

Übung 10: XML Schema

- Wir nähern uns den zahlreichen Möglichkeiten, die W3C XML Schema ermöglicht, in mehreren Schritten:
 - Teil A: Umstellung einer DTD auf XML Schema, mit einem vereinfachten XML-Dokument
 - Teil B: Definition eigener, angepasster Datentypen
 - Teil C: Modularisierung, Verwendung globaler Attribute
 - Teil D: Vervollständigung, Umstellung aller DTD-Elemente
- Grundlage:
 - Dateien aus Übung 7 (Klausurverwaltungsbeispiel)
- Vorbereitung:

```
cp ueb07.xml ueb10a.xml # Bereitgestellte
cp ueb07.dtd ueb10a.xsd # Quellen kopieren
```
- Aufruf des Schema-Validierers (Beispiel):

```
/usr/local/xerces-c-2.1.0/bin/SAXPrint -n -s -v=always ueb10a.xml
```

Übung 10: XML Schema

- **Aufgabe 10a:**
 - „Schema-validieren“ Sie ueb10a.xml
 - Vereinfachen Sie zunächst ueb10a.xml durch Entfernen / Auskommentieren aller Unterelemente von „Klausur“ außer „Aufgabenliste“.
 - Kommentieren Sie die DTD-Inhalte in ueb10a.xsd aus. Nicht löschen, damit Sie leicht nachschlagen können, was umzustellen ist.
 - Ändern Sie den Kopfbereich von ueb10a.xml nun so, dass die Schemadatei „ueb10a.xsd“ von SAXPrint gefunden wird.
 - Bauen Sie einen Schema-Rumpf in ueb10a.xsd ein, so dass der Parser das Schema zu verarbeiten beginnt.
 - Bauen Sie dann die benötigten Elementtyp- und Attributtyp-Deklarationen ein und testen Sie alle Zwischenschritte mit dem Parser.
 - Hinweis:
 - Wenn die Namespace-Deklarationen stimmen, ist schon viel erreicht!
 - Versuchen Sie in Teil A nur eine direkte Umsetzung der Deklarationen aus der DTD nach XML Schema, gemäß der Vorlesungsbeispiele.

Übung 10: XML Schema

- **Aufgabe 10b:**
 - Führen Sie angepasste eigene Datentypen ein
 - Kopien anlegen:
 - cp ueb10a.xml ueb10b.xml; cp ueb10a.xsd ueb10b.xsd
 - Header in ueb10b.xml anpassen (ueb10b.xsd verwenden)
 - Header in ueb10b.xsd anpassen (*namespaces!*)
 - Definieren und ggf. referenzieren Sie folgende Datentypen
 - MaxPtType Wertebereich 0 ... 100, für Attribut „MaxPunkte“
Für „Experten“: Lassen Sie auch halbe Punkte zu!
 - KursIdType Beschränkung auf 4-stellige Zahlen
 - SemesterType WS2xxx, SS2xxx
 - Schema-validieren Sie Ihr Ergebnis

Übung 10: XML Schema

- **Aufgabe 10c:** Modularisieren Sie das Schema
 - Kopien anlegen:
 - cp ueb10b.xml ueb10c.xml; cp ueb10b.xsd ueb10c.xsd
 - Header in ueb10c.xml und ueb10c.xsd anpassen (analog 10b)
 - Element „Teilaufgabe“:
 - Dieses Element tritt implizit nochmal innerhalb von Element „Aufgabe“ auf. Definieren Sie es nur einmal und referenzieren Sie es dann mehrfach, um die Redundanz zu beseitigen. Denken Sie auch an das Attribut.
 - Hinweis: Wenn nötig, bilden Sie eine Gruppe (<group>)
 - Attribut „MaxPunkte“:
 - Dieses Attribut wird in gleicher Art von mehreren Elementen benötigt. Versuchen Sie es als globales Attribut anzulegen und per Referenz anzusprechen.
 - Hinweis: Wenn nötig, bilden Sie eine Gruppe (<attributeGroup>)
 - Schema-validieren Sie Ihr Ergebnis

Übung 10: XML Schema

- **Aufgabe 10d:** Vervollständigen Sie das Schema
 - Kopien anlegen:
 - cp ueb10c.xml ueb10d.xml; cp ueb10c.xsd ueb10d.xsd
 - Header in ueb10d.xml und ueb10d.xsd anpassen (analog 10b), ggf. auskommentierte Daten wieder ent-kommentieren.
 - Elemente „Teilnehmerliste“, „Ergebnisliste“, „Notenschwellen“
 - Nachdem nun die wesentlichen Arbeitsschritte erprobt worden sind, übertragen Sie das Verfahren auf die anderen Elemente.
 - Ziel: Alle DTD-Bestandteile finden Ihre Fortsetzung - oder oftmals: Verbesserung - im Schema.
 - Hinweis:
 - An zwei Stellen benötigen Sie eine Ableitung eines komplexen Typs aus einem einfachen Datentyp *by extension*.
Aufgabe: Ermitteln Sie die Einzelheiten dieses nicht vorgestellten Konstrukts eigenständig durch Suche in den Spezifikationen!
Tip: Im Tutorial gibt es ein Beispiel dazu.
 - b.w.

