



# XHTML

Referat von Peter Weidling innerhalb des Seminars  
Advanced Structured Content Administration  
im Masterstudiengang Informationswirtschaft  
im WS 2005/2006 bei Prof. Dr. Wolf-Fritz Riekert  
Präsentationsmaterialien angepasst  
an die Lehrveranstaltung „Computernetze und Webapplikationen“  
durch Prof. Dr. Wolf-Fritz Riekert

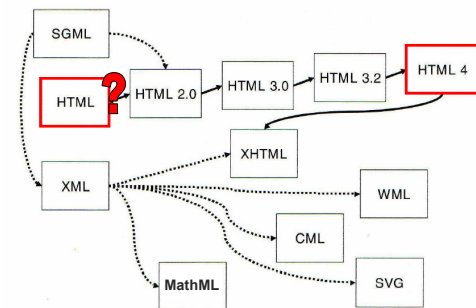
## Lernfragen zu XHTML

- Ist XHTML eine Erweiterung von HTML?
- Wie hängt XHTML mit XML zusammen?
- Wie ist ein XHTML-Dokument strukturiert ?
- Wo sind die klassischen Einsatzgebiete von XHTML?

## Was ist HTML?

- HTML steht für  
**H**ypertext **M**arkup **L**anguage
- Dokumentenformat zur Auszeichnung von Hypertext
- „Sprache des World Wide Web“
- entwickelt von Tim Berners-Lee  
(1989 am CERN in Genf)
- HTML ist eine Anwendung der Metasprache SGML

## Überblick



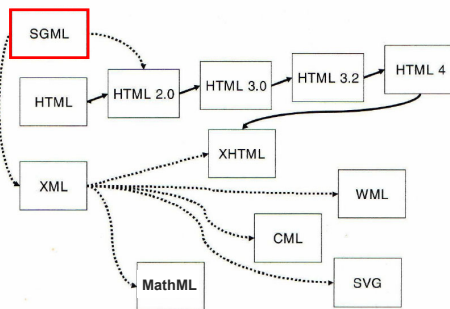
## Die Entstehungsgeschichte

- **HTML** (ohne Versionsnummer, 3. November 1992): Die Urversion, die sich nur an **Text** orientierte. [Weblink Erste Website](#)
- **HTML** (ohne Versionsnummer, 30. April 1993): Zu Text kommen neben Attributen wie fette oder kursive Darstellung auch die **Bildintegration** dazu.
- **HTML+** (November 1993) Geplante Erweiterungen, die in spätere Versionen einfließen, aber nie als HTML+ verabschiedet wurden. [Weblink](#)
- **HTML 2.0** (November 1995): Die mit RFC 1866 definierte Version führt u.a. **Formulartechnik** ein. Der Status dieses Standards ist „HISTORIC“. Auch die Vorgänger sind veraltet.

## Die Entstehungsgeschichte (II)

- **HTML 3.0**: Die Version erscheint nicht, weil sie mit der Einführung des Netscape-Browsers in der Version 3 bereits vor der Veröffentlichung obsolet ist.
- **HTML 3.2** (14. Januar 1997): Neu in dieser Version sind zahlreiche Feature wie **Tabellen**, Textfluss um Bilder, Einbindung von **Applets**.
- **HTML 4.0** (18. Dezember 1997): Mit dieser Version werden auch **Stylesheets**, Skripte und **Frames** eingeführt.
- **HTML 4.01** (24. Dezember 1999): Ersetzt HTML 4.0 mit vielen, kleineren Korrekturen. Es ist die **letzte HTML-Version**.

