



xml

Bernd Blümel
2001

Gliederung

Prüfung



- 1. Beispiel**
- 2. Syntax**
- 3. DTD's und XSL**
- 4. Ausblick**

XML

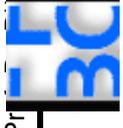
Prüfung



- **Format für den Datenaustausch**
- **Information über Dateninhalt**
- **"Verwandschaft" mit html**

Beispiel: Kinoprogramm einer Stadt

Prüfung



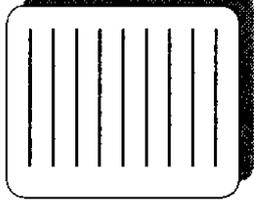
```
<movie>
  <title>Star Trek: Insurrection</title>
  <star>Patrick Stewart</star>
  <star>Brent Spiner</star>
  <theatre>
    <theatre-name>Mondo Plex 2000</theatre-name>
    <showtime>1415</showtime>
    <showtime>1630</showtime>
    <showtime>1845</showtime>
    <showtime>2100</showtime>
    <showtime>2315</showtime>
  </price>
  <adult-price>8.50</adult-price>
  <child-price>4.50</child-price>
</price>
</theatre>
```

...

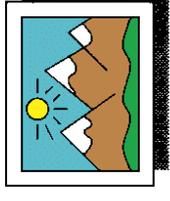
Woraus besteht ein Dokument?

Inhalt

- **Inhalt:** Text, Bilder, Tabellen, Musik, Videos, ...



Text

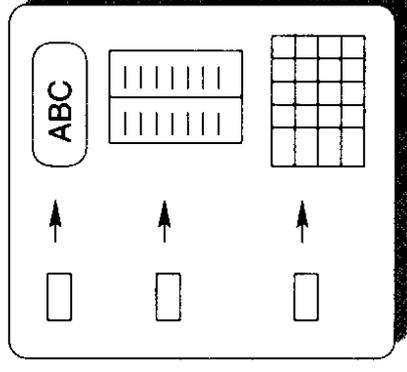


Bilder

- **Darstellungsinformation:**

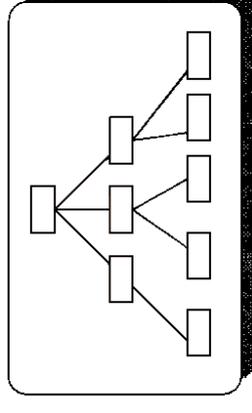
Layout, Formatierung, Schrift, ...

Darstellung



- **Struktur:** Aufteilung in Kapitel, Abschnitte, Fußnoten, Titel, ...

Struktur



(Metadaten)

Was sind Metadaten?

Prüfung



```
<!-- 1. Listing -->
<?xml version="1.0"?>
...
<p>Heinz</p>
<p>Müller</p>

<p>11</p>
<p>12</p>
<p>1970</p>
```

```
<!-- 2. Listing -->
<?xml version="1.0"?>
...
<Name>
  <Vorname>Heinz</Vorname>
  <Nachname>Müller</Nachname>
</Name>
<Geburtstag>
  <Tag>11</Tag>
  <Monat>12</Monat>
  <Jahr>1970</Jahr>
</Geburtstag>
```

Metadaten sind die beschreibende Information über die Struktur und Bedeutung von Daten und über die Anwendungen und Prozesse, die Daten manipulieren oder verarbeiten.

XML – Syntax

Prüfung

- Information wird durch start und end tags (`<name>` und `</name>`) markiert.
- Der so markierte Teil der Information heißt Element.
- Es existiert ein root-Element.
- Elemente können erweitert werden durch Attribute, d.i. ein Paar (Name, Wert).
- Es können Entity References benutzt werden, das sind (in der DTD deklarierte) Abkürzungen, z.B. `&musik;` für die Entity musik.

```
<?xml version="1.0"?>
...
<Adresse Land="D" >
  <Straße>Königsallee 234</Straße>
  <PLZ>44795</PLZ>
  <Stadt>Bochum</Stadt>
  &musik;
</Adresse>
```

XML – Struktur

```
<?xml version="1.0"?>  
<!DOCTYPE Liste [  
  <!ELEMENT Fließtext (#PCDATA | item ) * >  
  <!ELEMENT item (#PCDATA) >  
]>
```

} DTD
} Dokument-
typ-
Deklaration

```
<Fließtext>
```

Es gibt verschiedene Markup-Sprachen:

```
<item> XML, </item>  
<item> SGML, </item>  
<item> html. </item>  
</Fließtext>
```

} Dokument-
Instanz

Der Aufbau eines XML-Dokuments:

- Der Prolog, mindestens bestehend aus der Dokument-Deklaration und der DTD (Dokumenttyp-Definition)
- Der Dokument-Instanz mit dem Inhalt

Dokumenttyp-Definition (DTD)