Übung 10: XML Schema

- **Aufgabe 10d:** (Forts.)
 - Strenge Typisierung
 - Schränken Sie Wertebereiche sowohl von Attributen als auch die Werte (und Wiederholfaktoren) von Elementen so präzise wie möglich ein.
 - Schema-validieren Sie Ihr Ergebnis
 - Verwenden Sie den Schema-Validierer oftmals zwischendurch, um systematische Fehler frühzeitig zu erkennen.
 - Prüfung der Typisierungsangaben
 - Nach schließlich gelungener Validierung prüfen Sie, ob die strenge Typendefinition auch vom Parser beachtet wird.
 - Provozieren Sie dazu den Parser durch vorübergehend „falsche“ Werte in geeigneten Stellen von ueb10d.xml. Erhalten Sie stets die erwarteten Fehlermeldungen?
 - Abschluß
 - Entfernen Sie schließlich bei Erfolg die auskommentierte und nun nicht mehr benötigte DTD in der XSD-Datei und validieren Sie ein letztes Mal.

Übung 11: XML Information Set

- **Aufgabe:**
 - Erstellen Sie (auf Papier oder in einer Datei) die Beschreibung eines XML Infoset zu ueb04.xml
 - A) Nehmen Sie an, dass der Parser die externe *entity* nicht auflöst.
 - B) Unterstellen Sie einen validierenden Parser.
- **Anmerkungen**
 - Zweck der Aufgabe ist das Vertrautwerden mit dem XML Infoset durch eigene Anwendung des Vorlesungsmaterials.
 - Orientieren Sie sich am Beispiel aus der Vorlesung.
 - Im Unterschied zu diesem liegen hier keine expliziten namespace-Angaben vor, dafür aber externe *entities*.
 - Dies ist keine Fleißaufgabe:
 - Vielfalt ist Trumpf: Konzentrieren Sie Ihre Zeit auf die Abdeckung der unterschiedlichen Fälle, die auftreten können.
 - Lassen Sie sich wiederholende Angaben einfach aus.

Übung 12: XPath

- **Vorbemerkungen**
 - Grundlage
 - ueb07.xml, also unser Klausurverwaltungsbeispiel
 - Ziele
 - Vorübung: Vertrautwerden mit dem XPath-Datenmodell
 - Teil A: Passives Verstehen von XPath-Ausdrücken
 - Teil B: Aktive Fähigkeit zur Lösung von Suchaufgaben durch Erzeugung geeigneter XPath-Ausdrücke
 - Vorgehen
 - Die Teile A & B bestehen aus Abfolgen von Einzelaufgaben
 - Die Teilaufgaben werden tendenziell komplexer. Bearbeiten Sie diese daher am besten in der angegebenen Reihenfolge.
 - Allgemeines
 - Dies ist keine Rechnerübung - Papier und Bleistift reichen aus.

Übung 12: XPath

- Vorübung
 - Ermitteln Sie die XPath Dokumentenbaum-Darstellung des XML-Beispieldokuments zu XML Infoset („Phone home!“).
- Hinweise:
 - Beachten Sie insbesondere die Unterschiede zu Infoset: Was fehlt ganz, was wird vereinfacht, was ist analog?
 - Die Übung soll Ihnen Gelegenheit geben, über die Datenmodelle am konkreten Beispiel zu reflektieren. Es kommt nicht auf Antwort-Details an, sondern darauf, eventuell noch vorhandene Verständnisfragen zu provozieren.
 - Sollten Sie bei der Bearbeitung auf Verständnisfragen stoßen, zögern Sie nicht, Ihren Dozenten darauf anzusprechen.

Übung 12: XPath

- A: XPath-Ausdrücke verstehen
Was ist der Rückgabewert
(Typ, konkrete Angabe oder Beschreibung) von:
 - 1) `/Klausur/Aufgabenliste/Aufgabe`
 - 2) `descendant::/Klausur/*[@Note="1.3"]`
 - 3) `string(/Klausur//*[@Note="1.3"]/text())`
 - 4) `string(/Klausur//@MatrNr("123456")) [1]`

Übung 12: XPath

- B: XPath-Ausdrücke verfassen

Ermitteln Sie einen XPath-Ausdruck, der folgende Rückgaben liefert:

- 1) Alle „Teilaufgabe“-Knoten aller „Aufgabe“-Knoten
- 2) Die zweite Teilaufgabe der Aufgabe 1
- 3) Die Summe der max. Punkte aller Teilaufgaben der vorletzten Aufgabe
- 4) Alle Knoten, die auf Matrikelnummer „123456“ Bezug nehmen.