUT TYPE="CHECKBOX" NAME="Kontrollkästchen3">
am gemeinsamen Mittagessen,
<P>
<INPUT TYPE="CHECKBOX" NAME="Kontrollkästchen4" CHECKED>
an den Votr&auml;gen am Nachmittag und
<P>
<INPUT TYPE="CHECKBOX" NAME="Kontrollkästchen5" CHECKED>
am Abschl&szlig;vortrag
<P>
teilnehmen.
<P>
Die Teilnahmegeb&uuml;hr bezahle ich :
<SELECT NAME="Dropdown1" >
<OPTION>per Kreditkarte
<OPTION>per &Uuml;berweisung
<OPTION>per Lastschrift
<OPTION>per Euro-Scheck
<OPTION>bar</SELECT>
<P>
<INPUT TYPE="SUBMIT"  VALUE="Abschicken">

<P>
<INPUT TYPE="RESET"  VALUE="Verwerfen">

</FORM>
...
```

---

[wwwmaint@dante.de](mailto:wwwmaint@dante.de) bei Fehlermeldungen, Anregungen, Ergnzungen usw. zum Angebot dieses WWW-Servers  
[dante@dante.de](mailto:dante@dante.de) bei Fragen (Mitgliedschaft, Bestellungen, Nachfragen, Tagungen usw.) an DANTE e.V.

---