## Überblick



## Exkurs: SGML

- HTML & SGML werden häufig in einem Zug genannt
- SGML steht für **Standard Generalized Markup Language**
- Verfahren mit dem Auszeichnungssprachen (wie z.B. HTML) beschrieben werden können
- Internationaler Standard der ISO (genügt der ISO-Norm 8879)

## Exkurs: SGML

- Regeln für Zeichen, Zeichenfolgen und Schlüsselwörter
- formale Definition für „korrektes HTML“

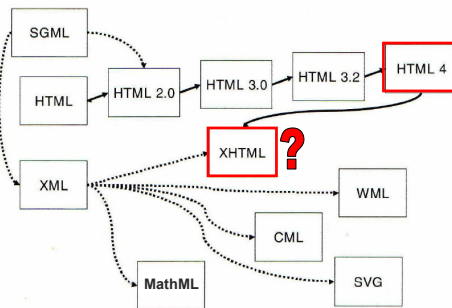
## Exkurs: SGML

- Dokumentendefinition legt nur die Grammatik (Syntax) fest, nicht die Bedeutung (Semantik) der Elemente
- Keine Aussage über die Formatierung von HTML durch den Browser

```

<html>
  <head>...</head>
  <body><h1>Diese Reihenfolge erlaubt die DTD</h1>
    <h2>Wie h2 im Browser angezeigt wird,
      geht nicht aus der DTD hervor</h2>
  </body>
</html>
  
```

## Überblick



## Was ist XHTML?

- XHTML steht für **Extensible Hypertext Markup Language**
- XHTML ist eine Neuformulierung von HTML in XML
- XHTML 1.0 enthält die Elemente und Attribute von HTML 4.01

## NEU: XML-Konformität

---

- Für XHTML-Dokumente gelten die Regeln von XML
- Ist ein XHTML-Dokument konform zu XML, also **syntaktisch korrekt**, dann sagt man, dass es **„wohlgeformt“** ist.

## Wohlgeformtheit

---

- XML und somit auch XHTML-Dokumente sind wohlgeformt, wenn sie folgende Regeln einhalten:
  - Nur ein Root-Element
  - Ein Element muss einen Start- und einen End-Tag haben!  
Ausnahme: Leere Elemente: `<empty />`

## Wohlgeformtheit

---

- Start- und End-Tag müssen identisch sein!  
(Groß/Klein-Schreibung beachten!)
- Verboten:  
`<el1>Inhalt</EL1>`

## Wohlgeformtheit

---

- Start- und End-Tag müssen identisch sein!  
(Groß/Klein-Schreibung beachten!)
- Korrekt:  
`<el1>Inhalt</el1>`

## Wohlgeformtheit

---

- Kein Überlappen der Elemente!
- Verboten:  
`<el1>`  
`<el2>Inhalt</el1>`  
`</el2>`

## Wohlgeformtheit

---

- Kein Überlappen der Elemente!
- Korrekt:  
`<el1>`  
`<el2>Inhalt</el2>`  
`</el1>`

## Wohlgeformtheit

---

- Alle Attribute müssen einen Wert haben und dieser wird in Anführungszeichen (" " oder ' ') gestellt.
- Verboten:  
`<el1 attribut=Wert>`  
`<el2 attribut>Inhalt</el2>`  
`</el1>`

## Wohlgeformtheit

---

- Alle Attribute müssen einen Wert haben und dieser wird in Anführungszeichen (" " oder ' ') gestellt.
- Korrekt:  
`<el1 attribut="Wert">`  
`<el2 attribut="Wert">Inhalt</el2>`  
`</el1>`

## Übung: Wohlgeformt

- Welches der folgenden Beispiele ist wohlgeformt?
- Wenn nicht, warum nicht?

## Übung 1

```
<a>  
<b>Eine leichte Übung zum Einstieg.</b>  
</a>
```

**Regel:** Elemente müssen immer geschlossen werden!

## Übung 2

```
<f>  
<c attri="Text">Hier steht ein  
<d>Text.</d>  
</c> Noch mehr Text.  
</f>
```

**Regel:** Werte von Attributen müssen in Anführungszeichen stehen!

## Übung 3

```
<j>  
  Beispiel Nummer drei<m />  
</j>
```

Korrekt!

