

XML

Eine Einführung

XML – Eine Einführung

Warum XML?

Warum XML?

- W3C

- World Wide Web Consortium

- beschreibt XML:

- *„Die Extensible Markup Language (XML) ist eine Teilmenge von SGML...*

- Das Ziel ist es zu ermöglichen, generisches SGML in der Weise über das Web auszuliefern, zu empfangen und zu verarbeiten, wie es jetzt mit HTML möglich ist. XML wurde entworfen, um eine einfache Implementierung und Zusammenarbeit sowohl mit SGML als auch mit HTML zu gewährleisten.“*

- *<http://www.w3c.org/TR/REC-xml>*

Warum XML?

- XML wurde speziell entworfen, um Informationen im WWW bereitzustellen, ebenso wie HTML
- Fragen:
 - Wenn HTML so universell ist, warum benötigte man eine neue Auszeichnungssprache?
 - Warum können wir nicht einfach SGML nehmen?
 - Soll XML HTML ergänzen oder ersetzen?

Warum XML? Warum XML Notwendig Ist.

- HTML stellt eine feste Menge vordefinierter *Elemente* bereit, mit denen Sie die Komponenten einer typischen, allgemeinen Zwecken dienenden Webseite auszeichnen können.
- **Beispiele:**
 - `fetter Text`
 - `` Bildbeschreibung
 - `Anker`

Warum XML? Warum XML Notwendig Ist.

- HTML Beispiel

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<html>
  <head>
    <title>Hinweise auf HTML Tutorials</title>
  </head>
  <body>
```

Falls Sie sich mit HTML vertraut machen wollen, sollten Sie unbedingt die folgende Standardreferenz in deutsch kennen lernen:


```
<a href=http://www.netzwelt.com/selfhtml/index.html>"SELFHTML: Version 8.0 vom 27.10.2001, Autor: Stefan M&uuml;nz</a>
```

```
<p>In dieser Referenz finden Sie weitere Hinweise auf die g&auml;ngige Literatur</p>
```

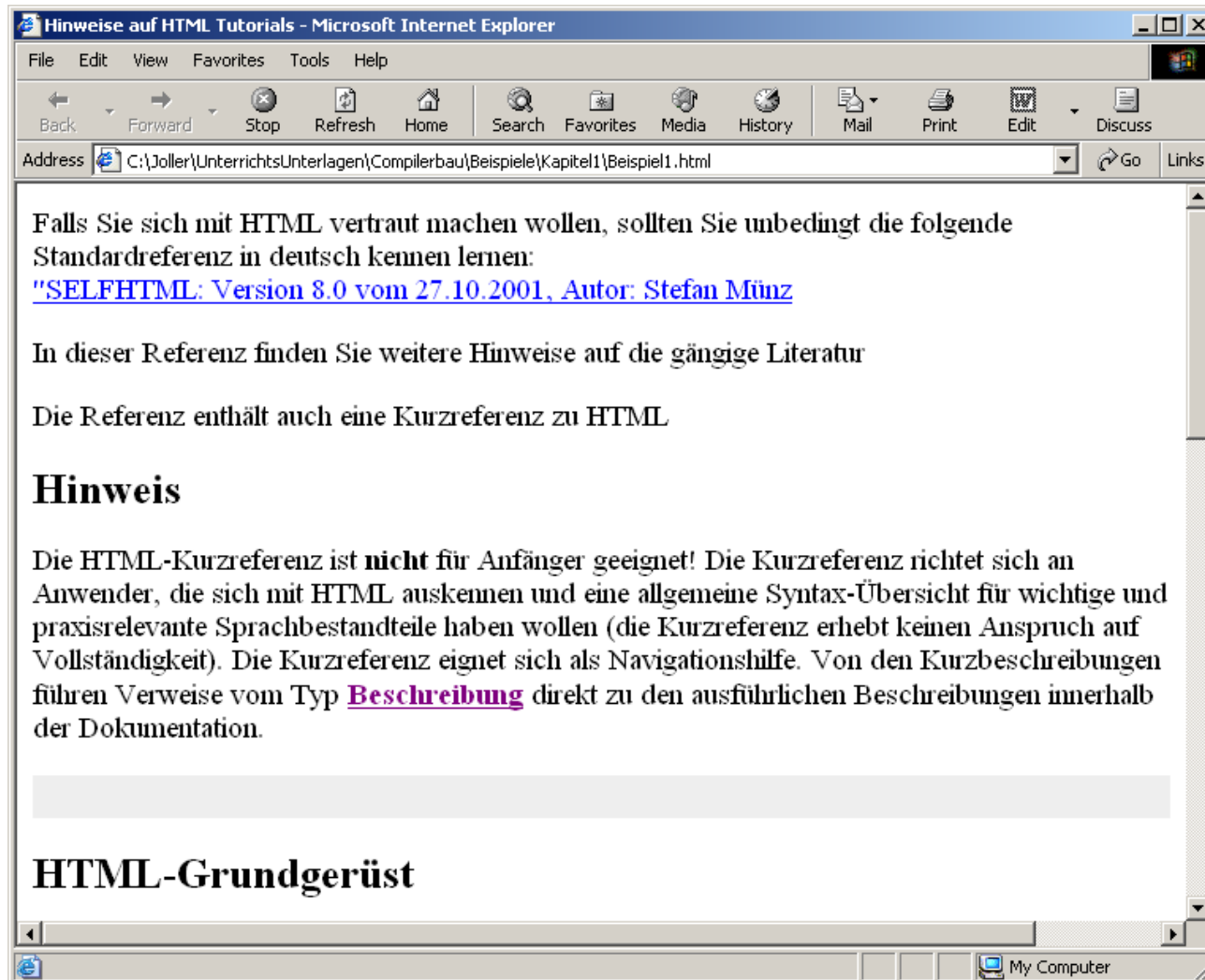
```
<p>Die Referenz enthält auch eine Kurzreferenz zu HTML</p>
```