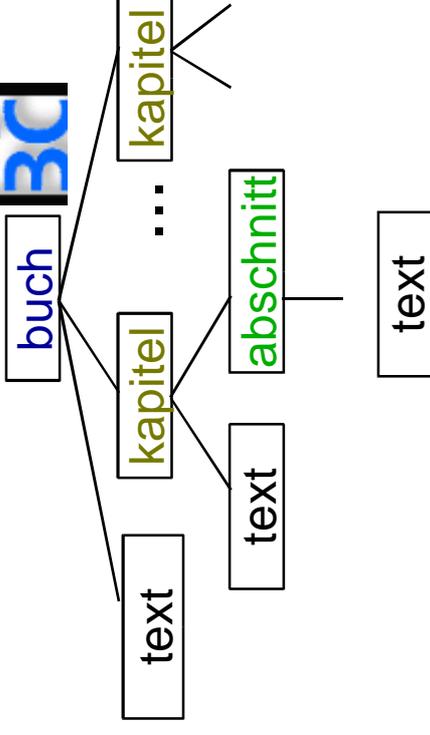
- Die Dokumenttyp-Deklaration `<!DOCTYPE ... >` informiert darüber, welche DTD für die folgende Instanz gültig ist und wo sie ist (externe DTD) oder wie sie aussieht (interne DTD)
- Die **DTD** (Document Type Definition) definiert,
 - welche Elementtypen es gibt und
 - welchen Inhalt sie haben dürfen,
 - welche Attribute sie haben und welche Werte sie annehmen dürfen
 - welche Entities (Referenzen, Variablen) es gibt
 - welche Notationen (Formate) es gibt

DTD

```
<?xml version="1.0"?>  
<!DOCTYPE buch SYSTEM "buch.dtd">
```



```
<!-- buch.dtd -->  
<!ELEMENT buch (text,kapitel+)>  
<!ELEMENT kapitel (text |  
                    abschnitt*)>  
<!ELEMENT abschnitt (text)>  
<!ELEMENT text (#PCDATA)>  
  
<!NOTATION MPEG SYSTEM  
    "/usr/local/bin/mpeg_player">  
  
<!ENTITY musik SYSTEM  
    "../urlaub.mpeg"  
    NDATA MPEG >
```



Sonderbedeutungen

Default: genau einmal
?: einmal oder kein mal
+: mindestens einmal
*: beliebig oft

Datentypen in XML

#PCDATA (parsed character data)

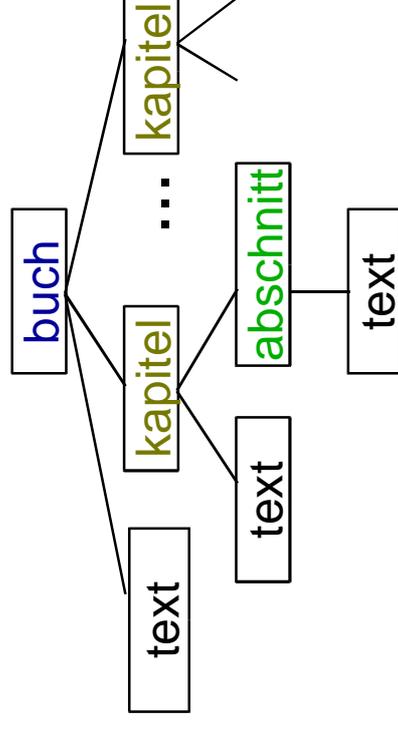
= reiner Text

NDATA (non-XML data)

= fremde Notation (Format)

Die Dokument-Instanz

```
<buch>
XML in 2 Schritten
<kapitel>
  Der erste Schritt
  <abschnitt>
    Am Anfang gab es nur Worte,
    erst später kam die Struktur.
  </abschnitt>
  <abschnitt>
    Man entspannt am einfachsten
    bei folgendem Stück:
    &musik;
  </abschnitt>
</kapitel>
</buch>
```

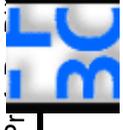


xml

- ist eine Metasprache (eine Sprache zur Beschreibung von Sprachen)
- ermöglicht den interaktiven Austausch von Dokumenten und Formularen, z.B. für Recherche-Anfragen, Auskünfte etc.
- ist eine Verallgemeinerung von HTML, der Standardsprache des Internets (HTML ist eine spezielle DTD)
- basiert auf SGML, dem Standard für Auszeichnungssprachen, ISO 8879 (daher auch der Name: Markups oder Auszeichnungen sind die Anweisungen, die die Lektoren für die Setzer in den Druckereien an den Texttrand schreiben)

XML – eine Familie von Technologien

Pr
mei



XLink: Standard für Hyperlinks und andere Beziehungen zwischen XML–Dokumenten (in Entwicklung)

XPointer & XFragment: Sprachen, um auf Teile von XML–Dokumenten zeigen. (XPointer ist etwa wie eine URL–Adresse, nur dass er auf Dokument–Teile verweist)

XSL (*Extensible Stylesheet Language*): Sprache für *Stylesheets* (Formatvorlagen)

DOM (*Document Object Model*): Standardmenge von Klassen und Methoden, um XML (und HTML) Dokumente durch Programmiersprachen wie Java, ECMAScript, VBScript oder C++ zu manipulieren

Einsatzmöglichkeiten (Auswahl)

Pr
mei



- **www (IE versteht Teilmenge von xml und xsl, Netscape wird es auch können)**
- **Austausch strukturierter Daten**
 - **Marktplätzen**
 - **Kunden Lieferanten**
- **Austausch von von Office–Anwendungen erzeugter Daten**
- **Konfigurationsdateien**