## Übung 4

<n>

<o>Blindtext findet man z. B. unter  
www.blindtexte.de </o>

</n>

**Regel:** Start- und End-Tags müssen identisch sein!

## Validität

- Valide XHTML-Dokumente erfüllen folgende Anforderungen:
  - sie sind wohlgeformt (→ syntaktisch korrekt)
  - sie enthalten nur definierte Elemente und Attribute (→ semantisch korrekt)
- Die vorangegangenen Beispiele waren konstruiert. (Die verwendeten Elemente und Attribute sind nicht definiert. → nicht valide)

## Validität

- Wie in XML sind die zulässigen Elemente und Attribute in einer sog. DTD (Doctype Definition) festgehalten.
- Jedes XHTML-Dokument muss der DTD genügen, d.h. validierbar sein.

## Aufbau eines XHTML-Dokuments

- XML-Deklaration  
<?xml version="1.0" encoding="iso-8859-1" ?>

## Aufbau eines XHTML-Dokuments

### ■ XML-Deklaration

```
<?xml version="1.0" encoding="utf-8"?>
```

### ■ XHTML-Dokumententyp

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"  
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
```

## Dokumententypen

### ■ XHTML 1.0 Strict

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"  
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
```

### ■ XHTML 1.0 Transitional

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"  
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
```

### ■ XHTML 1.0 Frameset

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Frameset//EN"  
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-frameset.dtd">
```

## W3C Empfehlungen

- **XHTML 2.0** 7/2004 Working Draft
- nicht mehr rückwärtskompatibel zu HTML
- Keine Präsentationselemente → CSS(!)
- dafür mehr Strukturierungsmöglichkeiten
- Bessere Unterstützung für Bildtypen

## Aufbau eines XHTML-Dokuments

### ■ Dokumenten-Kopf

```
<html xmlns="http://www.w3.org/1999/xhtml">  
<head>  
  <title>Einführung in XHTML</title>  
</head>
```

### ■ Dokumenten-Rumpf

```
<body>  
  <h1>Alles XML-konform!</h1>  
  (...)  
</body>  
</html>
```

```
<?xml version="1.0" encoding="iso-8859-1" ?>
```

XML-Prolog

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"  
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
```

DTD

```
<html xmlns="http://www.w3.org/1999/xhtml">  
<head>  
<title>Titel der Seite</title>  
</head>
```

Kopf

```
<body>  
<h1>Hello World!</h1>  
<p>  
Das ist meine erste XHTML-Seite!  
</p>  
</body>  
</html>
```

Rumpf

## Vorteile von XHTML

- XML-Konformität
- HTML oder XML Document Object Models (.htm, .html, .xml)
- Einfachere Verarbeitung durch den Browser ( valide Dokumente → weniger Ausnahmen)
- Verwendung von XML erlaubt eine einfachere Modularisierung, d.h. die Semantik kann dem jeweiligen Bedarf angepasst werden.

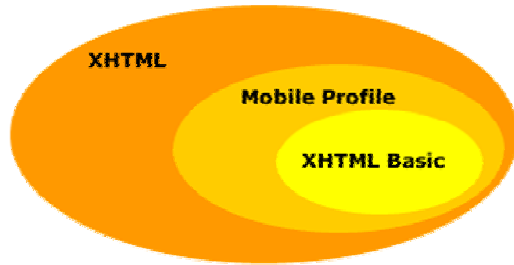
## Modularisierung

- Gliederung in Teilsprachen
- Teilsprache erfüllt bestimmte Funktion
- definiert durch angepasste DTDs:
- beliebige Verkettung der Teilsprachen

## Gründe für Modularisierung

- HTML/XHTML sehr komplex
- Hohe Hardware Anforderungen
- Darstellung auf unterschiedlichen Endgeräten, wie z.B. Mobiltelefonen oder Fernsehern
- Zusätzlich durch striktere Regeln effizienter

# Modularisierung - Beispiel



# Modularisierung – Beispiel II



XHTML auf verschiedenen Endgeräten

# Einordnung von XML, HTML und XHTML