```
<h2 >Hinweis</h2>
```

```
<p>Die HTML-Kurzreferenz ist <b>nicht</b> f&uuml;r Anf&auml;nger geeignet!
```

Die Kurzreferenz richtet sich an Anwender, ...

Warum XML? Warum XML Notwendig Ist



Warum XML? Warum XML Notwendig Ist

- Jedes HTML Element beginnt mit einem *Start*-Tag: einem Textblock, der in die Zeichen `<` und `>` eingeschlossen ist und den Elementnamen sowie womöglich weitere Informationen enthält.
- Die meisten Start-Tags enden mit einem *End*-Tag, das dem korrespondierenden Start Tag entspricht, allerdings nur aus dem Schrägstrich `/` und dem Elementnamen besteht

Warum XML? Warum XML Notwendig Ist.

- *Aufbau eines HTML Elements*

- **<h2 >Hinweis</h2>**
Start-Tag *<h2>*
Inhalt: *Hinweis*
End-Tag *</h2>*

- *Typische HTML Elemente*

- *HTML (<html>...</html>)* : *die gesamte Seite*
- *HEAD (<head>...</head>)* : *Seitenkopf*
- *TITLE (<title>...</title>)* : *Seitentitel*
- *BODY (<body>...</body>)* : *Haupttext*
- *H1 (<h1>.....</h1>)* : *Überschrift (Stufe 1)*

Warum XML? Warum XML Notwendig Ist.

- Der Browser:
 - erkennt jedes dieser Standardelemente und weiss, wie sie zu formatieren und anzuzeigen sind.
- Probleme:
 - Inkompatible Browser verunmöglichen strikte Standardisierung

Warum XML? Warum XML Notwendig Ist.

- Die Menge der vordefinierten HTML Elemente wurde zwar erheblich erweitert seit der ersten Version (aktuell: Version 4), aber HTML kann auch heute noch nicht zur Definition komplexer Dokumenteninhalte eingesetzt werden.

Warum XML? Warum XML Notwendig Ist.

- Dokumente, die nicht aus typischen Komponenten bestehen (Überschriften, Absätze, Listen, Tabellen und so weiter)
 - HTML fehlen beispielsweise die Elemente, die zur Auszeichnung einer Partitur oder beispielsweise die Elemente, die zur Auszeichnung einer Partitur oder einer Menge von mathematischen Gleichungen notwendig sind

Warum XML? Warum XML Notwendig Ist.

- Datenbanken wie etwa ein Buchbestand
 - Sie können eine HTML Seite verwenden, um darin statische Datenbankinformationen zu speichern, etwa eine Liste mit Buchbesprechungen
 - Wollen Sie diese Informationen allerdings filtern, durchsuchen, sortieren oder auf eine andere Weise damit arbeiten, müsste jede einzelne Informationseinheit ausgezeichnet werden (wie etwa in MS Access).
HTML verfügt nicht über die hierzu erforderlichen Elemente.

Warum XML? Warum XML Notwendig Ist.

- Dokumente, die in einer hierarchischen Baumstruktur organisiert werden sollen.
 - Nehmen Sie beispielsweise an, dass Sie an einem Buch schreiben und es in Teile, Kapitel, Abschnitt A und Abschnitt B und Abschnitt C und so weiter unterteilen wollen.
Ein Programm könnte dann dieses strukturierte Dokument verwenden, um ein Inhaltsverzeichnis zu generieren, unterschiedlich detaillierte Auszüge zu erstellen, bestimmte Abschnitte zu extrahieren und die Informationen auf andere Weise zu verarbeiten.

Warum XML? XML, Die Lösung.

- Die XML-Definition besteht nur aus einer Basissyntax.
 - Wenn Sie ein XML Dokument erstellen, dann greifen Sie nicht auf eine beschränkte Menge vordefinierter Elemente zurück, sondern erstellen eigene Elemente, denen Sie beliebige Namen zuweisen (daher *extensible*).
