



Praktikum zu XML: Übung 02

XML-Deklaration,
Benutzung einer externen DTD,
XML-Modus des Emacs (!),
Definition und Verwendung von *entities*



Organisatorisches



- Arbeitsverzeichnis:
`~/kurse/xml/02/`
- Dateinamen:
`02-unicode.xml`
`02-unicode.dtd`
`02-abbrevs.ent`
- Werkzeuge:
`emacs` # NICHT X-Emacs
`nsgmls` # implizit via Emacs
`mozilla` # zur Zeichenkontrolle



- A: Einbau einer externen DTD
 - Legen Sie **02-unicode.xml** an durch Kopieren Ihres Ergebnisses aus Aufgabe 01:
`cp ../01/01-unicode.xml ../02-unicode.xml`
 - Legen Sie **02-unicode.dtd** an durch Kopieren aus folgender Quelle:
`/local0/werntges/kurse/xml/
02/02-unicode.dtd`
 - Legen Sie eine leere Datei **02-abbrevs.ent** an:
`touch 02-abbrevs.ent`



- A: Einbau einer externen DTD (Forts.)
 - Fügen Sie nun per Editor folgende Zeile am Anfang von 02-unicode.xml ein:

```
<!DOCTYPE Codetabelle SYSTEM  
  "02-unicode.dtd" [  
]>
```
 - Sichern und schließen Sie die XML-Datei, laden Sie die Datei dann neu mit **Emacs**. Nun sollte die DTD beachtet werden!
 - Öffnen Sie **02-unicode.xml** ferner mit **mozilla**. Es sollten keine Fehlermeldungen auftreten!



- B: Der XML-Modus des Emacs
 - Fügen Sie per Editor einen neuen Tabelleneintrag ein
 - Gewünschtes Zeichen: Kleines griechische „delta“. Ermitteln Sie zunächst Unicode und ISO-Code/TabNr dazu.
 - Positionieren Sie nun den Cursor an die richtige Stelle.
 - Verwenden Sie dann Menüpunkt **Markup/Insert Element**
 - Wählen Sie das gewünschte Element „Eintrag“ aus.
 - Ergänzen Sie den eingefügten Text um Nutzdaten.
 - Ergänzen Sie analog das Element „ISO-Code“ im neuen Eintrag.
 - Validieren Sie nun Ihre Datei über **SGML/Validate!**
 - Es erscheint unten ein Fenster mit Warnungen und evtl. Fehlermeldungen. Ignorieren Sie die „üblichen“ Warnungen (vgl. mündliche Erläuterungen), beseitigen Sie gemeldete Fehler.
 - Tip: Das untere Fenster läßt sich ausblenden mit CTRL-X 0 (Null), wenn der Cursor in diesem Fenster steht (aktives Fenster).



- C: Einbau der XML-Deklaration; Zeichensätze.
 - Fügen Sie folgende Zeile am Anfang von 02-unicode.xml ein:

```
<?xml version="1.0"?>
```
 - Validieren Sie die Datei mit Emacs und mozilla.
 - Geben Sie nun passende Umlaute per Tastatur „direkt“ in Element „Beschreibung“ zu den Umlauten ä und Ä ein.
 - Validieren Sie erneut. Ergebnis?
 - Ergänzen Sie nun die *encoding declaration*:

```
<?xml version="1.0" encoding="UTF-8"?>
```
 - Validieren Sie erneut. Begründen Sie Ihre Beobachtung!



- C: Einbau der XML-Deklaration; Zeichensätze.
 - Ermitteln Sie einen geeigneten Zeichensatz und tragen Sie diesen ein:

```
<?xml version="1.0"
  encoding="(Ihr-Zeichensatzname)"?>
```
 - Validieren Sie erneut und passen Sie ggf. den Zeichensatznamen an, bis die Validierung gelingt und mozilla die Zeichen korrekt anzeigt.
 - Hinweis: Nutzen Sie das Vorlesungsmaterial bzw. die XML-Spezifikation zur Ermittlung eines geeigneten Zeichensatznamens.



- D: Einbau eines internen *general entity*.
 - Ergänzen Sie zu Beginn der internen DTD folgende Zeilen:

```
<!ENTITY ae "a-Umlaut">
<!ENTITY Ae "A-Umlaut">
```
 - Wenden Sie die *entity*-Referenzen im Element „Beschreibung“ von „ä“ und „Ä“ an, etwa so:

```
<Beschreibung>ae, &ae;</Beschreibung>
```
 - Validieren Sie mit Emacs/nsxml. Ok?
 - Validieren Sie nun mit mozilla. Ergebnis?
 - So weit die Vorbereitungen ...



Aufgaben, mit Arbeitsanweisungen



- D: Einbau eines externen *parameter entity*.
 - Ergänzen Sie zu Beginn der internen DTD folgende Zeilen:

```
<!ENTITY % abbrev SYSTEM "02-abbrevs.ent">
%abbrev;
```
 - Validieren Sie mit Emacs/nsgmls. Ok?
 - Validieren Sie nun mit mozilla. Ergebnis?
 - Vertauschen Sie nun die entities: Zuerst die Deklarationen der beiden *general entities*, dann die des *parameter entity* mitsamt Referenz.
 - Validieren Sie nun erneut mit Emacs/nsgmls und mit mozilla. Hat sich das Ergebnis verändert?



Aufgaben, mit Arbeitsanweisungen



- E: Nutzung vordefinierter externer *parameter entities*.
 - Tragen Sie in **02-abbrevs.ent** folgende Zeilen ein:

```
<!ENTITY % iso-lat1 SYSTEM
"/usr/share/sgml/entities/xml-iso-entities-
8879.1986/ISOlat1.ent">
<!ENTITY % iso-grk1 SYSTEM
"/usr/share/sgml/entities/xml-iso-entities-
8879.1986/ISOgrk1.ent">
%iso-lat1;
%iso-grk1;
```
 - Schauen Sie in diese Dateien hinein mit einem Tool Ihrer Wahl und suchen Sie sich je ein *entity* aus, etwa „Auml“ und „agr“ (aber nicht gerade diese!). Laden Sie dann die XML-Datei neu.
 - Verwenden Sie Ihre beiden entities an geeigneten Stellen, d.h. setzen Sie die entsprechenden *entity references* mittels Emacs-Menüpunkt „Markup/Insert Entity“ und validieren Sie mit Emacs/nsgmls.
 - Prüfen Sie schließlich mit mozilla. Funktioniert's ebenfalls?



- Die Verwendung von *entities* ist weitverbreitet und sollte bis zur sicheren Beherrschung geübt werden.
- Manche XML-Prozessoren halten sich nicht immer an den Standard oder sind schlicht noch fehlerhaft. Im Zweifelsfall sollte man daher mit mehreren Werkzeugen prüfen.
- Der „Umweg“ über Datei 02-abbrevs.ent ist nicht generell erforderlich, sondern veranschaulicht die Kaskadierbarkeit von *entities*.
- Die ISO-*Entities* sind ein weitverbreiteter Quasistandard.
- Wir werden noch eine bessere Methode zu ihrer Adressierung kennenlernen als eine direkte lokale Pfadangabe: PUBLIC *entities* und *catalogs*!