- Sie können daher XML einsetzen, um jeden Dokumenttyp zu beschreiben.

Warum XML? XML, Die Lösung

- Beispiel: Bücherliste.xml

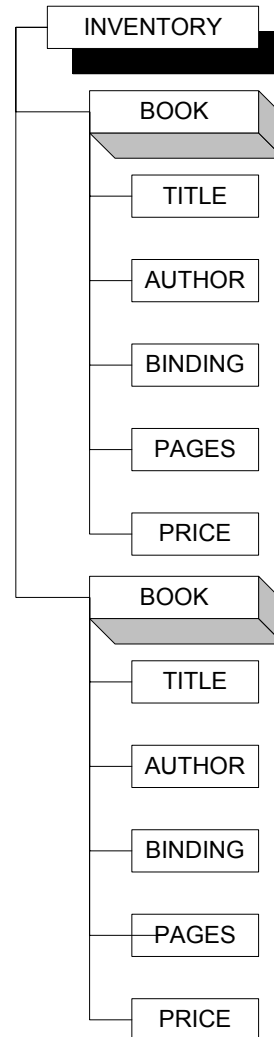
```
<?xml version="1.0"?>
<INVENTORY>
  <BOOK>
    <TITLE>The Adventures of Huckleberry Finn</TITLE>
    <AUTHOR>Mark Twain</AUTHOR>
    <BINDING>Taschenbuch</BINDING>
    <PAGES>336</PAGES>
    <PRICE>EU 12,75</PRICE>
  </BOOK>
  <BOOK>
    <TITLE>The Legend of Sleepy Hollow</TITLE>
    <AUTHOR>Washington Irving</AUTHOR>
    <BINDING>Taschenbuch</BINDING>
    <PAGES>98</PAGES>
    <PRICE>EU 4,95</PRICE>
  </BOOK>
</INVENTORY>
```


Warum XML? XML, Die Lösung.

- Es ist wichtig zu verstehen, dass die Elementnamen in einem XML Dokument (INVENTORY, BOOK, ...) *nicht* Teil der XML Definition sind.
 - Sie wählen diese Namen, wenn Sie ein bestimmtes Dokument erstellen
 - Sie können für Ihre Elemente beliebig gültige Namen wählen (BESTAND, ARTIKEL, ...)
 - Wenn Sie Elemente in einem XML Dokument benennen, dann wählen Sie beschreibende Namen, beispielsweise BUCH, NAME, ...

Warum XML? XML, Die Lösung

- Das vorige Beispiel zeigt eine hierarchische Struktur



XML eignet sich also zur Definition eines hierarchisch strukturierten Dokuments, wie z.B. eines Buches oder einer Buchliste

Warum XML? XML Dokumente Erstellen

- XML kennt keine vordefinierte Elemente
- Allerdings besitzt XML eine *streng definierte Syntax*
 - Beispiel:
 - Im Gegensatz zu HTML muss in XML jedes XML Element sowohl ein Start-Tag als auch ein End-Tag besitzen (oder das besondere Tag des *leeren Elements*)
 - Jedes verschachtelte Element muss vollständig innerhalb des Elements enthalten sein, das es umschließt.
- Die hohe Flexibilität bei der Erstellung erfordert eine strenge Syntax.

Warum XML? XML Dokumente Anzeigen.

- HTML:
 - Der Browser weiss wie die Elemente formatiert werden müssen.
- XML:
 - Der Browser kennt die Tags nicht.
 - Möglichkeiten:
 - Stylesheet Verknüpfungen
 - Datenbindung
 - Skript

Warum XML? XML Dokumente Anzeigen

- Stylesheet Verknüpfungen
 - Mit dieser Technik verknüpfen Sie ein Stylesheet mit dem XML Dokument
 - Ein Stylesheet ist eine separate Datei, die Anweisungen zum Formatieren der einzelnen XML Elemente enthält.
 - Sie können entweder ein CSS Stylesheet (cascaded style sheet), das auch bei HTML Dateien verwendet wird, einsetzen, oder
 - ein XSL Stylesheet verwenden.
XSL ist wesentlich flexibler als CSS.

Warum XML? XML Dokumente Anzeigen

- Datenbindung

- Diese Option erfordert, dass Sie eine HTML Seite erstellen, das XML Dokument damit verknüpfen und HTML Standardelemente in der Seite, etwa die Elemente SPAN oder TABLE, an die XML Elemente anbinden.

Die HTML Elemente zeigen dann automatisch die Informationen der XML Elemente an, an die sie gebunden sind.

Warum XML? XML Dokumente Anzeigen

- Skripts
 - Bei dieser Technik erstellen Sie eine HTML Seite und verknüpfen das XML Dokument damit.
Sie erstellen ein Skript (JavaScript, VBScript), um auf individuelle XML Elemente zuzugreifen bzw. sie anzuzeigen.
 - Der Browser legt das XML Dokument als Dokumentobjektmodell (DOM) offen, das eine grosse Anzahl von Objekten, Eigenschaften und Methoden bereitstellt, über die der Skriptcode auf die XML Elemente zugreifen, sie bearbeiten und ändern kann.

Warum XML? SGML, HTML Und XML

- SGML (Strucured Generalized Markup Language)
 - war die erste Auszeichnungssprache aus der
 - sowohl HTML
 - als auch XML abgeleitet wurden.
 - Sie beschreiben ein Dokument in SGML, indem Sie eine passende Elementmenge sowie eine Dokumentstruktur erstellen.

Warum XML? SGML, HTML Und XML

- XML
 - Stellt eine vereinfachte Version von SGML dar, die für das Web optimiert ist.
 - Wie bei SGML kann man eine *XML Anwendung*, ein *Vokabular*, definieren.

Warum XML? Wird HTML Durch XML Ersetzt?

- Zum aktuellen Zeitpunkt lautet die Antwort auf die obige Frage:
 - NEIN
- HTML ist immer noch die Primärsprache, um Browser anzuweisen, wie Informationen im Web abzuzeigen sind.
- Statt HTML zu ersetzen, wird XML zusammen mit HTML eingesetzt, um die Fähigkeiten von Webseiten für folgende Zwecke zu erweitern:
 - Darstellung fast jeden Dokumenttyps
 - Informationen auf andere Arten sortieren, filtern, ...
 - Präsentation stark strukturierter Informationen

Warum XML? Die Offiziellen Ziele Von XML

- Ziel 1:
 - XML soll im Internet leicht einsetzbar sein
- Ziel 2:
 - XML soll ein breites Spektrum von Anwendungen unterstützen.
 - Primär soll XML über Server und Browser Informationen im Web bereitgestellt werden.
 - XML kann aber auch von anderen Programmtypen verwendet werden.
 - Beispiele
Einsatz von XML für den Austausch von Finanzinformationen
Verteilung und Aktualisierung von Software.
Erstellung von Voice Skripts

Warum XML? Die Offiziellen Ziele Von XML

- Ziel 3:
 - XML soll mit SGML kompatibel sein
- Ziel 4:
 - Es soll leicht sein, Programme zu schreiben, die XML Dokumente verarbeiten
- Ziel 5:
 - Die Anzahl optionaler Elemente von XML soll auf das absolute Minimum, idealerweise auf Null, beschränkt sein.

Warum XML? Die Offiziellen Ziele Von XML

- Ziel 6:
 - XML Dokumente sollen für Menschen lesbar und hinreichend klar sein.
- Ziel 7:
 - Das XML Design soll rasch zur Verfügung stehen.
- Ziel 8:
 - Das Design von XML soll formal und prägnant sein
- Ziel 9:
 - XML Dokumente sollen leicht zu erstellen sein
- Ziel 10:
 - Knappheit soll beim XML Markupcode von geringer Bedeutung sein (also nicht kryptisch, schwer verständlich)

Warum XML? Standardanwendungen

- Unternehmen können XML Anwendungen für die Erstellung von Dokumenten definieren.
- Zu einer XML Anwendung kann eine DTD (Dokumenttyp Definition) oder ein Schema erstellt werden, welches dem Datenbankschema in Datenbankanwendungen entspricht.

Warum XML? XML Anwendungen Zur Erweiterung Von XML Dokumenten.

- Beispiele:
 - XSL : Extensible Stylesheet Language
 - XML Schema
 - XML Linking Language (XLink)
 - Gestattet die Verknüpfung von XML DOKumenten
 - XML Pointer Language (XPointer)
 - Gestattet die flexible Definition von Verknüpfungszielen.

Warum XML? Praktische XML Anwendungen

- Ref:
 - <http://www.oasis-open.org/cover/xml.html#applications>
- Datenbanken speichern
 - Beispiel: Bücherliste
- Dokumente strukturieren
 - Beispiel: Buchliste
- Multimediapräsentationen speichern (SMIL)
- Channels definieren (CDF)
- Software Pakete beschreiben (OSD)